



Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra



Automata and finite order elements in the Nottingham group

Jakub Byszewski^a, Gunther Cornelissen^{b,*}, Djurre Tijsma^c

^a Wydział Matematyki i Informatyki, Uniwersytet Jagielloński, ul. S. Łojasiewicza 6, Kraków, 30-348, Poland

^b Mathematisch Instituut, Universiteit Utrecht, Postbus 80.010, Utrecht, 3508 TA, The Netherlands

^c Mathematisches Institut der Heinrich-Heine-Universität, Universitätsstraße 1, Düsseldorf, 40225, Germany

ARTICLE INFO

Article history:

Received 1 October 2020

Available online 1 April 2022

Communicated by Kirsten

Eisenträger

MSC:

11B85

11-04

11G20

11S31

11Y16

20E18

20E45

68Q70

Keywords:

Nottingham group

Power series over finite fields

Automata theory

ABSTRACT

The Nottingham group at 2 is the group of (formal) power series $t + a_2t^2 + a_3t^3 + \dots$ in the variable t with coefficients a_i from the field with two elements, where the group operation is given by composition of power series. The depth of such a series is the largest $d \geq 1$ for which $a_2 = \dots = a_d = 0$.

Only a handful of power series of finite order (forcedly a power of 2) are explicitly known through a formula for their coefficients. We argue in this paper that it is advantageous to describe such series in closed computational form through automata, based on effective versions of proofs of Christol's theorem identifying algebraic and automatic series.

Up to conjugation, there are only finitely many series σ of order 2^n with fixed break sequence (i.e. the sequence of depths of σ^{o2^i}). Starting from Witt vector or Carlitz module constructions, we give an explicit automaton-theoretic description of: (a) representatives up to conjugation for all series of order 4 with break sequence $(1, m)$ for $m < 10$; (b) representatives up to conjugation for all series of order 8 with minimal break sequence $(1, 3, 11)$; and (c) an embedding of the Klein four-group into the Nottingham group at 2.

We study the complexity of the new examples from the algebraic-geometric properties of the equations they satisfy. For

* Corresponding author.

E-mail addresses: jakub.byszewski@gmail.com (J. Byszewski), g.cornelissen@uu.nl (G. Cornelissen), tijsma@uni-duesseldorf.de (D. Tijsma).

this, we generalise the theory of sparseness of power series to a four-step hierarchy of complexity, for which we give both Galois-theoretic and combinatorial descriptions. We identify where our different series fit into this hierarchy. We construct sparse representatives for the conjugacy class of elements of order two and depth $2^\mu \pm 1$ ($\mu \geq 1$). Series with small state complexity can end up high in the hierarchy. This is true, for example, for a new automaton we found, representing a series of order 4 with 5 states (the minimal possible number for such a series).

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Suppose $\sigma(t) = t + a_2t^2 + a_3t^3 + a_4t^4 + \cdots \neq t$ is a formal power series in the variable t with coefficients from the field $\mathbf{F}_2 = \mathbf{Z}/2\mathbf{Z}$ with two elements. Since $\sigma(t) = t + O(t^2)$, substituting $\sigma(t)$ into itself produces a power series $\sigma^{\circ 2}(t) = t + a_2(a_3 + 1)t^4 + \cdots$, and one may iterate this process to arrive at $\sigma^{\circ N}(t) := \sigma(\sigma(\cdots \sigma(t)))$. (We will systematically write $\sigma^{\circ N}(t)$ for the N -fold composition, and $\sigma(t)^N$ for the N -th power of the power series $\sigma(t)$; so here, for example, $\sigma(t)^2 = t^2 + a_2t^4 + \cdots$.) Our concern is *the explicit description of σ and N for which $\sigma^{\circ N}(t) = t$* (this is only possible if N is a power of 2). Our goal is not to compute finitely many coefficients a_i of such $\sigma(t)$, but rather to give a *finite description* of the complete series. To accomplish this, one might search for explicit formulas for the general coefficient a_i or for the set

$$E(\sigma) := \{i \in \mathbf{Z}_{\geq 0} : a_i \neq 0\}$$

of occurring exponents, and this has been done in a few cases. In this paper, we will argue that one may push the boundaries of what is currently feasible by describing the coefficients of the power series by means of a finite automaton (that such a description is possible was already pointed out in [9, Rem. 1.5]). We will construct the automaton using symbolic computation, based on Christol's characterisation of algebraic power series by automata [23,24]. We wish to stress that an automaton is a perfectly deterministic finite description of the corresponding power series $\sigma(t)$, but that a very small automaton (i.e. with very few states) may correspond to a power series for which an elementary description of the set $E(\sigma)$ is very complex. If one is interested in just the computation of the k -th coefficient of the power series $\sigma(t)$, the automaton can be used to do this in time logarithmic in k .

We will first review the mathematical relevance of this problem. Then we describe existing results and explain our method. Since the same question makes sense for the finite field \mathbf{F}_p with p elements (where p is prime, and then forcedly N is a power of p), we will consider this more general problem in the theoretical parts of the paper.

1.1. Connections

Fixing a prime number p , the *Nottingham group* $\mathcal{N}(\mathbf{F}_p)$ is the pro- p -Sylow subgroup of the group of ring automorphisms $\text{Aut}(\mathbf{F}_p[[t]])$ of the formal power series ring $\mathbf{F}_p[[t]]$ over the finite field \mathbf{F}_p , with composition as multiplication. There is a group isomorphism $\text{Aut}(\mathbf{F}_p[[t]]) \cong \mathcal{N}(\mathbf{F}_p) \rtimes \mathbf{F}_p^*$. A ring endomorphism σ of $\mathbf{F}_p[[t]]$ is determined uniquely by the image $\sigma(t) \in t\mathbf{F}_p[[t]]$ of t , and $\mathcal{N}(\mathbf{F}_p)$ is identified with the group of power series $\sigma(t) \in \mathbf{F}_p[[t]]$ with $\sigma(t) = t + O(t^2)$ under composition. We write $\sigma \circ \tau$ for the result of substituting the series $\tau \in \mathcal{N}(\mathbf{F}_p)$ for the variable t in $\sigma \in \mathcal{N}(\mathbf{F}_p)$. The Nottingham group arises in many areas:

- In *group theory*, as Ershov remarked in [32], $\mathcal{N}(\mathbf{F}_p)$ is ‘an excellent test example for many questions or conjectures in profinite group theory that have been settled for Chevalley groups’. In that reference, he proved that for $p \geq 5$, $\mathcal{N}(\mathbf{F}_p)$ admits no open embedding into a topologically simple group. On the other hand, every countably based pro- p group embeds into $\mathcal{N}(\mathbf{F}_p)$ (Camina [20]; Jennings [44]); in particular, every finite p -group embeds into $\mathcal{N}(\mathbf{F}_p)$ (an older unpublished result of Leedham-Green and Weiss; see [20, Thm. 3]).
- In *number theory*, the Nottingham group occurs naturally in the theory of wild ramification (as the group of wild automorphisms of $\mathbf{F}_p((t))$; see Fesenko [33]).
- The previous point relates to *algebraic geometry*, namely: if a group G acts on a smooth projective curve X over \mathbf{F}_p , then the stabiliser G_x of a point $x \in X$ acts on the completion of the local ring $\mathcal{O}_{X,x}$. This completion is isomorphic to $\mathbf{F}_p[[t]]$, leading to an embedding of the wild ramification group G_x^1 (the p -Sylow subgroup of G_x) into $\mathcal{N}(\mathbf{F}_p)$; one can, for example, study deformations of group actions on curves through deformations of this group homomorphism, much like deformations of linear group representations, e.g. of Galois groups, cf. [56].

The need for explicit representations of finite order elements in $\mathcal{N}(\mathbf{F}_p)$ has been articulated several times, both in group theory ([21, p. 216], [54, §5.4]), as well as in deformation theory, where conclusive results about formal deformation spaces and/or lifting are only known when standard forms for the series are available [8,15,28,30,16,36].

Our results are also relevant for the *theory of automata* (that it relies upon), in particular, issues of implementation of certain algorithms for solving algebraic equations (Section 3, e.g. [14]), the enumeration of automata with specific properties (cf. Section 4), and an extension of Cobham’s theory of complexity of automata/regular languages (cf. Section 10).

1.2. Review of previous work

Klopsch has proven that every element of order p in $\mathcal{N}(\mathbf{F}_p)$ is conjugate to

$$t/\sqrt[p]{1 - mat^m} = t + at^{m+1} + \dots \quad (1)$$

for some positive integer m coprime to p and $a \in \mathbf{F}_p^*$, and that these series are mutually not conjugate [48]. The expression (1) may be readily converted into a formula for the coefficients of the corresponding power series by applying the binomial expansion (see also the discussion in Example 1.3.1).

Jean [43] and Lubin [54] indicated how to use formal groups and explicit local class field theory to describe elements of any order p^n in $\mathcal{N}(\mathbf{F}_p)$, and iterative procedures for the calculation of the coefficients of such elements were described (compare [42], [47], [10, §6]). However, the only known formulas for elements of order p^n for $n > 1$ are for $p^n = 4$ in $\mathcal{N}(\mathbf{F}_2)$, given by Jean in [42, Ch. 7], Chinburg and Symonds [22], and Scherr and Zieve (cf. [9, Rem. 1.4]). The Chinburg–Symonds example represents the action of an automorphism of order 4 on the local completed ring at zero of the supersingular elliptic curve over \mathbf{F}_2 ; compare also [9, Sect. 1], where it is argued that this is essentially the only example that can be constructed by such a method; more precisely, up to conjugation, it is the only ‘almost rational’ example. The final section of [42] contains another (implicit) way of describing a solution to the problem, this time by using the method of Mellin [57] to solve algebraic equations—in this case, a trinomial—using hypergeometric series (the historically not entirely accurate reference in [57] is to a monograph by Belardinelli).

The *break sequence* of $\sigma \in \mathcal{N}(\mathbf{F}_p)$ of order p^n is a refined invariant with the property that there are only finitely many conjugacy classes of elements of fixed order p^n with a given break sequence. The method of Lubin [54] can in principle be used to count that number using results from local class field theory. There is an exact characterisation of possible break sequences [54, Obs. 5]. We briefly recall the definitions.

Definition 1.2.1. The *depth* of $\sigma = \sigma(t) \in \mathcal{N}(\mathbf{F}_p)$ is $d(\sigma) := \text{ord}_t(\sigma(t) - t) - 1$ (and $d(t) = \infty$), so if $\sigma(t) = t + a_k t^k + O(t^{k+1})$ with $a_k \neq 0$, then $d(\sigma) = k - 1$. The *lower break sequence* of an element $\sigma \in \mathcal{N}(\mathbf{F}_p)$ of finite order p^n is defined as

$$\mathbf{b}_\sigma = (b_i)_{i=0}^{n-1} = (d(\sigma^{\circ p^i}))_{i=0}^{n-1}.$$

The data \mathbf{b}_σ correspond bijectively to the so-called *upper break sequence* $\mathbf{b}^\sigma = \langle b^{(i)} \rangle_{i=0}^{n-1}$ that we will not define; for our purposes, it suffices to quote from [54, Def. 4] the formula that converts between lower and upper break sequences, which in our case of the cyclic group generated by σ becomes

$$b^{(0)} = b_0 \quad \text{and} \quad b^{(i)} = b^{(i-1)} + p^{-i}(b_i - b_{i-1}) \quad \text{for } i > 0. \quad (2)$$

We will always indicate lower sequences by $(\)$ -brackets, and the corresponding upper sequences by $\langle \ \rangle$ -brackets, and we will write $(b_i) = \langle b^{(i)} \rangle$ for corresponding lower and upper break sequences.

1.3. The method of construction

We will use the term *p-automaton* to describe a finite directed multigraph (allowing loops, as well as multiple edges between vertices) for which:

- vertices are labelled by elements of \mathbf{F}_p [‘output alphabet \mathbf{F}_p ’];
- one vertex (the so-called *start vertex*) is additionally marked ‘Start’;
- each vertex has exactly p outgoing edges, each labelled by a different element of the set $\{0, 1, \dots, p-1\}$; [‘input alphabet $\{0, 1, \dots, p-1\}$ ’]
- there is a path in the automaton from the start vertex to any vertex [‘accessibility’];
- an edge with label 0 always connects two vertices with the same label [‘leading zeros invariance’].

In the general theory of automata, this is called a ‘leading zeros invariant p -DFAO (deterministic finite p -automaton with output) with output alphabet \mathbf{F}_p and all states accessible’. Vertices are also called ‘states’. We omit the qualifier p when it is clear from the context.

Such an automaton produces the so-called *p-automatic sequence* $(a_k)_{k \geq 0}$, where a_k is the label carried by the final vertex of the walk that starts at the start vertex and follows the edges according to the successive digits of k in base p (starting from the least significant digit, also called the ‘reverse/backwards reading convention’, compare [5, 12.2]). The sequence $(a_k)_{k \geq 0}$ gives rise to the corresponding formal power series $\sum a_k t^k$ over \mathbf{F}_p in the variable t . Note that the ‘leading zeros invariance’ property means that we can allow the base- p expansion of k to have any number of leading zeros without affecting the resulting sequence. Should an automaton contain inaccessible vertices, they may be removed together with all their connecting edges without changing the corresponding series.

Example 1.3.1. We consider Klopsch’s series

$$\sigma_{K,3} := t/\sqrt[3]{1+t^3} = \sum_{k \geq 0} a_{3k+1} t^{3k+1} = t + t^4 + t^{13} + \dots \in \mathcal{N}(\mathbf{F}_2)$$

of order 2 with lower break sequence (3). The coefficients of this series can be described explicitly: a_{3k+1} is equal to the binomial coefficient $\binom{-1/3}{k}$ modulo 2. Writing $-1/3$ as a 2-adic integer $-1/3 = \sum_{k \geq 0} 4^k$, we get an infinite product representation

$$\sigma_{K,3} = t \prod_{k \geq 0} (1 + t^{3 \cdot 4^k}),$$

which shows that $a_k = 1$ if and only if the base-4 expansion of $k-1$ contains only the digits 0 or 3. An automaton corresponding to this series is depicted in Fig. 1; one way

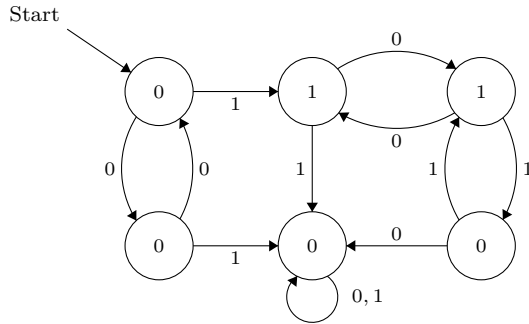


Fig. 1. A 2-automaton representing Klopsch's series $\sigma_{K,3} \in \mathcal{N}(\mathbf{F}_2)$ of order 2 with lower break sequence (3).

to construct it is to solve the algebraic equation $(t^3 + 1)\sigma^3 = t^3$ with initial coefficients $\sigma = t + t^4 + O(t^5)$ using one of the algorithms in Section 2 below.

To illustrate our reading conventions, we compute the coefficient a_{13} of the corresponding power series: write $13 = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$ in base 2 as 1101; begin at the start vertex and follow the directed edges with respective labels 1, 0, 1, 1; we end up in a vertex with label 1, so $a_{13} = 1$. (If one adds leading zeros, e.g. by writing $13 = 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$, the result is the same even though the final vertex might be different.)

Our construction of elements of order p^n in $\mathcal{N}(\mathbf{F}_p)$ proceeds as follows:

- (i) Use Witt vectors to construct a cyclic Galois extension of order p^n of the field of Laurent series $\mathbf{F}_p((z))$ with certain ramification behaviour (this is similar to the method employed by Leedham-Green and Weiss, see [20, Thm. 3]; for a discussion using class field theoretic methods instead, see Remark 2.1.2). This field extension is described in terms of a finite set of generators α_i satisfying a set of explicit algebraic relations over $\mathbf{F}_p((z))$ and with explicit formulas for the action of a generator σ of the Galois group on the variables α_i . Moreover, one can choose this field extension in such a way that α_i are algebraic over the field of rational functions $\mathbf{F}_p(z)$, so all computations involve algebraic functions only (cf. Examples 2.2.2 & 2.2.3).
- (ii) Choose a rational function in the variables α_i that is a uniformiser for the field extension, say t . One can consider σ as an automorphism of $\mathbf{F}_p((t))$, and one has an explicit expression for $\sigma(t)$ as a rational function of the variables α_i . This leads to a set of algebraic equations involving $\sigma(t)$, t and α_i (note that 'algebraic' is w.r.t. the usual addition and multiplication of power series, not composition). By elimination of the variables α_i from those equations (in general with the help of a Groebner basis algorithm), one finds an explicit equation $F(t, X) = 0$ for $\sigma = \sigma(t)$ over the field $\mathbf{F}_p(t)$.
- (iii) Use an algorithmic version of a proof of Christol's theorem (based on using Ore polynomials, Furstenberg's diagonal method, or differential forms on algebraic curves) to find automata whose series correspond to the solutions of the equa-

tion $F(t, X) = 0$ in $\mathbf{F}_p[[t]]$. By Hensel's Lemma, sufficiently many initial coefficients of a solution will determine such a solution uniquely, so different solutions can be distinguished by solving iteratively for enough coefficients of a putative power series solution.

- (iv) The equation found in (iii) might have several solutions, and at least one of these solutions is a power series of order p^n . Identify the solution(s) that correspond to elements of order p^n .

We describe the steps in some detail in the next section. In the first two steps, there are many possible choices of extensions and uniformisers, and hence there are many possible algebraic equations. The size of the resulting automaton depends heavily on the choices made in the first two steps of the method, and the minimal size of an automaton representing a power series can vary greatly in a conjugacy class (theoretical bounds depending on the equations can be found in Bridy [13]).

Once the equation is fixed, the third and fourth step in the construction (which replace the naive method of trying to solve the equation recursively for the coefficients of a putative power series solution) have been automated by Rowland (see [58] for the source code and [59] for the description) and partly in [14]; we have used these implementations to produce the automata.

1.4. Results

We start by describing the case of elements of order 4.

Theorem 1.4.1 (Corollary 5.1.2 & Propositions 3.4.1, 4.2.1, 5.2.1, 5.3.1). *The following is a complete list representing all possible elements of order 4 in $\mathcal{N}(\mathbf{F}_2)$ with break sequence $(1, m) = \langle 1, (m+1)/2 \rangle$ for all admissible values $m < 10$, up to conjugation in $\mathcal{N}(\mathbf{F}_2)$:*

- with break sequence $(1, 3) = \langle 1, 2 \rangle$: two (previously known) series σ_{CS} and $\sigma_{\text{CS}}^{\circ 3}$ given in Equations (12) & (13), with the corresponding automata displayed in Table 1. The series σ_{CS} is conjugate in $\mathcal{N}(\mathbf{F}_2)$ to a new series σ_{\min} described by the automaton in Fig. 2, which is the unique series of order 4 described by a 2-automaton with at most 5 states.
- with break sequence $(1, 5) = \langle 1, 3 \rangle$: a series $\sigma_{(1,5)}$ corresponding to the 13-state automaton displayed in Fig. 5.
- with break sequence $(1, 9) = \langle 1, 5 \rangle$: a series $\sigma_{(1,9)}$ with 110-state automaton described in Table 2.

In Section 4 we present an algorithm for finding, for fixed integers N and n , all minimal 2-automata representing an element of finite order 2^n in $\mathcal{N}(\mathbf{F}_2)$ with at most N states.

For some of the automata it is possible to extract a manageable *closed formula* for the power series. We will present eight such formulas for power series of order 4 with minimal break sequence, of which five are new: $\sigma_j^{\circ 3}$ displayed in Equations (16) & (17) and $\sigma_{T,1}, \sigma_{T,2}, \sigma_{T,3}$ and $\sigma_{T,4}$ in Table 3. Note that although it is easy to determine which of these are mutually conjugate, the conjugating power series itself may be hard to describe: as far as we know, it may be transcendental over $\mathbf{F}_2(t)$, and we are not aware of any criteria that guarantee the existence of an algebraic conjugating power series (but cf. Remark 10.2.4).

For order 8, we have the following result (for the notion of ‘minimal’ break sequence, see Example 2.4.3).

Theorem 1.4.2 (Propositions 7.1.1, 7.2.1 & 7.3.1). *Up to conjugation in $\mathcal{N}(\mathbf{F}_2)$, there are precisely 4 elements $\sigma_8, \sigma_8^{\circ 3}, \sigma_{8,2}, \sigma_{8,2}^{\circ 3}$ of order 8 with ‘minimal’ break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$ in $\mathcal{N}(\mathbf{F}_2)$, where σ_8 corresponds to the 320-state automaton given in Table 5, and $\sigma_{8,2}$ corresponds to the 926-state automaton described in 7.3.*

The automata are also stored in standard Mathematica form in [17].

Since every finite 2-group embeds in $\mathcal{N}(\mathbf{F}_2)$, Klopsch asked for a description of an embedding of the Klein four-group $V = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ in $\mathcal{N}(\mathbf{F}_2)$. We have the following result.

Theorem 1.4.3 (Propositions 8.1.2 & 8.2.1). *For every embedding of the Klein four-group V in the Nottingham group $\mathcal{N}(\mathbf{F}_2)$, some nontrivial element of V has depth at least 5. Furthermore, the series $\sigma_{V,1}$ and $\sigma_{V,2}$ corresponding to the automata depicted in Table 6 have break sequences (1) and (5) and exhibit an explicit embedding of two generators of the Klein four-group into $\mathcal{N}(\mathbf{F}_2)$.*

One notices in the examples that for fixed order and break sequence, some series with an explicit ‘easy’ formula are produced by a rather large automaton, while at the same time there exist series requiring fewer states for which an ‘easy’ formula does not seem to exist. We study this phenomenon in Section 10, generalising the concept of *sparseness*. Recall that a series $\sigma = \sum a_i t^i$ is in the class S of sparse series if the number of nonzero coefficients a_i with $i \leq N$ grows like a power of a logarithm of N . Klopsch’s series $\sigma_{K,m}$ are not sparse, but at least for some values of m their conjugacy class contains a sparse series.

Theorem 1.4.4 (Proposition 10.2.1). *Any power series of order 2 and depth $m = 2^\mu \pm 1$, $\mu \geq 1$, is conjugate to a sparse power series $\sigma_{S,m}$ given in Equations (22), (23) & (24), the first two of which correspond to the automata displayed in Table 8.*

We classify general series into three classes that we consider to have ‘easy formulas’:

$$S \subset \widehat{S} \subset \widehat{\widehat{S}} \subset \mathbf{F}_2[[t]],$$

where \widehat{S} is the class of series that are sparse up to multiplication with a rational function, and $\widehat{\widehat{S}}$ is the class of series that are in \widehat{S} up to composition with an automorphism of $\mathbf{F}_p(t)$. Whether or not a series is in a certain class can be studied both using Galois theory (Section 11) and combinatorics of automata (Section 12). Even for the ‘larger’ automata with several hundred states, the combinatorial method can be automated relatively easily using the computer algebra representation (cf. Table 11). Among the series described above there occur examples at all levels of this hierarchy of complexity.

Theorem 1.4.5 (*Theorem 11.2.6 & Table 9*). *The series $\sigma_{T,1}, \dots, \sigma_{T,4}, \sigma_{CS}^{\circ 3}$ are in S ; the series $\sigma_{CS}, \sigma_{CS}^{\circ 2}$ are in \widehat{S} but not in S ; the series $\sigma_J, \sigma_J^{\circ 3}$ are in $\widehat{\widehat{S}}$ but not in \widehat{S} ; the series $\sigma_{K,m}(m \geq 3), \sigma_{V,1}, \sigma_{V,2}, \sigma_{V,3}, \sigma_{\min}, \sigma_{(1,5)}, \sigma_{(1,9)}, \sigma_8$ are not in $\widehat{\widehat{S}}$.*

Finally, in Section 13 we briefly discuss the synchronisation properties of some of our automata, in relation to a ‘structured/random’ decomposition of automatic sequences in [19].

1.5. Some open problems

- We have provided one example of an embedding of a non-cyclic p -group (the Klein four-group V) into $\mathcal{N}(\mathbf{F}_p)$ (for $p = 2$), with the break sequences of the nontrivial elements of V being $(1), (1)$ and (5) . Study the possible break sequences for embeddings of V into $\mathcal{N}(\mathbf{F}_2)$, and more generally for embeddings of arbitrary (finite) p -groups into $\mathcal{N}(\mathbf{F}_p)$ (cf. Proposition 8.1.2 and Subsection 8.3 for some explicit challenges).
- Is there a sparse series of order 2 with break sequence (11) ? This is equivalent to asking whether Klopsch’s series $t/\sqrt[11]{1} + t^{11} \in \mathcal{N}(\mathbf{F}_2)$ is conjugate to a sparse series. More generally, is every element of finite order in $\mathcal{N}(\mathbf{F}_2)$ sparse (or in \widehat{S} or $\widehat{\widehat{S}}$) up to conjugation?
- Provide an automaton-theoretic characterisation of series that are sparse up to multiplication with a rational function, in a manner analogous to how [63] gives a necessary and sufficient condition for a series to be sparse in terms of properties of a corresponding automaton. This appears to be a hard problem, see Remark 12.2.3.
- As the automaton method allows us to extend the catalogue of known elements of finite order in $\mathcal{N}(\mathbf{F}_p)$, one may argue that it is advantageous to manipulate elements of finite order in $\mathcal{N}(\mathbf{F}_p)$ in their automatic form directly, ignoring any explicit form for the coefficients of the corresponding power series. Thus, it would make sense to study ‘ p -automata of finite order’ as a subject of its own. How to characterise an automaton that represents a series of finite order?
- If it exists, describe an algorithm that finds all automata on at most N states that represent series of finite order. For any *given* finite order this is easy (see Section 4), so an affirmative solution of this problem would most likely require finding a bound

on the order of a series in terms of the number of states of an automaton that generates it.

Notation. We will use the notation σ and $\sigma(t)$ for elements of $\mathcal{N}(\mathbf{F}_p)$ interchangeably, and also use σ for the corresponding element of the Galois group of an extension of fields of formal Laurent series. We will also write ‘ $\sigma(t)$ ’ when σ is considered as an element of a Galois group and t is a specified uniformiser.

2. Detailed method: finding an algebraic equation

2.1. Extensions of Laurent series fields and elements of $\mathcal{N}(\mathbf{F}_p)$

Let $k = \mathbf{F}_p((z))$ be a field of formal Laurent series with corresponding valuation v_z , and let K/k be a cyclic totally ramified Galois extension of degree p^n . Let t be a uniformiser for K with corresponding valuation v_t , so that $K = \mathbf{F}_p((t))$. Any $\sigma \in \text{Gal}(K/k)$ is an automorphism of $\mathbf{F}_p((t))$ fixing $\mathbf{F}_p((z))$, and it automatically preserves the valuation. It follows that $\sigma(t) = a_1 t + a_2 t^2 + a_3 t^3 + \cdots$ for some $a_i \in \mathbf{F}_p$; since the order of σ is a power of p , we have $a_1 = 1$, meaning that σ is an element of $\mathcal{N}(\mathbf{F}_p)$. In this way, elements of order p^n in $\mathcal{N}(\mathbf{F}_p)$ arise from totally ramified cyclic p^n -extensions of fields of Laurent series.

We first explicitly describe cyclic p^n -extensions using Witt vectors and then discuss how to detect whether they are totally ramified. By Artin–Schreier theory any abelian extension K/k of order p^n can be decomposed as a tower of field extensions

$$k = K_0 \subsetneq K_1 \subsetneq \cdots \subsetneq K_n = K \quad (3)$$

with $K_{i+1} = K_i(\alpha_i)$ for $0 \leq i \leq n-1$ and K_{i+1}/K_i an Artin–Schreier extension with $\alpha_i^p - \alpha_i \in K_i$. In the opposite direction Witt vectors allow one to guarantee that such an iterative procedure produces a *cyclic* extension K/k .

Any $\sigma \in \mathcal{N}(\mathbf{F}_p)$ of order p^n arises from such a construction: Harbater [38, §2] proved that every such σ describes the action of a generator of the Galois group on the completed local ring at a totally ramified point of a global $\mathbf{Z}/p^n\mathbf{Z}$ -Galois cover of \mathbf{P}^1 having a unique ramification point. The choice of a uniformiser at the ramified point (i.e. the choice of an isomorphism of the completed local ring with $\mathbf{F}_p[[t]]$) corresponds to a conjugation of the representing power series. It follows that any σ of order p^n is conjugate to an algebraic power series; note that the conjugating power series is an element of $\mathcal{N}(\mathbf{F}_p)$, but is not necessarily algebraic over $\mathbf{F}_p(t)$. The genus of the cover can be computed in terms of the break sequence from the (wild) Riemann–Hurwitz formula (compare [9, §3.3 & 3.4]).

Remark 2.1.1. The general theory of Harbater–Katz–Gabber covers ([46, 1.4.1], compare [9, §4.3, Cor. 4.10]) implies that any finite subgroup of $\text{Aut}(\mathbf{F}_p[[t]])$ can be conjugated into a subgroup consisting of algebraic power series (but, again, the conjugating series

itself need not be algebraic). Harbater proved the result for p -groups over perfect fields. For a cohomological characterisation of the occurring Galois covers, see [50].

Remark 2.1.2. There exist alternative methods for the explicit construction of equations for the Galois extensions. One may use explicit local class field theory, using the theory of formal groups/moduli of Lubin and Tate [55]. An essentially equivalent global method is to use explicit global class field theory of function fields, employing torsion of the Carlitz module [61], and then localising at a totally ramified place. This shows, at least theoretically, that the resulting series can be described by recursion relations or automata and immediately leads to a recursive algorithm to compute the coefficients of the power series. In Remark 5.1.3 and Subsection 7.3, we describe how to find series of order 4 and 8 in this way. In particular, we use this method to construct a complete set of representatives for all conjugacy classes of order 8 elements with minimal break sequence. We have performed more experiments implementing these methods and observed that they tend to lead to automata with more states compared to the above method. A possible reason is that class field theory methods give Ore-style equations that in algorithms produce state spaces of size doubly exponential in the degree of the equation (cf. Subsection 3.3 below).

2.2. Witt vectors and construction of p^n -extensions

Let k be a field of characteristic $p > 0$ and let $n \geq 1$ be an integer. Let $W_n(k)$ denote the ring of (n -truncated p -typical) Witt vectors over k . As a set $W_n(k)$ is equal to k^n , and we write its elements as vectors of length n . The zero and identity element of $W_n(k)$ are $0 = (0, \dots, 0)$ and $1 = (1, 0, \dots, 0)$. Addition and multiplication of two elements $a, b \in W_n(k)$ are defined by polynomial expressions in the coordinates $a_0, \dots, a_{n-1}, b_0, \dots, b_{n-1}$ of a and b (see e.g. Example 2.2.2 and 2.2.3 below that we will use later). The ring $W_n(k)$ comes with a Frobenius endomorphism $\text{Frob}: W_n(k) \rightarrow W_n(k)$ mapping the element (a_0, \dots, a_{n-1}) to $(a_0^p, \dots, a_{n-1}^p)$. The map $\wp := \text{Frob} - \text{Id}$ is an endomorphism of the underlying abelian group of $W_n(k)$. Writing k^{sep} for a separable closure of k , for any given $\beta \in W_n(k)$ there exists some $\alpha \in W_n(k^{\text{sep}})$ such that $\wp(\alpha) = \beta$. Such α is unique up to addition of an element of $\ker \wp = W_n(\mathbf{F}_p)$ and the extension $k(\wp^{-1}(\beta)) := k(\alpha_0, \dots, \alpha_{n-1})$ of k is independent of the choice of α . Note that $W_1(k)$ is just the field k .

Theorem 2.2.1 (Witt; cf. [52, p. 107, Thm. 5]). *Let k denote a field of characteristic $p > 0$, let k^{sep} denote a separable closure of k , and let n denote any positive integer. For any field K with $k \subseteq K \subseteq k^{\text{sep}}$, K/k is a cyclic Galois extension of degree p^n if and only if there exists a $\beta \in W_n(k)$ with $\beta_0 \notin \wp(k)$ such that $K = k(\wp^{-1}(\beta))$. If $\alpha \in W_n(k^{\text{sep}})$ satisfies $\wp(\alpha) = \beta$, then $k(\wp^{-1}(\beta)) = k(\alpha_0, \dots, \alpha_{n-1})$ and a generator σ of the Galois group $\text{Gal}(K/k)$ is determined by the equations*

$$\sigma(\alpha_i) = (\alpha + 1)_i, \quad i = 0, \dots, n-1. \quad (4)$$

Example 2.2.2. We consider the ring of Witt vectors $W_2(k)$ of length two over a field k of characteristic 2. For $a = (a_0, a_1), b = (b_0, b_1) \in W_2(k)$ the formulas for addition and multiplication are

$$a + b = (a_0 + b_0, a_1 + b_1 + a_0 b_0) \quad \text{and} \quad a \cdot b = (a_0 b_0, a_0^2 b_1 + a_1 b_0^2),$$

and the map \wp is given by $\wp(a) = (a_0^2 + a_0, a_1^2 + a_1 + a_0^2 + a_0^3)$. Observe that this implies that $-(a_0, a_1) = (a_0, a_1 + a_0^2)$. According to Theorem 2.2.1, an extension K/k is a cyclic Galois extension of degree 4 if and only if $K = k(\alpha_0, \alpha_1)$, where α_0, α_1 satisfy

$$\begin{cases} \alpha_0^2 + \alpha_0 = \beta_0; \\ \alpha_1^2 + \alpha_1 = \beta_1 + \beta_0 \alpha_0 \end{cases}$$

for some $\beta_0, \beta_1 \in k$ with β_0 not of the form $x^2 + x$ for $x \in k$. The Galois group of K/k is generated by the field automorphism σ defined on the generators α_0, α_1 by

$$\begin{cases} \sigma(\alpha_0) = \alpha_0 + 1; \\ \sigma(\alpha_1) = \alpha_1 + \alpha_0. \end{cases} \quad (5)$$

Example 2.2.3. We consider the ring of Witt vectors $W_3(k)$ of length three over a field k of characteristic 2. For $a = (a_0, a_1, a_2), b = (b_0, b_1, b_2) \in W_3(k)$ the formula for addition is

$$a + b = (a_0 + b_0, a_1 + b_1 + a_0 b_0, a_2 + b_2 + a_1 b_1 + a_0 a_1 b_0 + a_0 b_0 b_1 + a_0^3 b_0 + a_0 b_0^3)$$

and for multiplication is

$$a \cdot b = (a_0 b_0, a_0^2 b_1 + a_1 b_0^2, a_1^2 b_1^2 + a_0^4 b_2 + a_2 b_0^4 + a_0^2 a_1 b_0^2 b_1).$$

By Theorem 2.2.1, cyclic degree-8 extensions K/k of a field k of characteristic 2 are of the form $K = k(\alpha_0, \alpha_1, \alpha_2)$, where

$$\begin{cases} \alpha_0^2 + \alpha_0 = \beta_0; \\ \alpha_1^2 + \alpha_1 = \beta_1 + \beta_0 \alpha_0; \\ \alpha_2^2 + \alpha_2 = \beta_2 + \alpha_1 \beta_1 + \alpha_0 \alpha_1 \beta_0 + \alpha_0 \beta_0 \beta_1 + \alpha_0^3 \beta_0 + \alpha_0 \beta_0^3, \end{cases} \quad (6)$$

with β_0 not of the form $x^2 + x$ for $x \in k$. The Galois group of K/k is generated by the field automorphism defined on the generators $\alpha_0, \alpha_1, \alpha_2$ by

$$\begin{cases} \sigma(\alpha_0) = \alpha_0 + 1; \\ \sigma(\alpha_1) = \alpha_1 + \alpha_0; \\ \sigma(\alpha_2) = \alpha_2 + \alpha_0 \alpha_1 + \alpha_0^3 + \alpha_0. \end{cases} \quad (7)$$

2.3. Ramification

The ramification in an Artin–Schreier extension of $\mathbf{F}_p((z))$ can be described using the following easy result (see e.g. [34, III.(2.5)]).

Lemma 2.3.1. *Let $k = \mathbf{F}_p((z))$ and let $K = k(\alpha)$ be an extension of k with $\alpha^p - \alpha = \gamma$ for some $\gamma \in k$. If $v_z(\gamma)$ is negative and not divisible by p , then K/k is a cyclic extension of degree p , and for any uniformiser π of K we have $v_\pi(\alpha) = v_z(\gamma)$; for $x \in k$ we have $v_\pi(x) = pv_z(x)$.*

If we decompose a general cyclic totally ramified p^n -extension as a tower of Artin–Schreier extensions as in (3) and we write z_i for a uniformiser of K_i (so $z_0 = z$ and $z_n = t$), then $v_{z_{i+1}}(\alpha_i) = v_{z_i}(\alpha_i^p - \alpha_i)$ for $i = 0, \dots, n-1$.

The general approach is now to take the following steps:

- (i) Write down explicit equations for a cyclic p^n -extension in the variables α_i arising from the Witt construction, or other generators of the field (this may make equations simpler or help in applying Lemma 2.3.1 to check that the extension is totally ramified).
- (ii) Choose a uniformiser t as an algebraic function of the α_i (or the chosen field generators); using Lemma 2.3.1 allows us to control the valuations of rational functions in the field generators.
- (iii) Compute the action of a generator σ of the Galois group on the uniformiser t using the action in terms of Witt vectors given by Equation (4); this gives an equation for $\sigma(t)$ in terms of the α_i (or the chosen field generators).

These three steps lead to a set of algebraic equations from which one should eliminate all but t and $\sigma(t)$, leading to an algebraic equation $F(t, X) = 0$ with $F \in \mathbf{F}_p[t, X]$ satisfied by $X = \sigma = \sigma(t)$. For elimination, one may use a Groebner basis algorithm (we used the implementation in SINGULAR [31]; in order to be able to eliminate all the variables it might be necessary to first make a primary decomposition of the ideal generated by the equations and extract a one-dimensional component).

Example 2.3.2. We start describing what will be our ‘running example’ for the next few sections, leading up to a particularly small (as it will turn out, the smallest possible one in terms of number of states) automaton for a series of order 4 with ‘minimal’ break sequence.

Let $k = \mathbf{F}_2((z))$, $\beta = (z^{-1}, 0) \in W_2(k)$, and write $\alpha = (x, y) \in W_2(k^{\text{sep}})$ for a solution of $\wp(\alpha) = \beta$. Since $v_z(\wp(k)) = 2\mathbf{Z} \cup \mathbf{Z}_{\geq 0}$ we have $z^{-1} \notin \wp(k)$, and by Theorem 2.2.1 the extension $K/k = \mathbf{F}_2((z))(x, y)/\mathbf{F}_2((z))$, with x and y satisfying

$$\begin{cases} x^2 + x = z^{-1}; \\ y^2 + y = xz^{-1} = x^3 + x^2, \end{cases}$$

is a cyclic Galois extension of degree 4. It is totally ramified; an example of a uniformiser t for K is given by

$$t = (y + 1)/(y + x^2).$$

Indeed, breaking up the extension into Artin–Schreier extensions as in Equation (3), we have

$$k = K_0 = \mathbf{F}_2((z_0)) \subsetneq K_1 = K_0(x) = \mathbf{F}_2((z_1)) \subsetneq K_2 = K_1(y) = \mathbf{F}_2((z_2)) = K$$

with $z_0 = z, z_1, z_2$ uniformisers of the fields in the tower of extensions. So $v_{z_0}(z^{-1}) = -1$, $v_{z_1}(x) = -1$ and $v_{z_1}(z) = 2$. Hence $v_{z_1}(x^3 + x^2) = -3$, so K_1/K_0 is totally ramified. Then $v_{z_2}(y) = -3$, $v_{z_2}(x) = -2$ and $v_{z_2}(z) = 4$, so K_2/K_1 is also totally ramified. Hence t is a uniformiser for K since

$$v_{z_2}(t) = v_{z_2}(y + 1) - v_{z_2}(y + x^2) = 1.$$

Formula (5) shows that a generator σ of the Galois group is determined by the equations

$$\begin{cases} \sigma(x) = x + 1; \\ \sigma(y) = y + x, \end{cases}$$

the other generator is given by $\tau = \sigma^{\circ 3}$. We compute

$$\tau(t) = \sigma^{\circ 3} \left(\frac{y + 1}{y + x^2} \right) = \frac{y + x}{y + x^2 + x}.$$

To find an algebraic equation for $\tau = \tau(t)$ over $\mathbf{F}_2(t)$, we need to eliminate x and y from the three equations

$$\begin{cases} y^2 + y = x^3 + x^2 & \text{[equation of extension];} \\ (y + x^2)t = y + 1 & \text{[definition of uniformiser];} \\ (y + x^2 + x)\tau(t) = y + x & \text{[action of } \tau \text{ on uniformiser],} \end{cases}$$

from which we get that $X = \tau = \tau(t) \in \mathbf{F}_2[[t]]$ satisfies the (irreducible) equation

$$F(t, X) = (t + 1)^3 X^3 + (t^3 + t) X^2 + (t^3 + t + 1) X + t^3 + t = 0. \quad (8)$$

This equation has a unique solution of the form $t + O(t^2)$, as can be seen, e.g. from the corresponding t -adic Newton polygon; its initial coefficients are given by $t + t^2 + t^4 + t^5 + O(t^6)$.

2.4. Break sequence

By computing the first few coefficients of $\sigma \in \mathcal{N}(\mathbf{F}_p)$ of order p^n (using the algebraic equation for σ over $\mathbf{F}_p(t)$), it is easy to determine the lower break sequence of σ . If one has an explicit upper bound for the number of inequivalent series with given break sequences, we can enumerate all classes of such series by ‘trying’ enough equations, which sometimes works in practice. Such bounds are implicit in [54, Theorem 2.2] and have been made explicit in a few cases (cf. the discussion in Sections 5 & 7). Alternatively, using explicit local class field theory constructions as in [54] we are guaranteed to obtain representatives of all the conjugacy classes.

A method of Kanesaka and Sekiguchi directly computes the upper break sequence in terms of the Witt vector data for a given extension of $k := \mathbf{F}_p((z))$ [45, Thm. 5], which we rephrase as follows.

Definition 2.4.1. Fix a positive integer n . Call a vector $a = (a_i) \in \bigoplus_{\mathbf{N}} W_n(\mathbf{F}_p)$ of Witt vectors of length n (with finitely many nonzero entries) *suitable* if $a_i = 0$ for $p|i$ and for at least one i we have $a_i \in W_n(\mathbf{F}_p)^*$ (i.e. the zero component of a_i is not zero). If

$$\beta = (\beta_0, \dots, \beta_{n-1}) := \sum_{i \geq 0} a_i(z^{-i}, 0, \dots, 0) + \wp(b) \in W_n(k) \quad (9)$$

for a suitable $a = (a_i)$ and any $b \in W_n(k)$, define

$$\rho_n(\beta) := p^{-1} \max\{i \cdot \text{ord}(a_i) : a_i \neq 0\},$$

where $\text{ord}(a_i)$ is the order of a_i in the additive group $W_n(\mathbf{F}_p)$ (that itself is of exponent p^n). This is well-defined, since one can show that if a vector β admits such a representation, then the corresponding suitable vector is uniquely determined (since the vectors $(z^{-i}, 0, \dots, 0)$ are independent modulo $\wp(W_n(k))$). Also note that $\rho_n(\beta)$ is independent of $b \in W_n(k)$.

Define, for $m \leq n$, the truncation map $\lfloor (x_0, \dots, x_{n-1}) \rfloor_m := (x_0, \dots, x_{m-1})$. The truncation of a vector of the form as in Equation (9) in $W_n(k)$ is of that same form in $W_m(k)$.

Proposition 2.4.2 ([45]). For $k = \mathbf{F}_p((z))$ and a positive integer n , choose β of the form as in Equation (9) for a suitable vector $a = (a_i)$, some $b \in W_n(k)$, and assume $\beta_0 \neq 0$. Then the extension $k(\wp^{-1}(\beta))/k$ is a totally ramified cyclic extension of degree p^n , and the upper break sequence of a generator of the corresponding Galois group is $\langle \rho_1(\lfloor \beta \rfloor_1), \dots, \rho_n(\lfloor \beta \rfloor_n) \rangle$.

Although in this paper, we usually use lower break sequences, the above result is most naturally formulated in terms of upper break sequences; as remarked before, these can be easily changed into each other using Formula (2). The above result allows one to fix

not just p^n , but also the break sequence from the start, by choosing a suitable Witt vector $\beta \in W_n(k)$. Note that we get the same extension for every $b \in W_n(k)$, but it will be convenient to rewrite certain natural choices of β using nonzero b .

Example 2.4.3. We give some examples of constructions with break sequences that we will use later.

- (a) Choose $\beta = (z^{-1}, 0, \dots, 0) \in W_n(\mathbf{F}_p((z)))$ of length n , so all $a_i = 0$ for $i \neq 1$ and $a_1 = (1, 0, \dots, 0)$. Now a_1 is of order p^n in $W_n(\mathbf{F}_p)$ and the break sequence, called the *minimal* one, is

$$\langle p^i \rangle_{i=0}^{n-1} = \left(\frac{p^{2i+1} + 1}{p + 1} \right)_{i=0}^{n-1}.$$

- (b) For $\beta = (z^{-1}, z^{-pm}) \in W_2(\mathbf{F}_p((z)))$, with $m > p$ coprime to p , rewrite

$$\beta = a_1(z^{-1}, 0) + a_m(z^{-m}, 0)$$

with $a_1 = (1, 0)$ and $a_m = (0, 1)$. Now $\text{ord}(a_1) = p^2$ and $\text{ord}(a_m) = p$ in $W_2(\mathbf{F}_p)$, so we find the upper break sequence

$$\langle 1, m \rangle = (1, pm - p + 1).$$

- (c) For $\beta = (z^{-1}, z^{-m}) \in W_2(\mathbf{F}_p((z)))$ with $m > p$ coprime to p , we get the same break sequence as the previous example, since

$$(z^{-1}, z^{-m}) = (z^{-1}, z^{-pm}) - \wp((0, z^{-m})).$$

3. Detailed method: computing p -automata using proofs of Christol's theorem

3.1. Abstract algorithm

The following theorem of Christol relates algebraic power series to p -automatic sequences (see [23,24]):

Theorem 3.1.1 (Christol). *A power series $\sigma = \sum_{k \geq 0} a_k t^k \in \mathbf{F}_p[[t]]$ is algebraic over $\mathbf{F}_p(t)$ if and only if the sequence $(a_k)_{k \geq 0}$ is p -automatic.*

For our applications it is important that there are constructive proofs of this theorem: given an algebraic equation $F(t, X) = 0$ with $F(t, X) \in \mathbf{F}_p[t, X]$, the proofs can be turned into algorithms that compute p -automata representing the different solutions $X = \sigma \in \mathbf{F}_p[[t]]$. These algorithms start from a finite \mathbf{F}_p -vector space V with a distinguished nonzero vector $s_0 \in V$ and a set Λ of ‘Cartier-style’ operators $\Lambda_r: V \rightarrow V$ for

$r \in \{0, \dots, p-1\}$. From these data, they produce the directed graph structure of an automaton. A finite computation (using Hensel's Lemma) then fills in the vertex labels for the different solutions. For three such proofs/algorithms, we briefly indicate the triples (V, s_0, Λ) and point to other sources for proofs of correctness, optimised implementations and complexity analysis.

It follows from the proofs that for a given irreducible equation all solutions can be represented by automata with the same directed graph structure (including edge labels, but excluding vertex labels). Hence the desired algorithm can be broken down into two parts: first, the computation of that directed graph, and second, computing the correct output labels corresponding to the different solutions.

We will make the following assumptions and use the following notations throughout:

- $F(t, X) \in \mathbf{F}_p[t, X]$ is irreducible,
 - ◊ $d = \deg_X F$,
 - ◊ $h = \deg_t F$,
 - ◊ $m = \text{ord}_t \text{Res}_X(F(t, X), \frac{\partial F}{\partial X}(t, X))$ denotes the t -valuation of the resultant of F and its derivative in X , and
 - ◊ g denotes the geometric genus of the normalisation \mathcal{X} of the projective curve corresponding to the plane affine curve $F(t, X) = 0$.
- For $0 \leq r < p$, the *Cartier operator* \mathcal{C}_r acting on formal power series in $\mathbf{F}_p[[x_1, \dots, x_k]]$ is defined by

$$\mathcal{C}_r\left(\sum a_{i_1, i_2, \dots, i_k} x_1^{i_1} \cdots x_k^{i_k}\right) := \sum a_{pi_1+r, pi_2+r, \dots, pi_k+r} x_1^{i_1} \cdots x_k^{i_k}.$$

For the first part—the construction of the directed graph underlying the automaton—the proofs are based on constructing a graph from the specific set of data (V, s_0, Λ) , as follows:

▮ **Algorithm 3.1.2** (*Labeled Directed Graph Structure*).

Input A finite \mathbf{F}_p -vector space V , $s_0 \in V$, and maps $\Lambda = \{\Lambda_r: V \rightarrow V \text{ for } 0 \leq r < p\}$.
Output A finite directed graph with edge labels.

Write Γ for the monoid generated by the maps Λ_r with $0 \leq r < p$. Compute the set of vertices S as the orbit of s_0 under the action of Γ (by applying the maps Λ_r until no new elements appear), let the vertex s_0 be labelled ‘Start’, and put a directed edge between s_1 and s_2 with label r precisely if $s_2 = \Lambda_r(s_1)$. ▮

The second part can always be dealt with in the following way:

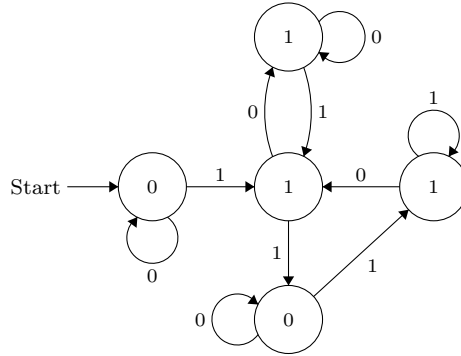


Fig. 2. A 2-automaton representing the element σ_{\min} of $\mathcal{N}(\mathbf{F}_2)$ of order 4 with lower break sequence (1, 3), corresponding to Equation (8).

⌈ **Algorithm 3.1.3** (*Vertex Labels*).

Input A polynomial $F(t, X) \in \mathbf{F}_p[t, X]$ and the directed graph structure, including edge labels, of automata representing all solutions $X = \sigma \in \mathbf{F}_p[[t]]$ of $F(t, X) = 0$.

Output A finite list of automata corresponding to all these solutions.

For an integer i , consider the truncated equation

$$F(t, \sigma_0) = O(t^{i+1}) \text{ with } \sigma_0 = a_0 + a_1t + a_2t^2 + \cdots + a_it^i. \quad (10)$$

- (i) Solve the truncated Equation (10) with $i = 2m$ for all the (finitely many) possible σ_0 . Hensel's Lemma implies that for each such σ_0 there is a unique solution $X = \sigma \in \mathbf{F}_p[[t]]$ of $F(t, X) = 0$ with $\sigma(t) = \sigma_0(t) + O(t^{m+1})$ (see e.g. the introduction of [12]).
- (ii) For each fixed σ_0 , run through the automaton following all base- p expansions of the integers $j = 0, 1, 2, \dots$ and give the final vertex of the walk corresponding to the base- p expansion of j the label a_j . For this, it may be necessary to compute the coefficients a_j of the solution of $F(t, X) = 0$ corresponding to σ_0 for some $j > 2m$, which can be done by solving the truncated equation inductively for $i = 2m + 1, \dots, j$, and use the leading zeros condition. ⌋

As we will indicate below, sometimes the vertex labels can be determined in a more efficient way, depending on the method used to compute the directed graph structure.

Example (continued) 3.1.4. Suppose we know that the directed graph structure of the solutions for Example 2.3.2 is as given in Fig. 2, but the possible vertex labels are still unknown. In this case, we have $m = 6$, and we are looking for a solution σ with $\sigma = t + O(t^2)$ (already known to exist). Substituting a tentative solution, we compute its

initial coefficients: $\sigma_{\min} = t + t^2 + t^4 + t^5 + t^7 + O(t^8)$. Using the coefficients of t^0, t^1, t^3, t^7 , the vertex labels are fixed uniquely, except for the label of the vertex reached from the start vertex by following the path 01. However, the assumption of leading zeros invariance fixes this value to be the same as that of the vertex reached by following the path 1. The resulting unique vertex labels are given in Fig. 2.

3.2. Three methods of constructing the input data

What is different in various proofs/algorithms is the construction of V, s_0 and Λ used as input for the construction of the directed graph. We briefly describe three possible approaches to this.

3.2.1. Using spaces of differential forms

This method is based on a proof by David Speyer and Andrew Bridy [13]. The fact that the algorithm is correct is explained in [13, §3]. A plug-and-play implementation of this algorithm is not available at the current time, but the built-in algorithms for function fields in MAGMA [11] include Kähler differentials and Cartier operators, making it relatively easy to implement the computations (but not the visualisations). The file [14] contains a description of a Magma routine that produces output that can be easily visualised in Mathematica and manipulated using [58].

Let Ω denote the \mathbf{F}_p -vector space of Kähler differentials on \mathcal{X} and K the function field of \mathcal{X} . Writing $\eta \in \Omega$ as $\eta = (u_0^p + u_1^p t + \cdots + u_{p-1}^p t^{p-1})dt$ for unique $u_i \in K$, define the Cartier operator $\mathcal{C}: \Omega \rightarrow \Omega$ by the formula $\mathcal{C}(\eta) := u_{p-1}dt$. Set $\omega := Xdt \in \Omega$ and define the effective divisor $D := (\omega)_\infty + (t)_\infty$, the sum of polar divisors of the differential ω and the function t . In this case:

- $V = \Omega(D)$ is the \mathbf{F}_p -vector space of differential forms on \mathcal{X} with divisor $\geq -D$ (of finite dimension $\leq h + 3d + g - 1$ over \mathbf{F}_p by Riemann–Roch, see [13, proof of Cor. 3.10]).
- $s_0 = \omega$.
- For any $r = 0, \dots, p-1$, define Λ_r as $\Lambda_r(\eta) := \mathcal{C}(t^{p-1-r}\eta)$. The maps Λ_r map V to itself (see [13, proof of Cor. 3.10]).

Example (continued) 3.2.1. Continuing the previous Example 2.3.2, we find (using MAGMA) that the curve corresponding to Equation (8) is of genus $g = 1$, the space $\Omega((Xdt)_\infty + (t)_\infty)$ is of dimension 8 and the subset $S = \Gamma(Xdt)$ has 5 elements corresponding to the vertices in the automaton. Representing these by the vectors

$$S = \{(1, 1, 0, 1, 0, 0, 1, 0), (1, 0, 0, 0, 0, 0, 0, 0), (0, 1, 0, 0, 0, 0, 1, 0), \\ (1, 1, 0, 1, 0, 0, 0, 0), (1, 0, 0, 0, 0, 1, 0, 0)\},$$

where the third vector is the start vertex, the action of the operators Λ_0 and Λ_1 is given by right multiplication with the following explicit 8×8 matrices over \mathbf{F}_2 :

$$\Lambda_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Lambda_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The resulting automaton is the one in Fig. 2.

3.2.2. Using equations in Ore form

This method is based on the proof from [24]. The fact that the algorithm is correct follows, e.g. from tracing through the proof of Christol's theorem in [5, Thm. 12.2.5] using [5, 12.2.4] for the expression for the corresponding p -kernel and the construction of the automaton corresponding to such a kernel as in the proof of the equivalence of ' p -automatic' and 'finite p -kernel', see e.g. [5, Thm. 6.6.2]. (The vector space described there is slightly larger, but the arguments show that the space defined below also works.) An implementation is described in [60, Rem. 4.7] and an actual implementation was done by Rowland in [58] (compare [59]).

One first computes a new polynomial $G(t, X) \in \mathbf{F}_p[t, X]$ in 'Ore form', i.e. $G(t, X) = \sum_{i=0}^d B_i X^{p^i}$ with $B_i \in \mathbf{F}_p[t]$, $B_0 \neq 0$, whose solution set in X contains the \mathbf{F}_p -vector space spanned by the solution set of F in X . Then the data are defined as follows:

- V is the set of linear combinations of elements from $\{X, X^p, \dots, X^{p^{d-1}}\}$ with coefficients being elements from $\mathbf{F}_p[t]$ of degree at most

$$N := \max(\deg B_0, \max\left\{\left\lceil \frac{\deg B_i + (p^i - 2) \deg B_0}{p - 1} \right\rceil - 1 \mid 1 \leq i \leq d\right\}).$$

- $s_0 := B_0 X$.
- For $0 \leq r < p$ and $D_k \in \mathbf{F}_p[t]$ of degree at most N , define

$$\Lambda_r \left(\sum_{k=0}^{d-1} D_k X^{p^k} \right) := \sum_{k=1}^{d-1} \mathcal{C}_r(D_k - D_0 B_k B_0^{p^k-2}) X^{p^{k-1}} - \mathcal{C}_r(D_0 B_d B_0^{p^d-2}) X^{p^{d-1}}.$$

The bound N on the degrees of D_k is chosen so that s_0 belongs to V and the operators Λ_r map V to itself (for this, note that for a polynomial $D \in \mathbf{F}_p[t]$ we have $\deg \mathcal{C}_r(D) \leq \lfloor \frac{\deg D}{p} \rfloor$).

One may circumvent the use of Algorithm 3.1.3: for the solution σ_0 whose truncation was fixed in (10) (with $\ell := \text{ord}_t B_0 \geq 1$) we can directly compute the labels of the vertices, as follows. Write

$$\frac{\sigma_0}{B_0} = b_1 t^{-(\ell-1)} + b_2 t^{-(\ell-2)} + \cdots + b_{\ell-1} t^{-1} + b_\ell + O(t) \quad (11)$$

with $b_i \in \mathbf{F}_p$; then the vertex corresponding to $\sum_{k=0}^{d-1} D_k X^{p^k} \in V$, where $D_k = \sum_{j \geq 0} [D_k]_j t^j$ with $[D_k]_j \in \mathbf{F}_p$, has vertex label equal to $\sum_{k=0}^{d-1} \sum_{\substack{0 \leq i \leq N \\ p^k | i}} [D_k]_i \cdot b_{\ell-i/p^k}$.

Example (continued) 3.2.2. The series τ from the previous Example 2.3.2 satisfies the following equation in Ore form:

$$G(t, X) = (t^8 + 1)X^8 + (t^8 + t^4 + t^2 + 1)X^4 + (t^7 + t^6 + t^5 + t^4 + t^2)X^2 + (t^7 + t^5)X = 0.$$

Now $\dim V = 150$ and S consists of the following five elements, resulting in the automaton in Fig. 2:

$$\begin{aligned} s_0 &= (t^7 + t^5)X, \\ s_1 &= (t^6 + t^3)X + (t^{14} + t^{13} + t^{11} + t^{10} + t^9 + t^7)X^2 \\ &\quad + (t^{28} + t^{27} + t^{26} + t^{25} + t^{20} + t^{19} + t^{18} + t^{17})X^4, \\ s_2 &= (t^7 + t^6 + t^5)X + (t^{13} + t^{11} + t^{10} + t^8)X^2 + (t^{28} + t^{26} + t^{20} + t^{18})X^4, \\ s_3 &= t^2X + (t^{13} + t^8 + t^7 + t^6)X^2 + (t^{26} + t^{24} + t^{18} + t^{16})X^4, \\ s_4 &= (t^6 + t^4)X. \end{aligned}$$

3.2.3. Using diagonals of two-variable power series

This method splits the problem into two cases (‘non-singular’ and ‘general’) and is based on a theorem of Furstenberg [35, Prop. 2] in combination with the proof in [23] and an observation in [2]. In the special case, the algorithm is described in [60, Algorithms 1 & 2]. The general algorithm is implemented in [58]. It is somewhat different from the preceding two methods: the non-singular case follows the setup considered before, in that it produces a triple (V, s_0, Λ) . The general case, however, might produce a different automaton for every solution.

Special case. Suppose $G \in \mathbf{F}_p[t, X]$ is *non-singular*, meaning that $G(0, 0) = 0$ and $c := \partial G / \partial X(0, 0)$ is nonzero. We search solutions $\sigma \in \mathbf{F}_p[[t]]$ of $G(t, \sigma) = 0$ with $\sigma(0) = 0$. In this case, by Hensel’s lemma, there is a *unique* such solution σ ; Furstenberg’s theorem says that

$$\sigma(t) = \Delta \left(\frac{P(t, X)}{Q(t, X)} \right) (t) \quad \text{with } P(t, X) := c^{-1} X \frac{\partial G}{\partial X}(tX, X) \text{ and} \\ Q(t, X) := c^{-1} X^{-1} G(tX, X),$$

where the *diagonal* ΔG of a two-variable power series $G(t, X) = \sum a_{r,s} t^r X^s \in \mathbf{F}_p[[t, X]]$ is defined as the one-variable power series $(\Delta G)(t) := \sum a_{r,r} t^r \in \mathbf{F}_p[[t]]$. To avoid confusion: in the definition of P , the derivative is that of $G(t, X)$ w.r.t. X , after which the result is evaluated at (tX, X) , and the constant c^{-1} is introduced so that $Q(0, 0) = 1$. The relevant data are:

- V is the space of polynomials in $\mathbf{F}_p[t, X]$ of degree at most $\max(\deg_t P, \deg_t Q)$ in t and of degree at most $\max(\deg_X P, \deg_X Q)$ in X .
- $s_0 := P(t, X)$.
- For $0 \leq r < p$, $\Lambda_r(s) := \mathcal{C}_r(sQ^{p-1})$.

In this case, Algorithm 3.1.3 may be avoided: $v \in V$ is a two-variable polynomial, and the corresponding (unique) vertex label is the value of this polynomial at $(0, 0)$.

General case. Following [2, §3.1], compute the finite list of all possible polynomials $q \in \mathbf{F}_p[t]$ of degree $\leq 2m$ such that $F(t, q(t)) = O(t^{2m+1})$. For each such q , set $s = m + \text{ord}_t(\frac{\partial F}{\partial X}(t, q(t)))$, $G(t, X) = t^{-s} F(t, t^m X + q(t))$. Now G is non-singular; apply the previous case to construct an automaton for the (unique) power series solution $\tau(t)$ of $G(t, X) = 0$ with $\tau(0) = 0$. Modify the automaton producing τ to an automaton producing a power series solution $\sigma = q + t^m \tau$ of $F(t, X) = 0$ using standard constructions with automata (see e.g. [5, Thm. 5.4.1 & Cor. 6.8.5], which have constructive proofs).

Example (continued) 3.2.3. For Example 2.3.2, the polynomial is non-singular and we have

$$P(t, X) = t^3 X^6 + t^2 X^5 + (t^3 + t) X^4 + X^3 + t X^2 + X, \\ Q(t, X) = t^3 X^5 + (t^3 + t^2) X^4 + (t^3 + t) X^3 + (t^3 + t + 1) X^2 + t X + t + 1.$$

The space V is of dimension 28 and V consists of 6 elements:

$$s_0 = P = t^3 X^6 + t^2 X^5 + (t^3 + t) X^4 + X^3 + t X^2 + X, \\ s_1 = \Lambda_0(s_0) = t^3 X^5 + (t^3 + t) X^3 + t X, \\ s_2 = \Lambda_1(s_0) = t^2 X^4 + t^2 X^3 + (t + 1) X^2 + t X + 1, \\ s_3 = \Lambda_0(s_2) = t^2 X^4 + X^2 + 1, \\ s_4 = \Lambda_1(s_2) = t^2 X^4 + (t^2 + t + 1) X^2 + X, \\ s_5 = \Lambda_1(s_4) = t^2 X^4 + (t^2 + 1) X^2 + 1,$$

with

$$\Lambda_0(s_1) = s_1, \Lambda_1(s_1) = s_2, \Lambda_0(s_3) = s_3, \Lambda_1(s_3) = s_2, \Lambda_0(s_4) = s_4, \Lambda_0(s_5) = s_2, \Lambda_1(s_5) = s_5.$$

This leads to an automaton with 6 states, but the states corresponding to s_0 and s_1 have the same outgoing edges and the same output labels, and hence can be merged into one state without affecting the automatic sequence produced by the automaton. Doing so leads again to the automaton in Fig. 2.

3.3. Bounds on the complexity

The exact complexity of the algorithms does not appear to be known, but upper bounds on the number of states $\#S$ have been given in terms of d and h . In essentially all the known examples, these are obtained by first bounding the dimension of the vector space V , and then using the trivial inequality $\#S \leq p^{\dim V}$. In practice, it is often the case that the set S is much smaller than the vector space V (as seen, e.g. in Examples 3.2.1–3.2.3). We will show in Proposition 9.2.1 that $d = h$ for series of finite compositional order, and then we have the following upper bounds:

- Differential forms: $\log_p \#S \leq 4d + g - 1 \leq d(d + 2) \approx d^2$ ([13, Cor. 3.10] and the inequality $g \leq (d - 1)(h - 1)$ of Castelnuovo–Riemann [62, Cor. 3.11.4]);
- Ore polynomials: $\log_p \#S \leq d^3 p^d (p^d - 1) / (p - 1) \approx d^3 p^{2d-1}$ (using the upper bound dhp^d for the height of the Ore form equation from [1, Lem. 8.1]);
- Diagonals (non-singular case): $\log_p (\#S - 1) \leq d(d + 1) \approx d^2$ (for this bound it is shown that all states in S except possibly for s_0 lie in a vector subspace of V of dimension $d(h + 1)$ [60, Rem. 4.7], [3, Thm. 3.1]; the latter reference also contains an argument that shows that in the general case, the diagonal method gives a similar upper bound asymptotically in d as the differential forms method).

In our running example, $\#S$ is 5 or 6, and the respective bounds on $\#S$ are 2^{12} , 2^{1512} and $2^{12} + 1$. For more information on the exact complexity of our examples (that appear to require far fewer states than the theoretical general bounds), we refer to Section 9.

3.4. Our application

Our construction using Witt vectors produces a polynomial $F(t, X) \in \mathbf{F}_p[t, X]$ of which we first check irreducibility (if the polynomial were not irreducible, we would factor it and work with the factors). We know the polynomial has at least one solution $\sigma(t) = t + O(t^2) \in \mathbf{F}_p[[t]]$, and we search only for such solutions. Most of the time, we can prove that there will be a unique solution of this form, and we then know that this σ has the desired finite order under composition. In some cases, we find more than one solution, but in these cases, we can identify the correct series in a different way. For actual computations, we relied on implementations of all three algorithms; see the

section ‘How computations and visualisations were done’ at the end of the paper for details.

The results obtained in our running Example 2.3.2, 3.1.4, and either one of 3.2.1, 3.2.2 or 3.2.3 may be summarised as follows:

Proposition 3.4.1. *The series σ_{\min} corresponding to the automaton in Fig. 2 is of order 4 in $\mathcal{N}(\mathbf{F}_2)$ and has break sequence (1, 3) and initial coefficients $\sigma_{\min} = t + t^2 + t^4 + t^5 + O(t^6)$. \square*

A given finite order element of the Nottingham group can have in its conjugacy class many algebraic power series, which satisfy polynomial equations of various degrees. It would be interesting to find a theoretical upper bound on the minimal degree d of such a polynomial. This would also give an upper bound on the genus g of the curve \mathcal{X} (see Subsection 3.3).

4. An enumeration algorithm for automata on at most N states representing finite order series

4.1. An abstract algorithm

Before we start applying our construction in concrete cases, we discuss an enumeration algorithm for finding all ‘small’ (in terms of number of states) minimal automata representing an element in $\mathcal{N}(\mathbf{F}_2)$ of given finite order. The theoretical algorithm, which can readily be generalised to p -automata and order p^n elements in $\mathcal{N}(\mathbf{F}_p)$, consists of two parts.

▮ **Algorithm 4.1.1** (*Compositional Power Automaton*).

Input A 2-automaton A and an integer $n \geq 0$.

Output If σ denotes the series corresponding to A , a 2-automaton A_n corresponding to the series $\sigma^{\circ 2^n}$.

- (i) Find a polynomial $F(t, X) \in \mathbf{F}_2[t, X]$ with $F(t, \sigma) = 0$. This can be done by following the proof of Christol’s Theorem 3.1.1 (in the direction different from the one used in Section 3)—from the automaton, determine the 2-kernel using [5, Thm. 6.6.2] and then follow the first part of the proof in [5, Thm. 12.2.5].
- (ii) Composing with $\sigma(t)$ on the right gives $F(\sigma(t), \sigma^{\circ 2}(t)) = 0$. Eliminate Y from $F(t, Y) = F(Y, X) = 0$ to produce an algebraic equation $F_1(t, X) = 0$ satisfied by $X = \sigma^{\circ 2}$. Repeat this procedure to produce an algebraic equation $F_n(t, X) = 0$ for $\sigma^{\circ 2^n}$.
- (iii) Construct an automaton A_n for $\sigma^{\circ 2^n}$ from the equation $F_n(t, X) = 0$, using the methods of Section 3. ▮

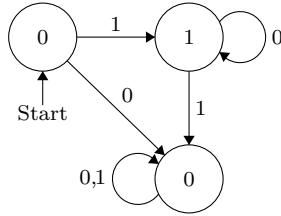


Fig. 3. Automaton for the power series t .

We will use the well-known fact that to each automaton A corresponds a unique minimal deterministic finite automaton \hat{A} with the same corresponding series, and that \hat{A} can be computed from A by an algorithm, see e.g. [51, §2.4]. In particular, one can check by an algorithm whether or not two automata A and B correspond to the same series—this happens precisely when $\hat{A} = \hat{B}$.

⌈ **Algorithm 4.1.2** (*Enumeration Bounded Size Automata of Fixed Compositional Order*).

Input Integers $n \geq 0$ and $N \geq 1$.

Output A finite list of all minimal 2-automata on at most N states representing an element of finite order 2^n in $\mathcal{N}(\mathbf{F}_2)$.

- (i) Go over all 2-automata on at most N states and eliminate those for which the corresponding power series is not of the form $\sigma = t + O(t^2)$.
- (ii) Remove duplicates from the list by comparing their minimal automata.
- (iii) For each remaining automaton A use Algorithm 4.1.1 to compute the automaton A_n .
- (iv) Compute the minimal automaton \hat{A}_n corresponding to A_n and check whether it equals the 3-state minimal automaton generating the series t , depicted in Fig. 3. ⌋

We do not know of an algorithm that lists all automata of size at most N corresponding to series of arbitrary but finite compositional order.

4.2. A practical implementation with application

A practical implementation of a more optimal algorithm in C++ was given by Groot Koerkamp [37] and produces a list of candidates for automata on at most 5 states representing series of order 2 and 4. Running that algorithm, we find a unique candidate automaton corresponding to a series of order 4. Since we already know from Proposition 3.4.1 that σ_{\min} is an order-4 series which is represented by an automaton with 5 states, this proves the following.

Proposition 4.2.1 (*Groot Koerkamp, [37]*). *The unique minimal (leading zeros invariant) 2-automaton with at most 5 states representing a power series of compositional order 4 is the one corresponding to the series σ_{\min} and depicted in Fig. 2. \square*

5. Construction and classification of some order-4 elements

5.1. Order 4, break sequence $(1, 3) = \langle 1, 2 \rangle$

Below are two known explicit power series with this order and break sequence: the one discovered by Chinburg and Symonds [22] and its compositional inverse, computed by Scherr and Zieve [9, Remark 1.4]:

$$\sigma_{\text{CS}} := t + t^2 + \sum_{k \geq 0} \sum_{\ell=0}^{2^k-1} t^{6 \cdot 2^k + 2\ell} = t + t^2 + O(t^6); \quad (12)$$

$$\sigma_{\text{CS}}^{\circ 3} = \sum_{k \geq 0} \left(t^{3 \cdot 2^k - 2} + t^{4 \cdot 2^k - 2} \right) = t + t^2 + t^4 + O(t^6). \quad (13)$$

An unpublished result of Lubin ([53], see [41, Thm. 2.2] for a proof) implies that there are precisely two conjugacy classes of such elements in $\mathcal{N}(\mathbf{F}_2)$. We now present a slightly more detailed lemma that allows us to distinguish between these conjugacy classes based on the first few coefficients alone.

Lemma 5.1.1. *Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ be an automorphism of order 4 with break sequence $(1, 3) = \langle 1, 2 \rangle$, and write $\sigma = \sum_{i=1}^{\infty} a_i t^i$ with $a_i \in \mathbf{F}_2$. Then $a_1 = a_2 = 1$, $a_3 = 0$, and exactly one of the following cases holds:*

- (a) $a_4 = a_5$ and σ is conjugate to σ_{CS} ;
- (b) $a_4 \neq a_5$ and σ is conjugate to $\sigma_{\text{CS}}^{\circ 3}$.

Proof. We have $a_1 = 1$ since $\sigma \in \mathcal{N}(\mathbf{F}_2)$, and $a_2 = 1$, $a_3 = 0$ since σ has lower break sequence $(1, 3)$; for the latter statement, compute the power series $\sigma^{\circ 2} = t + (1 + a_3)t^4 + O(t^5)$. The only possibilities for such series up to $O(t^6)$ are hence the four truncated series $\sigma = t + t^2 + a_4 t^4 + a_5 t^5 + O(t^6)$ with $a_4, a_5 \in \mathbf{F}_2$. Two of these correspond to (12) and (13), and for the other two, we observe that conjugating by $\phi : t \mapsto t + t^3$ gives

$$\begin{aligned} \phi^{-1} \circ \sigma_{\text{CS}} \circ \phi &= t + t^2 + t^4 + t^5 + O(t^6); \\ \phi^{-1} \circ \sigma_{\text{CS}}^{\circ 3} \circ \phi &= t + t^2 + t^5 + O(t^6). \end{aligned}$$

The quoted result of Lubin in [41, Thm. 2.2] implies that there are precisely two conjugacy classes of power series with break sequence $(1, 3) = \langle 1, 2 \rangle$. To finish the proof it is therefore enough to show that any automorphisms $\sigma, \tau \in \mathcal{N}(\mathbf{F}_2)$ with

$$\sigma = t + t^2 + O(t^6) \quad \text{and} \quad \tau = t + t^2 + t^4 + O(t^6)$$

are not conjugate in $\mathcal{N}(\mathbf{F}_2)$. Suppose this is the case, and let $\psi \in \mathcal{N}(\mathbf{F}_2)$ be such that $\sigma \circ \psi = \psi \circ \tau$. This implies that

$$\psi(t) + \psi(t)^2 + O(t^6) = \psi(t + t^2 + t^4) + O(t^6). \quad (14)$$

Writing $\psi(t) = t + \sum_{i=2}^{\infty} b_i t^i$ with $b_i \in \mathbf{F}_2$ and comparing the coefficients of t^4 and t^5 in (14) gives

$$b_2^2 + b_4 = 1 + b_2 + b_3 + b_4 \quad \text{and} \quad b_5 = b_3 + b_5,$$

which gives a contradiction since $b_2 \in \mathbf{F}_2$. \square

Corollary 5.1.2. *The series σ_{CS} and $\sigma_{\text{CS}}^{\circ 3}$ form a full set of representatives for the conjugacy classes of elements of order 4 with break sequence $(1, 3) = \langle 1, 2 \rangle$ in $\mathcal{N}(\mathbf{F}_2)$. \square*

The following different power series of order 4 and break sequence $(1, 3)$ was found earlier by Jean in [42] as a solution to the equation $(t + 1)\sigma^2 + (t^2 + 1)\sigma + t = 0$:

$$\sigma_J := \sum_{k \geq 0} \frac{t^{2^k}}{(t + 1)^{3 \cdot 2^k - 1}} = t + t^2 + t^5 + O(t^6). \quad (15)$$

Lemma 5.1.1 implies that it is conjugate to $\sigma_{\text{CS}}^{\circ 3}$.

Let us show how the power series of Chinburg–Symonds and Jean fit into our construction, and present the corresponding automata, using the same totally ramified cyclic extension $\mathbf{F}_2((z))(x, y)/\mathbf{F}_2((z))$ of degree 4 as in Example 2.3.2, but choosing different uniformisers t .

(i) First, let $t = yx^{-2}$. After elimination, we find the (irreducible) equations

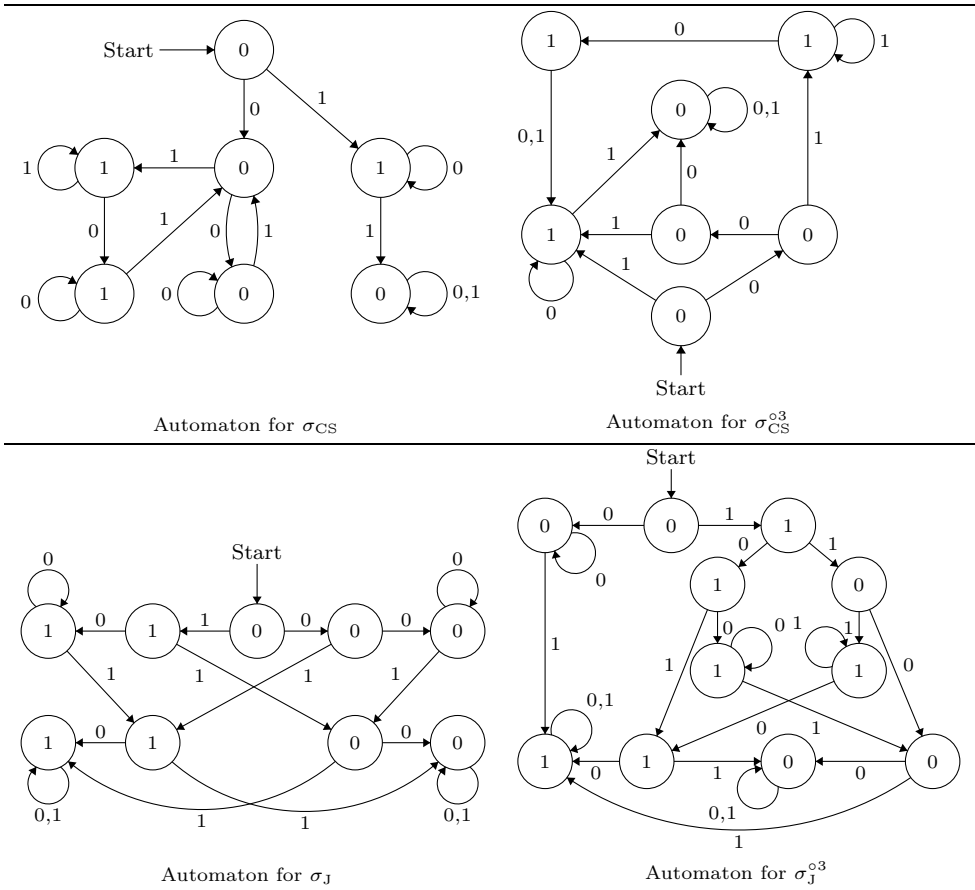
$$\begin{aligned} t^2 X^2 + X + t^2 + t &= 0; \\ (t^2 + 1)X^2 + X + t &= 0 \end{aligned}$$

for σ and τ , respectively. Looking at the valuations of the coefficients, we see that these equations have unique solutions of the form $t + O(t^2)$. The corresponding automata are given in the top right (σ) and the top left (τ) of Table 1. We now briefly indicate how these automata can be used to construct explicit formulas for σ and τ , showing that $\sigma = \sigma_{\text{CS}}^{\circ 3}$ and $\tau = \sigma_{\text{CS}}$.

- Write $\tau = \sum_{i \geq 1} a_i t^i$ with $a_i \in \mathbf{F}_2$. We will use the automaton corresponding to τ to determine for which $i \geq 1$ we have $a_i = 1$. For such i , starting at the start vertex and walking through the automaton following the successive digits of i

Table 1

Automata corresponding to series of Chinburg–Symonds and Jean and their inverses.



in base 2 (beginning with the least significant digit), we end up in a vertex with label 1. Since we can disregard any leading zeros, this vertex has an incoming edge with label 1. For τ note that this property holds precisely for those i for which the base-2 expansion is either 1, 10 or of the form $11d_k \cdots d_1 0$ for some $k \geq 0$, $d_1, \dots, d_k \in \{0, 1\}$, i.e. for i equal to 1, 2 or such that $6 \cdot 2^k \leq i < 8 \cdot 2^k$ for some $k \geq 0$. It follows that τ is given by the formula in (12).

- For the power series $\sigma = \sum_{i \geq 1} b_i t^i$ we see that the positive integers i for which $b_i = 1$ are precisely those which have a base-2 expansion of the form 1, 100, $1^k 10$ or $101^k 10$ with $k \geq 0$, and these are exactly the base-2 expansions of the numbers 1, 4, $4 \cdot 2^k - 2$ and $12 \cdot 2^k - 2$. This proves the formula for σ given in (13).

The fact that we can find such an explicit expression appears to be quite special. This relates to the fact that the automaton is ‘sparse’ in the sense of Section 10 below. The automaton for τ is not sparse, but the base-2 expansion of the occurring

powers has an explicit ‘closed’ form. It turns out that this series is sparse up to multiplication by a rational function.

(ii) Second, let $t = xy^{-1}$. Then we find the (irreducible) equations

$$\begin{aligned}(t+1)X^2 + (t^2+1)X + t &= 0; \\ tX^2 + (t^2+1)X + t^2 + t &= 0\end{aligned}$$

for σ and τ , respectively. From formula (15) we deduce that σ_J satisfies the same algebraic equation as σ , and since this equation has a unique solution of the form $t + O(t^2)$, we have $\sigma_J = \sigma$. Solving the equations for σ and τ by automata, we find that σ correspond to the bottom left, and τ to the bottom right automaton depicted in Table 1. Converting the automata into explicit series as above, we find (after some rewriting) that

$$\begin{aligned}\sigma_J = \sigma &= t + (t^7 + t^2) \sum_{k \geq 0} t^{8k} + \sum_{k, \ell \geq 0} \left(t^{4 \cdot 2^k(4\ell+1)+1} + t^{4 \cdot 2^k(4\ell+3)} \right) \\ &= t + \frac{t^7 + t^2}{t^8 + 1} + \sum_{k \geq 2} \frac{t^{3 \cdot 2^k} + t^{2^k+1}}{t^{4 \cdot 2^k} + 1},\end{aligned}$$

and

$$\begin{aligned}\sigma_J^{\circ 3} = \tau &= t + (t^{11} + t^5) \sum_{k \geq 0} t^{16k} + \sum_{k \geq 1, \ell \geq 0} \left(t^{2^k(2\ell+1)} + t^{4 \cdot 2^k(4\ell+1)-1} + t^{4 \cdot 2^k(4\ell+3)+1} \right) \\ &= t + \frac{t^{11} + t^5}{t^{16} + 1} + \frac{t^2}{t^2 + 1} + \sum_{k \geq 3, \ell \geq 0} \left(t^{2^k(4\ell+1)-1} + t^{2^k(4\ell+3)+1} \right).\end{aligned}\quad (16)$$

On the other hand, from the algebraic equation for τ (which has a unique solution of the form $t + O(t^2)$), we can find directly another explicit form for τ : the series $\tilde{\tau} := \frac{t}{t^2+1} \cdot \tau$ satisfies $\tilde{\tau} = t^2/(t+1)^3 + \tilde{\tau}^2$, and hence (iteratively) $\tilde{\tau} = \sum_{k \geq 0} (t^2/(t+1)^3)^{2^k}$, leading to the formula

$$\sigma_J^{\circ 3} = \tau = \sum_{k \geq 0} \frac{t^{2 \cdot 2^k - 1}}{(t+1)^{3 \cdot 2^k - 2}}.\quad (17)$$

The series σ and τ are further closed forms of elements of order 4 in $\mathcal{N}(\mathbf{F}_2)$ with break sequence (1, 3) and conjugate to $\sigma_{CS}^{\circ 3}$ and σ_{CS} , respectively.

The element σ_{\min} in Proposition 3.4.1 is conjugate to σ_{CS} .

Remark 5.1.3. We outline a construction of an automaton for such a series of order 4 with minimal break sequence using the Carlitz module construction of abelian extensions of

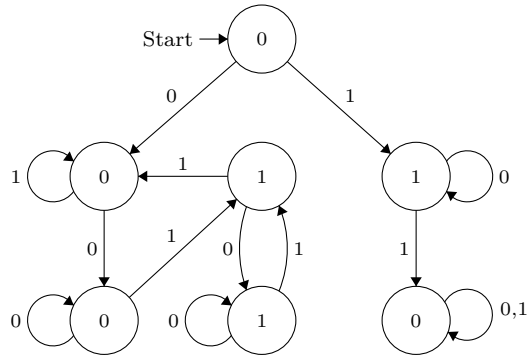


Fig. 4. Automaton corresponding to $\sigma_{\text{CS}}^{\circ 2} \in \mathcal{N}(\mathbf{F}_2)$ of order 2 with break sequence (3).

function fields, see e.g. [39] (this is a global class field theory version essentially equivalent to the local method based on Lubin–Tate theory used by Jean).

Let $\rho: \mathbf{F}_2[z] \rightarrow \text{End}(\mathbf{G}_a)$ denote the Carlitz module for $K := \mathbf{F}_2(z)$ defined by $\rho_z(X) = zX + X^2$. Now the extension $K(\rho[z^3])/K$ given by adjoining the roots of $\rho_{z^3}(X)$ is Galois with Galois group $G = (\mathbf{F}_2[z]/z^3)^* \cong \mathbf{Z}/4\mathbf{Z}$, generated by the class of $z+1$ (of order 4), where an element $g \in G$ acts on $\alpha \in K(\rho[z^3])$ by $g(\alpha) := \rho_g(\alpha)$. A minimal polynomial for the extension is $f := \rho_{z^3}(X)/\rho_{z^2}(X) = X^4 + (z^2 + z)X^2 + z^2X + z$, its splitting field is a cyclic degree-4 extension in which z is totally ramified (and no other place ramifies, cf. [39, Prop. 2.2, Thm. 3.2]), and a root t is a uniformiser for the extension locally above z . The action of a generator of the Galois group is given by $\sigma(t) = \rho_{z+1}(t) = t + zt + t^2$.

Eliminating z , we find an equation $(t+1)X^2 + (t^2+1)X + t = 0$ for σ . This is exactly the equation for σ_J , previously obtained using Witt vectors, and solved by a series corresponding to the automaton in Table 1 with 9 states.

Remark 5.1.4. If τ is an element of order 4 with break sequence (1, 3), then $\tau^{\circ 2}$ has break sequence (3), and hence is conjugate to the Klopsch’s series $\sigma_{K,3}$ (see Example 1.3.1). Taking $\tau = \sigma_{\text{CS}}$ produces the power series $\sigma := \sigma_{\text{CS}}^{\circ 2} = t + t^4 + O(t^5)$, which satisfies $(t^2+1)X^2 + X + t^2 + t = 0$. The corresponding automaton is presented in Fig. 4, leading to the following explicit formula for an element of order 2 with break sequence (3):

$$\sigma_{\text{CS}}^{\circ 2} = t + \sum_{k \geq 0} \sum_{\ell=0}^{2^k-1} t^{4 \cdot 2^k + 2\ell} = t + \frac{1}{t^2+1} \sum_{k \geq 1} (t^{2 \cdot 2^k} + t^{3 \cdot 2^k}).$$

5.2. Order 4, break sequence (1, 5) = (1, 3)

By Lubin’s result ([53], [41, Thm. 2.2]), there is a unique conjugacy class of such power series. No formula for such a series is known, but following our philosophy, we can represent the solution by an automaton.

Proposition 5.2.1. *Up to conjugation, every element in $\mathcal{N}(\mathbf{F}_2)$ of order 4 with break sequence $(1, 5) = \langle 1, 3 \rangle$ is given by the power series $\sigma_{(1,5)}$ corresponding to the automaton in Fig. 5 with 13 states, with initial coefficients*

$$\sigma_{(1,5)} = t + t^2 + t^3 + t^4 + t^6 + O(t^7).$$

Proof. Suitable algebraic equations are found from Witt's theory using Example 2.2.2; following Example 2.4.3, we start with the element $\beta := (z^{-1}, z^{-3}) \in W_2(\mathbf{F}_2((z)))$, and rewrite the resulting equation in terms of the variables $x := \alpha_0$ and $y := \alpha_1 + \alpha_0^3 + \alpha_0^2$ as

$$\begin{cases} x^2 + x = z^{-1}; \\ y^2 + y = x^5 + x^3. \end{cases} \quad (18)$$

(The variable y is used instead of α_1 since that choice allows us to use Lemma 2.3.1.) Writing $z_0 = z, z_1, z_2$ for uniformisers of the fields in the tower of extensions

$$K_0 := \mathbf{F}_2((z)) \subsetneq K_1 = K_0(x) = \mathbf{F}_2((z_1)) \subsetneq K_2 = K_1(y) = \mathbf{F}_2((z_2)),$$

we have $v_{z_1}(x) = v_{z_0}(z^{-1}) = -1$, so $v_{z_1}(x^5 + x^3) = -5$, and hence $v_{z_2}(y) = -5$ and $v_{z_2}(x) = -2$. Hence the extensions are all totally ramified and we can choose $t = x^2 y^{-1}$ as uniformiser for K_2 (since $v_{z_2}(t) = 1$). A generator σ for the Galois group of K_2/K_0 is determined by

$$\begin{cases} \sigma(x) = x + 1; \\ \sigma(y) = y + x^2 + 1, \end{cases}$$

and with $t = x^2 y^{-1}$ we compute that $\sigma(t) = (x^2 + 1)/(y + x^2 + 1)$. By eliminating x and y from these last two equations and the two equations in (18), we find that $\sigma = \sigma(t)$ satisfies the following (irreducible) equation over $\mathbf{F}_2(t)$:

$$t^2 X^3 + (t + 1)^3 X + t^3 + t = 0. \quad (19)$$

Considering the sum and product of the three solutions, we find that there is a unique solution with $\sigma = t + O(t^2)$. The corresponding automaton with initial coefficients $t + t^2 + t^3 + t^4 + t^6 + O(t^7)$ and Equation (19) produced by the algorithm is displayed in Fig. 5. \square

5.3. Order 4, break sequence $(1, 9) = \langle 1, 5 \rangle$

Again by Lubin's result in [53], there is a unique conjugacy class of such power series. A corresponding automaton is found as follows.

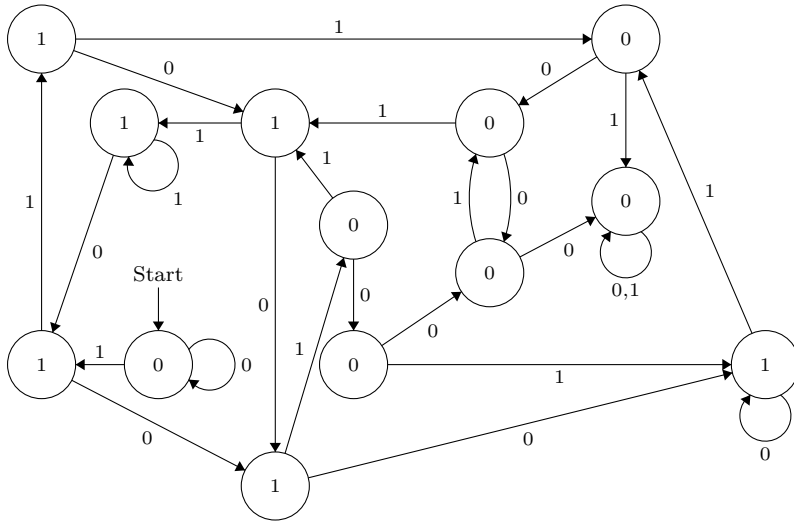


Fig. 5. Automaton representing a power series $\sigma_{(1,5)}$ of order 4 with break sequence $(1, 5)$ (unique up to conjugation).

Proposition 5.3.1. *Up to conjugation, every element in $\mathcal{N}(\mathbf{F}_2)$ of order 4 with break sequence $(1, 9) = \langle 1, 5 \rangle$ is given by the power series $\sigma_{(1,9)}$ corresponding to the automaton described as follows using the data in Table 2: it has 110 states, corresponding to the 110 triples on the displayed ordered list, where the start vertex is the first triple on the list and a triple (ℓ, i, j) occurs on the list precisely if the following three conditions hold: it has label ℓ , there is a directed edge with label 0 to the i -th triple on the list and there is a directed edge with label 1 to the j -th triple on the list. The initial coefficients of $\sigma_{(1,9)}$ are*

$$\sigma_{(1,9)} = t + t^2 + t^3 + t^4 + t^5 + t^6 + t^7 + t^9 + t^{11} + t^{12} + t^{13} + t^{17} + t^{18} + O(t^{19}).$$

Proof. Following Example 2.4.3(c), we start with $\beta = (z^{-1}, z^{-10}) \in W_2(\mathbf{F}_2((z)))$. In the resulting equations $\wp(\alpha) = \beta$, change variables to $x := \alpha_0$ and $y := \alpha_1 + \alpha_0^{10} + \alpha_0^9 + \alpha_0^6 + \alpha_0^3 + \alpha_0$ to find

$$\begin{cases} x^2 + x = z^{-1}; \\ y^2 + y = x^9 + x. \end{cases}$$

Writing $z_0 = z$, z_1 , z_2 for uniformisers of the fields in the tower of extensions

$$K_0 := \mathbf{F}_2((z)) \subsetneq K_1 = K_0(x) = \mathbf{F}_2((z_1)) \subsetneq K_2 = K_1(y) = \mathbf{F}_2((z_2)),$$

we have $v_{z_1}(x) = -1$, so $v_{z_1}(x^9 + x) = -9$, and hence $v_{z_2}(y) = -9$, $v_{z_2}(x) = -2$ and $v_{z_2}(z) = 4$. Hence all extensions are totally ramified and we can choose $t = x^{-1}yz^2$ as uniformiser for K_2 (since $v_{z_2}(t) = 1$). A generator σ for the Galois group of K_2/K_0 is determined by

$$\begin{cases} \sigma(x) = x + 1; \\ \sigma(y) = y + x^4 + x^2 + x + 1. \end{cases}$$

By elimination of variables, we find that $\sigma = \sigma(t)$ satisfies the following (irreducible) equation over $\mathbf{F}_2(t)$:

$$t^2\sigma^7 + t^3\sigma^6 + (t^5 + t^4 + t^2)X^5 + (t^5 + t^3)X^4 + \\ (t^7 + t^5 + t^4 + t^3 + t)X^3 + t^5X^2 + (t^3 + t + 1)X + t = 0.$$

There is a unique solution of the form $t + O(t^2)$, and its initial coefficients are as indicated in the proposition; the corresponding 2-automaton can be found in Table 2 and in [17] (the visual representation in Table 2 is more of an illustration but can be manipulated directly in [17] using standard graph theory algorithms). \square

6. Some new explicit formulas for power series of order 4

The explicit power series σ_{CS} and its inverse are a full set of representatives for the conjugacy classes of order-4 elements with break sequence $(1, 3)$. The series σ_{J} is another power series with a nice closed formula. We did a larger search for automata corresponding to such power series and found five more for which we could write down reasonably sized closed formulas. One of these is the inverse of Jean's series displayed in Equations (16), (17). We list the other four in Table 3.

We start with the equation from Example 2.3.2, but choose different uniformisers t . Recall that we write $\tau = \sigma^{\circ 3}$.

- (i) First, let $t = (1 + x^2 + y)/(x^2 + xy)$. Then $\sigma = \sigma_{\text{T},1}$ satisfies

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

and $\tau = \sigma_{\text{T},2}$ satisfies

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

Solving these (irreducible) equations by automata, we find that $\sigma_{\text{T},1}$ and $\sigma_{\text{T},2}$ correspond to the top left, respectively top right automaton depicted in Table 4. It is relatively straightforward to convert the automata into explicit series following the method explained after Corollary 5.1.2, and the result is shown in Table 3 (including the initial coefficients).

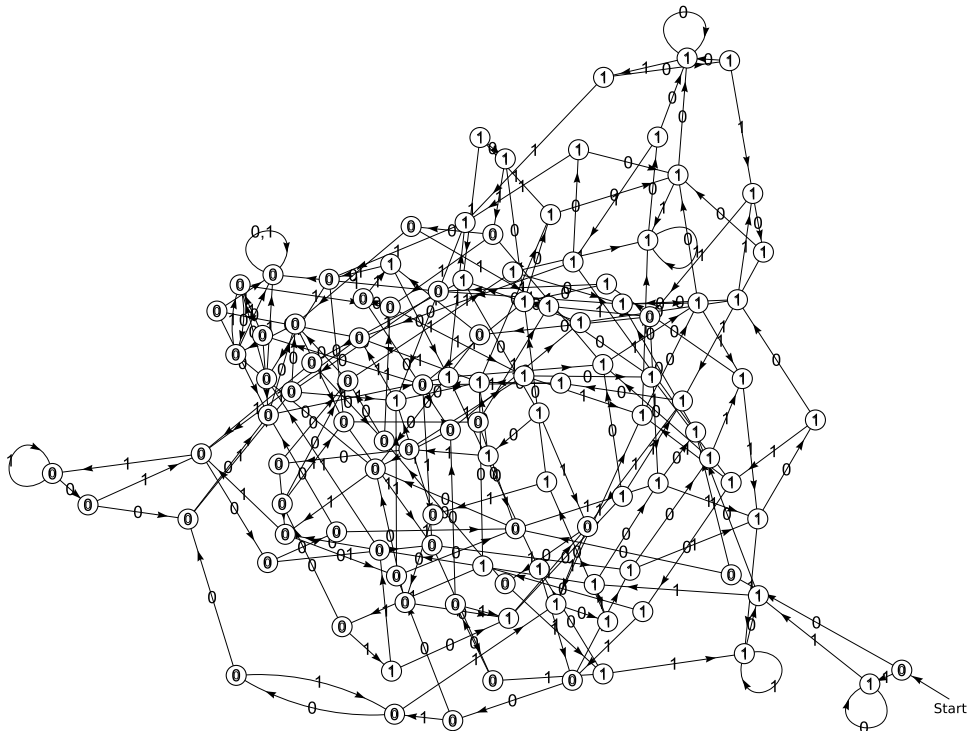
- (ii) Second, let $t = xy/(x^3 + y)$. Then $\sigma = \sigma_{\text{T},3}$ satisfies

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

Table 2

Representation of the automaton for the power series $\sigma_{(1,9)}$ of order 4 with break sequence (1, 9). The meaning of the representation by a set of triples is found in Proposition 5.3.1: a triple (ℓ, i, j) represents a vertex (the first one on the list being the start vertex) with vertex label ℓ , a directed edge with label 0 to the i -th triple, and with label 1 to the j -th triple.

$((0, 2, 3), (0, 7, 8), (1, 3, 69), (0, 5, 6), (0, 12, 13), (1, 85, 72), (0, 20, 16), (1, 88, 89), (0, 10, 11), (0, 24, 37),$
 $(1, 84, 74), (0, 24, 40), (1, 104, 76), (0, 15, 16), (0, 20, 37), (1, 8, 72), (0, 18, 19), (0, 30, 13), (1, 73, 74),$
 $(0, 48, 49), (0, 22, 23), (0, 48, 60), (1, 91, 92), (0, 20, 25), (1, 81, 94), (0, 27, 6), (0, 35, 28), (0, 29, 19),$
 $(0, 62, 28), (0, 31, 9), (0, 22, 32), (1, 88, 78), (0, 34, 11), (0, 31, 49), (0, 7, 36), (1, 105, 52), (1, 23, 61),$
 $(0, 7, 39), (1, 106, 14), (1, 95, 17), (0, 42, 36), (0, 20, 40), (0, 22, 36), (0, 45, 46), (0, 31, 60), (1, 40, 55),$
 $(0, 22, 39), (0, 50, 51), (0, 35, 54), (0, 50, 50), (0, 48, 9), (0, 53, 54), (0, 21, 50), (0, 59, 57), (0, 49, 56),$
 $(0, 51, 50), (0, 58, 57), (0, 62, 54), (0, 66, 4), (0, 38, 4), (0, 60, 56), (0, 21, 43), (0, 30, 54), (0, 43, 50),$
 $(0, 21, 51), (0, 21, 7), (0, 65, 4), (0, 47, 9), (1, 8, 98), (1, 70, 71), (1, 100, 92), (1, 75, 76), (1, 97, 98),$
 $(1, 79, 78), (1, 81, 82), (1, 69, 76), (1, 70, 78), (1, 83, 78), (1, 77, 80), (1, 102, 72), (1, 88, 90), (1, 103, 74),$
 $(1, 70, 89), (1, 39, 82), (1, 91, 80), (1, 77, 87), (1, 93, 94), (1, 79, 64), (1, 99, 52), (1, 101, 14), (1, 79, 68),$
 $(1, 36, 61), (1, 106, 68), (1, 107, 63), (1, 91, 96), (1, 108, 17), (1, 106, 41), (1, 16, 55), (1, 77, 92), (1, 70, 90),$
 $(1, 77, 96), (1, 106, 64), (1, 109, 26), (1, 23, 26), (1, 86, 52), (1, 86, 67), (1, 110, 17), (1, 105, 67), (1, 105, 44),$
 $(1, 105, 33))$



and $\tau = \sigma_{T,4}$ satisfies the same equation as σ (it turns out that another solution is $\sigma_{T,3}^{\circ 2} = \sigma_{T,4}^{\circ 2}$). Solving this (irreducible) equation by automata, we find that σ and τ correspond to the bottom left and bottom right automaton depicted in Table 4.

Table 3

Four explicit power series of order 4 with break sequence (1, 3) (the representation is minimal in the sense that no monomial occurs twice in the same formula).

$\sigma_{T,1} = t + \sum_{k \geq 2} \left(t^{2^k-2} + t^{2 \cdot 2^k-1} + t^{4 \cdot 2^k-5} \right) + \sum_{k, \ell \geq 2} t^{2^k(2^\ell-3)+1} = t + t^2 + O(t^5).$
$\sigma_{T,2} = t + t^2 + \sum_{k \geq 3} \left(t^{2^k-4} + t^{2^k-3} + t^{2^k-1} + t^{4 \cdot 2^k-6} + t^{4 \cdot 2^k-5} + t^{8 \cdot 2^k-22} + t^{8 \cdot 2^k-21} \right) +$ $(t+1) \sum_{k, \ell \geq 3} t^{2^k(2^\ell-6)+2} + (t+1) \sum_{k, \ell, m \geq 2} t^{2^{k+\ell}(2^m-3)+2 \cdot 2^k-2} = t + t^2 + t^4 + t^5 + O(t^7).$
$\sigma_{T,3} = t + t^8 + t^{44} + \sum_{k \geq 2} \left(t^{2^k-2} + t^{3 \cdot 2^k-2} + t^{8 \cdot 2^k-4} + t^{8 \cdot 2^k+4} + t^{8 \cdot 2^k+20} + t^{16 \cdot 2^k+44} + t^{24 \cdot 2^k-4} \right) +$ $\sum_{k, \ell \geq 2} \left(t^{2^k(2^\ell+3)-2} + t^{4 \cdot 2^k(2^\ell+2)+4} + t^{8 \cdot 2^k(2^\ell+3)-4} + t^{8 \cdot 2^k(2^\ell+2)+12} \right) +$ $\sum_{k, \ell \geq 2, m \geq 1} \left(t^{2^{k+\ell}(2^m+1)+2^k-2} + t^{8 \cdot 2^{k+\ell}(2^m+1)+8 \cdot 2^k-4} \right) = t + t^2 + t^6 + t^8 + t^{10} + O(t^{13}).$
$\sigma_{T,4} = t + t^4 + t^8 + t^{20} + \sum_{k \geq 2} \left(t^{2^k-2} + t^{8 \cdot 2^k-4} + t^{8 \cdot 2^k+20} + t^{16 \cdot 2^k+12} + t^{16 \cdot 2^k+44} \right) +$ $\sum_{k, \ell \geq 2} \left(t^{2^k(2^\ell+1)-2} + t^{8 \cdot 2^k(2^\ell+1)-4} + t^{4 \cdot 2^k(2^\ell+2)+4} + t^{8 \cdot 2^k(2^\ell+2)+12} + t^{2^k(2^\ell+3)-2} + t^{8 \cdot 2^k(2^\ell+3)-4} \right) +$ $\sum_{k, \ell \geq 2, m \geq 1} \left(t^{2^{k+\ell}(2^m+1)+2^k-2} + t^{8 \cdot 2^{k+\ell}(2^m+1)+8 \cdot 2^k-4} \right) = t + t^2 + t^4 + t^6 + t^8 + O(t^{13}).$

Converting the automata into explicit series as before, we find the formulas in Table 3 (again including the initial coefficients).

By the criterion in Lemma 5.1.1, we see easily that $\sigma_{T,2}$, $\sigma_{T,3}$ and σ_{CS} are conjugate, and so are $\sigma_{T,1}$, $\sigma_{T,4}$ and σ_{CS}^3 .

7. Construction and classification of some order-8 elements

7.1. Order 8, break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$

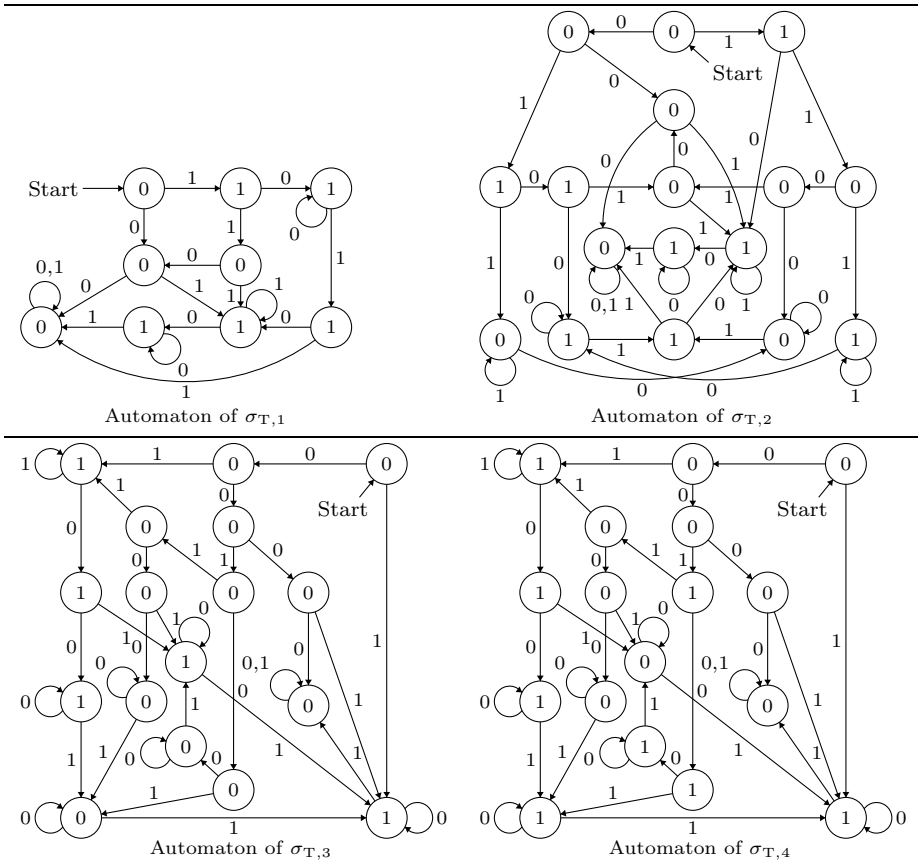
Up to now, no finite description of any element of $\mathcal{N}(\mathbf{F}_2)$ of order 8 was known. Our method produces an example.

Proposition 7.1.1. *An element σ_8 in $\mathcal{N}(\mathbf{F}_2)$ of order 8 with break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$ is given by the automaton described by the data in Table 5: it has 320 states, corresponding to the 320 triples on the displayed ordered list, where the start vertex is the first triple on the list and a triple (ℓ, i, j) occurs on the list precisely if the following three conditions hold: its vertex label is ℓ , there is a directed edge with label 0 to the i -th triple on the list and there is a directed edge with label 1 to the j -th triple on the list. The initial terms of σ_8 are*

$$\sigma_8 = t + t^2 + t^5 + t^6 + t^{12} + O(t^{13}).$$

Table 4

Automata corresponding to the order 4, break sequence (1, 3) series in Table 3.



We refrain from including a pictorial representation, but the automaton is stored in standard Mathematica form in [17], making it easy to manipulate.

Proof. We refer to Example 2.2.3 on how to use Witt vectors of length 3 to construct cyclic order-8 extensions. We choose $\beta = (z^{-1}, 0, 0) \in W_3(\mathbf{F}_2((z)))$ and rewrite the resulting equations in (6) in terms of the variables $x := \alpha_0$, $y := \alpha_1$ and $w := \alpha_2 + \alpha_0^2 \alpha_1$ to find

$$\begin{cases} x^2 + x = z^{-1}; \\ y^2 + y = xz^{-1}; \\ w^2 + w = x^4y + x^3y. \end{cases}$$

Choosing uniformisers $z_0 = z, z_1, z_2, z_3$ for the intermediate fields in the tower of field extensions

$$K_0 = \mathbf{F}_2((z)) \subsetneq K_1 = K_0(x) = \mathbf{F}_2((z_1)) \subsetneq K_2 = K_1(y) = \mathbf{F}_2((z_2)) \subsetneq K_3 = K_2(w) \\ = \mathbf{F}_2((z_3))$$

and using Lemma 2.3.1 as in Example 2.3.2, we see that the extension K_3/K_0 is totally ramified. We find the relevant valuations (following Lemma 2.3.1):

$$\begin{aligned} v_{z_1}(x) &= -1, \quad v_{z_1}(z) = 2; \\ v_{z_2}(y) &= -3; \quad v_{z_2}(x) = -2, \quad v_{z_2}(z) = 4; \\ v_{z_3}(w) &= -11, \quad v_{z_3}(x) = -4, \quad v_{z_3}(y) = -6, \quad v_{z_3}(z) = 8. \end{aligned}$$

We choose the uniformiser t as $t = (w + y)/(x^3 + y)$. Then indeed $v_{z_3}(t) = 1$, and the action of the generator of the Galois group on α_i is given by (7), which implies that for our choice of variables we have

$$\begin{cases} \sigma(x) = x + 1; \\ \sigma(y) = y + x; \\ \sigma(w) = w + xy + y, \end{cases} \quad (20)$$

and so by elimination we find the (irreducible) equation

$$t^6 X^6 + (t^6 + t^2)X^4 + (t^6 + t^5 + t^4 + t^3 + t^2 + 1)X^2 + (t + 1)^3 X + t^6 + t^5 + t^2 + t = 0$$

for $\sigma = \sigma_8$. The initial coefficients are as indicated, and we readily verify the lower break sequence $(1, 3, 11)$ from

$$\sigma_8 = t + t^2 + O(t^3), \quad \sigma_8^2 = t + t^4 + O(t^5), \quad \sigma_8^4 = t + t^{12} + O(t^{13}). \quad \square$$

7.2. Detecting conjugacy using local class field theory

Proposition 7.2.1. *The number of conjugacy classes of elements of order 8 in $\mathcal{N}(\mathbf{F}_2)$ with ‘minimal’ break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$ is 4.*

Proof. We follow the method of Lubin [54]. For $k \geq 1$, write U_k for the multiplicative group of units $U_k = 1 + z^k \mathbf{F}_2[[z]]$. By [54, Thm. 2.2] elements of exact order 2^n in $\mathcal{N}(\mathbf{F}_2)$ up to conjugation correspond bijectively to continuous surjective characters $\eta: U_1 \rightarrow \mathbf{Z}/2^n \mathbf{Z}$ up to so-called strict equivalence (the bijection arises from the restriction of the local reciprocity map to U_1). Strict equivalence of characters η and η' means that there exists $u \in \mathcal{N}(\mathbf{F}_2)$ with $\eta(u(z)/z) = 0$ and $\eta'(x) = \eta(x \circ u)$ for all $x \in U_1$. Moreover, the upper break sequence $\langle b^{(0)}, \dots, b^{(n-1)} \rangle$ can be read off from the corresponding character: $\eta(U_{b^{(i)}}) = 2^i \mathbf{Z}/2^n \mathbf{Z}$ and $\eta(U_{b^{(i)}+1}) = 2^{i+1} \mathbf{Z}/2^n \mathbf{Z}$ [54, Prop. 3.2].

In our case of order 8 elements with minimal break sequence, this implies that the corresponding characters factor through U_1/U_5 and map U_3 to $4\mathbf{Z}/8\mathbf{Z}$. Since we have an isomorphism of groups

$$\begin{aligned}\mathbf{Z}/8\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} &\rightarrow U_1/U_5 \\ (c, d) &\mapsto (1+z)^c(1+z^3)^d U_5,\end{aligned}$$

there are eight such characters $\eta_{a,b}$ determined by $\eta_{a,b}(1+z) = a \in (\mathbf{Z}/8\mathbf{Z})^*$ and $\eta_{a,b}(1+z^3) = 4b$ with $b \in \mathbf{Z}/2\mathbf{Z}$. We need to determine which of these are strictly equivalent. Write any $u \in \mathcal{N}(\mathbf{F}_2)$ in the form $u(z) = z(1+z)^\alpha(1+z^3)^\beta u_5$ with $\alpha \in \{0, \dots, 7\}$, $\beta \in \{0, 1\}$ and $u_5 \in U_5$. We have $\eta_{a,b}(u(z)/z) = a\alpha + 4b\beta \pmod 8$, and hence $\eta_{a,b}(u(z)/z) = 0$ if and only if $u(z) \equiv z \pmod{z^6}$ or $u(z) \equiv z + z^4 + bz^5 \pmod{z^6}$. Suppose then that

$$\eta_{a',b'}(x) = \eta_{a,b}(x \circ u), \quad (21)$$

and evaluate both sides for $x = 1+z$ and $x = 1+z^3$, respectively. For the first choice of u , we immediately find that $a' = a$ and $b' = b$. For the second choice of u , for $x = 1+z$, the left hand side of (21) evaluates to a' and the right hand side to $\eta_{a,b}((1+z)^5 U_5) = 5a$. For $x = 1+z^3$, the left hand side is b' and the right hand side $\eta_{a,b}((1+z^3)U_5) = b$.

We conclude that the strict equivalence class of $\eta_{a,b}$ consists of $\eta_{a,b}$ and $\eta_{5a,b}$, and there are indeed four strict equivalence classes in total. \square

We state below an analogue of Lemma 5.1.1 that allows us to distinguish between these four conjugacy classes based on the first few coefficients of the power series.

Proposition 7.2.2. *Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ be an automorphism of order 8 with break sequence given by $(1, 3, 11) = \langle 1, 2, 4 \rangle$, and write $\sigma = \sum_{i=1}^{\infty} a_i t^i$ with $a_i \in \mathbf{F}_2$. Then $a_1 = a_2 = 1$, $a_3 = 0$, $a_5 \neq a_7$, and σ is conjugate to a series $\sigma_{8,(b_4,b_{11})}$ of order 8 that has initial coefficients*

$$\sigma_{8,(b_4,b_{11})} = t + t^2 + b_4 t^4 + t^7 + b_{11} t^{11} + O(t^{12})$$

for a unique choice of $b_4, b_{11} \in \mathbf{F}_2$. In particular, the conjugacy class of σ depends only on $\sigma \pmod{t^{12}}$.

The series σ_8 is conjugate to $\sigma_{8,(1,1)}$ and $\sigma_8^{\circ 3}$ is conjugate to $\sigma_{8,(0,1)}$. These give representatives of two of the four conjugacy classes of minimally ramified series of order 8.

Proof. We will show that any such σ is conjugate to some $\sigma_{8,(b_4,b_{11})}$ modulo t^{12} , and that the series $\sigma_{8,(b_4,b_{11})}$ are not conjugate modulo t^{12} for the four different choices of (b_4, b_{11}) . Since we know that there are 4 conjugacy classes of series σ satisfying the required assumptions, this shows that actual series $\sigma_{8,(b_4,b_{11})}$ of order 8 with minimal break sequence do exist.

We first note that $d(\sigma) = 1$ implies $a_1 = a_2 = 1$; computing $\sigma^{\circ 2}$, we get $\sigma^{\circ 2} = t + (1+a_3)t^4 + O(t^5)$, and $d(\sigma^{\circ 2}) = 3$ gives $a_3 = 0$; finally, $\sigma^{\circ 4} = t + (a_5 + a_7)t^{12} + O(t^{13})$, and since $d(\sigma^{\circ 4}) = 11$, we get $a_5 \neq a_7$.

Table 5

Representation of the automaton for the power series σ_8 of order 8 with break sequence (1, 3, 11). The meaning of the representation by a set of triples is found in Proposition 7.1.1: a triple (ℓ, i, j) represents a vertex (the first one on the list being the start vertex) with vertex label ℓ , a directed edge with label 0 to the i -th triple, and with label 1 to the j -th triple.

$((0, 2, 3), (0, 58, 59), (1, 82, 185), (0, 5, 3), (0, 65, 66), (0, 7, 8), (0, 136, 137), (1, 278, 43), (0, 10, 11), (0, 140, 141),$ $(1, 281, 43), (0, 13, 8), (0, 147, 38), (0, 15, 11), (0, 151, 152), (0, 17, 18), (0, 76, 77), (1, 279, 117), (0, 20, 18),$ $(0, 78, 79), (0, 22, 23), (0, 60, 61), (1, 280, 117), (0, 25, 23), (0, 70, 72), (0, 9, 27), (1, 89, 190), (0, 24, 27), (0, 30, 31),$ $(0, 44, 41), (1, 87, 190), (0, 32, 31), (0, 32, 34), (1, 72, 189), (0, 33, 36), (1, 224, 160), (0, 33, 38), (1, 214, 154),$ $(0, 40, 41), (0, 51, 109), (1, 84, 185), (0, 43, 3), (0, 115, 116), (0, 96, 101), (0, 46, 3), (0, 80, 68), (0, 35, 48),$ $(1, 272, 112), (0, 37, 50), (1, 290, 45), (0, 35, 8), (0, 37, 53), (1, 282, 39), (0, 55, 18), (0, 99, 100), (0, 57, 27),$ $(0, 142, 143), (0, 60, 93), (1, 238, 128), (0, 60, 106), (1, 236, 87), (0, 63, 64), (0, 58, 112), (1, 265, 48), (0, 151, 129),$ $(1, 242, 66), (0, 65, 68), (1, 260, 59), (0, 70, 71), (0, 151, 179), (1, 293, 305), (1, 231, 97), (0, 74, 75), (0, 95, 199),$ $(1, 232, 97), (0, 51, 26), (1, 268, 48), (0, 44, 192), (1, 267, 143), (0, 81, 82), (0, 195, 154), (1, 256, 143), (0, 81, 84),$ $(1, 240, 66), (0, 153, 28), (1, 71, 189), (0, 153, 42), (1, 215, 154), (0, 131, 134), (1, 225, 160), (0, 129, 27), (0, 55, 164),$ $(0, 94, 89), (0, 111, 107), (0, 96, 97), (0, 125, 123), (1, 23, 34), (0, 99, 79), (0, 125, 128), (1, 251, 273), (1, 222, 220),$ $(0, 103, 101), (0, 52, 114), (0, 105, 89), (0, 85, 89), (0, 4, 107), (1, 3, 221), (0, 54, 109), (1, 221, 221), (0, 56, 107),$ $(0, 129, 130), (0, 113, 114), (0, 138, 139), (1, 262, 303), (0, 131, 132), (1, 266, 303), (0, 118, 116), (0, 148, 149),$ $(0, 120, 114), (0, 150, 50), (0, 62, 68), (0, 73, 123), (1, 264, 59), (0, 90, 123), (0, 126, 127), (0, 126, 75), (1, 296, 305),$ $(1, 240, 66), (0, 153, 28), (1, 71, 189), (0, 153, 42), (1, 215, 154), (0, 131, 134), (1, 225, 160), (0, 129, 27), (0, 55, 164),$ $(1, 213, 156), (0, 98, 172), (1, 288, 42), (0, 63, 16), (1, 291, 45), (0, 63, 6), (1, 270, 112), (0, 140, 145), (1, 283, 39),$ $(0, 142, 11), (0, 52, 122), (0, 193, 187), (1, 212, 156), (0, 49, 104), (0, 148, 194), (1, 289, 42), (0, 58, 124), (0, 155, 121),$ $(0, 95, 203), (0, 157, 121), (0, 33, 206), (0, 159, 46), (0, 69, 179), (0, 161, 46), (0, 144, 182), (0, 163, 122), (0, 47, 19),$ $(0, 165, 122), (0, 99, 21), (0, 167, 124), (0, 133, 106), (0, 169, 124), (0, 83, 40), (0, 171, 26), (0, 176, 29), (0, 173, 28),$ $(0, 65, 46), (0, 175, 26), (0, 184, 26), (0, 44, 189), (0, 178, 32), (0, 142, 168), (0, 86, 32), (0, 181, 29), (0, 210, 192),$ $(0, 183, 29), (0, 211, 186), (0, 51, 45), (0, 154, 186), (0, 191, 192), (0, 129, 188), (0, 194, 192), (0, 179, 188),$ $(0, 162, 186), (0, 198, 164), (0, 187, 193), (0, 193, 193), (0, 205, 166), (0, 148, 191), (0, 197, 162), (0, 67, 194),$ $(0, 98, 177), (0, 200, 164), (0, 37, 180), (0, 202, 162), (0, 146, 208), (0, 204, 166), (0, 126, 108), (0, 49, 110),$ $(0, 207, 168), (0, 135, 16), (0, 209, 168), (0, 88, 24), (0, 55, 156), (0, 52, 119), (1, 319, 100), (1, 294, 313),$ $(1, 310, 71), (1, 316, 308), (1, 212, 164), (1, 218, 172), (1, 319, 79), (1, 218, 177), (1, 68, 187), (1, 66, 187),$ $(1, 223, 158), (1, 318, 36), (1, 311, 145), (1, 304, 284), (1, 227, 203), (1, 297, 97), (1, 227, 199), (1, 230, 192),$ $(1, 297, 101), (1, 230, 189), (1, 233, 190), (1, 233, 31), (1, 235, 154), (1, 244, 72), (1, 237, 160), (1, 241, 141),$ $(1, 239, 196), (1, 271, 122), (1, 241, 170), (1, 257, 16), (1, 243, 129), (1, 248, 194), (1, 243, 179), (1, 246, 162),$ $(1, 248, 191), (1, 246, 154), (1, 249, 187), (1, 249, 193), (1, 216, 199), (1, 252, 201), (1, 258, 42), (1, 217, 172),$ $(1, 255, 174), (1, 277, 104), (1, 257, 6), (1, 260, 112), (1, 260, 124), (1, 258, 28), (1, 261, 93), (1, 261, 106),$ $(1, 263, 12), (1, 292, 26), (1, 244, 44), (1, 228, 102), (1, 229, 14), (1, 247, 9), (1, 269, 104), (1, 294, 12), (1, 302, 222),$ $(1, 317, 53), (1, 312, 134), (1, 271, 119), (1, 275, 119), (1, 218, 286), (1, 277, 110), (1, 317, 50), (1, 309, 84),$ $(1, 277, 306), (1, 320, 127), (1, 315, 314), (1, 226, 128), (1, 295, 303), (1, 285, 30), (1, 245, 87), (1, 287, 30),$ $(1, 258, 307), (1, 249, 221), (1, 259, 130), (1, 276, 48), (1, 219, 274), (1, 223, 8), (1, 292, 45), (1, 232, 48),$ $(1, 294, 19), (1, 297, 39), (1, 280, 123), (1, 299, 112), (1, 252, 132), (1, 301, 43), (1, 247, 82), (1, 242, 68),$ $(1, 250, 128), (1, 244, 71), (1, 152, 130), (1, 298, 307), (1, 284, 130), (1, 300, 307), (1, 245, 89), (1, 247, 84),$ $(1, 241, 145), (1, 256, 11), (1, 253, 274), (1, 254, 274), (1, 259, 27), (1, 252, 134), (1, 318, 38), (1, 233, 34),$ $(1, 280, 128), (1, 320, 75))$

We will now prove that σ is conjugate to $\sigma_{8,(b_4,b_{11})}$ for some $b_4, b_{11} \in \mathbf{F}_2$. We do this by conjugating with selected elements of $\mathcal{N}(\mathbf{F}_2)$ in the following steps (in each step the symbols a_i denote the coefficients of the ‘new’ power series, obtained by performing the conjugations described in the previous steps):

Step I (conjugating with $\chi_3: t \mapsto t + t^3$). We have $\chi_3^{\circ-1} = t + t^3 + t^5 + t^9 + t^{11} + O(t^{12})$, yielding

$$\chi_3 \circ \sigma \circ \chi_3^{\circ-1} = t + t^2 + (1 + a_4)t^4 + (1 + a_5)t^5 + O(t^6),$$

so conjugating if necessary by χ_3 we may and do assume that $a_5 = 0$; then $a_7 = 1$, since $a_5 \neq a_7$.

Step II (conjugating with $\chi_5: t \mapsto t + t^5$). We have $\chi_5^{\circ-1} = t + t^5 + t^9 + O(t^{12})$, yielding

$$\chi_5 \circ \sigma \circ \chi_5^{\circ-1} = t + t^2 + a_4 t^4 + (1 + a_6) t^6 + O(t^7),$$

so conjugating if necessary by χ_5 we may and do assume that $a_6 = 0$.

Step III (conjugating with $\chi_2: t \mapsto t + t^2$). We have $\chi_2^{\circ-1} = t + t^2 + t^4 + t^8 + O(t^{12})$, yielding

$$\chi_2 \circ \sigma \circ \chi_2^{\circ-1} = t + t^2 + a_4 t^4 + t^7 + (1 + a_8) t^8 + (1 + a_9) t^9 + (a_9 + a_{10}) t^{10} + (1 + a_{11}) t^{11} + O(t^{12}),$$

so conjugating if necessary by χ_2 we may and do assume that $a_9 = 0$.

Step IV (conjugating with $\chi_6: t \mapsto t + t^6$). We have $\chi_6^{\circ-1} = t + t^6 + O(t^{12})$, yielding

$$\chi_6 \circ \sigma \circ \chi_6^{\circ-1} = t + t^2 + a_4 t^4 + t^7 + (1 + a_8) t^8 + (1 + a_{10}) t^{10} + a_{11} t^{11} + O(t^{12}),$$

so conjugating if necessary by χ_6 we may and do assume that $a_8 = 0$.

Step V (conjugating with $\chi_4: t \mapsto t + t^4$). We have $\chi_4^{\circ-1} = t + t^4 + O(t^{12})$, yielding

$$\chi_4 \circ \sigma \circ \chi_4^{\circ-1} = t + t^2 + a_4 t^4 + t^7 + (1 + a_{10}) t^{10} + a_{11} t^{11} + O(t^{12}),$$

so conjugating if necessary by χ_4 we may and do assume that $a_{10} = 0$.

This ends the proof that σ is conjugate to $\sigma_{8,(b_4,b_{11})}$ for some $b_4, b_{11} \in \mathbf{F}_2$.

We will now prove that the power series $\sigma_{8,(b_4,b_{11})}$ and $\sigma_{8,(c_4,c_{11})}$ are not conjugate in $\mathcal{N}(\mathbf{F}_2)$ unless $(b_4, b_{11}) = (c_4, c_{11})$. Indeed, suppose that $\sigma_{8,(b_4,b_{11})}$ and $\sigma_{8,(c_4,c_{11})}$ are conjugate, and let $\tau \in \mathcal{N}(\mathbf{F}_2)$ be a conjugating power series, so that $\sigma_{8,(b_4,b_{11})} \circ \tau = \tau \circ \sigma_{8,(c_4,c_{11})}$. Write $\tau = t + \sum_{i=2}^{\infty} d_i t^i$. Computing $\sigma_{8,(b_4,b_{11})} \circ \tau - \tau \circ \sigma_{8,(c_4,c_{11})}$, we get

$$\begin{aligned} & \sigma_{8,(b_4,b_{11})} \circ \tau - \tau \circ \sigma_{8,(c_4,c_{11})} \\ &= (d_3 + b_4 + c_4) t^4 + d_3 t^5 + (d_5 + d_3 c_4) t^6 + \\ & \quad (d_2 + d_6 + d_7 + d_2 b_4 + d_2 c_4 + d_3 c_4 + d_5 c_4) t^8 + (d_2 + d_5 + d_7 + d_3 c_4) t^9 + \\ & \quad (d_2 + d_4 + d_6 + d_7 + d_9 + d_3 c_4 + d_7 c_4) t^{10} + (d_2 + d_2 d_3 + d_7 + b_{11} + c_{11}) t^{11} + O(t^{12}). \end{aligned}$$

Considering the coefficients at t^5 , t^6 and t^9 gives $d_3 = d_5 = d_2 + d_7 = 0$; looking then at the coefficients at t^4 and t^{11} gives $b_4 = c_4$ and $b_{11} = c_{11}$.

Applying the algorithm from the above proof, we find that σ_8 is conjugate to $\sigma_{8,(1,1)}$ and $\sigma_8^{\circ 3}$ is conjugate to $\sigma_{8,(0,1)}$. (This requires computing more coefficients than we have specified in Steps I and II, but the computations are easy.) \square

Corollary 7.2.3. *Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ be an automorphism of order 8 with break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$. Then σ and $\sigma^{\circ 5}$ are conjugate in $\mathcal{N}(\mathbf{F}_2)$, while σ and $\sigma^{\circ 3}$ are not.*

Proof. This follows from the proof of Proposition 7.2.1—if an element σ corresponds to the character $\eta_{a,b}$, then for k odd the element $\sigma^{\circ k}$ corresponds to $k\eta_{a,b} = \eta_{ka,kb} = \eta_{ka,b}$. Since $\eta_{a,b}$ and $\eta_{5a,b}$ are strictly equivalent, while $\eta_{a,b}$ and $\eta_{3a,b}$ are not, the claim follows.

It is also possible to give a direct proof using the method of Proposition 7.2.2, as follows. Denote the relation of being conjugate by \sim . By Proposition 7.2.2, we may assume without loss of generality that $\sigma = t + t^2 + b_4 t^4 + t^7 + b_{11} t^{11} + O(t^{12})$ for some $b_4, b_{11} \in \mathbf{F}_2$. Then

$$\sigma^{\circ 2} = t + t^4 + t^8 + t^9 + (1 + b_4)t^{10} + t^{11} + O(t^{12}), \quad \sigma^{\circ 4} = t + O(t^{12}),$$

and hence $\sigma = \sigma^{\circ 5} + O(t^{12})$ and

$$\sigma^{\circ 3} = t + t^2 + (1 + b_4)t^4 + t^7 + t^9 + b_4 t^{10} + (1 + b_{11})t^{11} + O(t^{12}).$$

Following the algorithm of the proof of Proposition 7.2.2 (and using the notation therein), we may conjugate $\sigma^{\circ 3}$ in turn by χ_2, χ_6 and in the case where $b_4 = 1$ also χ_4 to arrive at

$$\sigma^{\circ 3} \sim t + t^2 + (1 + b_4)t^4 + t^7 + b_{11}t^{11} + O(t^{12}),$$

i.e. if $\sigma \sim \sigma_{8,(b_4,b_{11})}$, then $\sigma^{\circ 3} \sim \sigma_{8,(b_4+1,b_{11})}$. Applying Proposition 7.2.2 again shows that $\sigma \sim \sigma^{\circ 5}$ and $\sigma \approx \sigma^{\circ 3}$. \square

7.3. Finding representatives via explicit class field theory

We have already constructed representatives of two out of four conjugacy classes of minimally ramified series of order 8. In order to construct the representatives for the remaining conjugacy classes, we will extend the method using the Carlitz module from Remark 5.1.3.

Let ρ be the Carlitz module for $K = \mathbf{F}_2(z)$. We know from [54, Obs. 4 & Sect. 5] that the characters $\eta: U_1 \rightarrow \mathbf{Z}/8\mathbf{Z}$ corresponding to minimally ramified order-8 elements factor through U_5 , and the corresponding Galois extensions can be obtained as a subextension of $K(\rho[z^5])/K$. The extension $K(\rho[z^5])/K$ has Galois group

$$G = (\mathbf{F}_2[z]/z^5)^* \cong \mathbf{Z}/8\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} = \langle z + 1 \bmod z^5 \rangle \times \langle z^3 + 1 \bmod z^5 \rangle.$$

The group G has two subgroups with quotient $\mathbf{Z}/8\mathbf{Z}$:

$$H_1 = \langle z^3 + 1 \bmod z^5 \rangle \quad \text{and} \quad H_2 = \langle z^4 + z^3 + 1 \bmod z^5 \rangle.$$

The field $K(\rho[z^5])$ is generated by a root α of the degree-16 polynomial $\rho_{z^5}(X)/\rho_{z^4}(X)$. The fixed fields L_1 and L_2 of H_1 and H_2 , respectively, are generated by the elements

$$\beta_1 := \alpha \cdot \rho_{z^3+1}(\alpha) \quad \text{and} \quad \beta_2 := \alpha \cdot \rho_{z^4+z^3+1}(\alpha).$$

Recalling that L_i/K has Galois group cyclic of order 8 generated by σ acting as $\sigma(\alpha) = z\alpha + \alpha + \alpha^2$, we can compute $\sigma(\beta_i)$ and we find that

$$\begin{cases} \beta_1 = \alpha^9 + (z^4 + z^2 + z)\alpha^5 + (z^4 + z^3 + z^2)\alpha^3 + (z^3 + 1)\alpha^2; \\ \sigma(\beta_1) = \alpha^{10} + (z + 1)\alpha^9 + (z^4 + z^2 + z)\alpha^6 + (z^5 + z^4 + z^3 + z)\alpha^5 + \\ (z^4 + z^3 + z^2 + 1)\alpha^4 + (z^5 + z^3 + z^2)\alpha^3 + (z^4 + z^3 + z^2 + z + 1)\alpha^2 + \\ (z^2 + z)\alpha; \end{cases}$$

and

$$\begin{cases} \beta_2 = \alpha^9 + (z^4 + z^2 + z)\alpha^5 + (z^4 + z^3 + z^2)\alpha^3 + (z^3 + 1)\alpha^2 + z\alpha; \\ \sigma(\beta_2) = \alpha^{10} + (z + 1)\alpha^9 + (z^4 + z^2 + z)\alpha^6 + (z^5 + z^4 + z^3 + z)\alpha^5 + \\ (z^4 + z^3 + z^2 + 1)\alpha^4 + (z^5 + z^3 + z^2)\alpha^3 + (z^4 + z^3 + z^2 + z + 1)\alpha^2 + \\ (z^2 + z)\alpha. \end{cases}$$

Since z is the only ramified place and it is totally ramified in $K(\rho[z^5])$, the same is true in L_i . We can choose $t = \beta_i$ as a uniformiser for the place above z in L_i . Elimination of z and α leads to the following equation for the element $\sigma_{8,1} = \sigma_{8,1}(t)$ of order 8 with $t = \beta_1$:

$$\begin{aligned} tX^6 + (t+1)X^5 + (t^5 + t^3 + t)X^4 + (t^5 + t^2 + t)X^3 + \\ (t^6 + t^3 + t)X^2 + t^4X + t^6 + t^5 + t^4 + t^3 = 0; \end{aligned}$$

and to the following equation for the element $\sigma_{8,2} = \sigma_{8,2}(t)$ of order 8 with $t = \beta_2$:

$$\begin{aligned} tX^6 + (t+1)X^5 + (t^5 + t^3)X^4 + (t^5 + t + 1)X^3 + \\ (t^6 + t^5 + t^4 + t^3 + t)X^2 + (t^4 + t^2)X + t^4 + t^3 = 0. \end{aligned}$$

These equations define algebraic curves of geometric genus 7, solved by the series

$$\sigma_{8,1}(t) = t + t^2 + t^5 + t^{11} + O(t^{13}) \quad \text{and} \quad \sigma_{8,2}(t) = t + t^2 + t^5 + t^9 + t^{11} + O(t^{13})$$

of order 8, which are produced by automata with 668 and 926 states, respectively. Furthermore, $\sigma_{8,1}$ is conjugate to $\sigma_{8,(1,1)}$ and $\sigma_{8,2}$ is conjugate to $\sigma_{8,(1,0)}$ by the method from Proposition 7.2.2. We may summarise the above discussion as follows:

Proposition 7.3.1. *There are four conjugacy classes of order-8 elements with break sequence $(1, 3, 11) = \langle 1, 2, 4 \rangle$ and their representatives are the series $\sigma_{8,1}$, $\sigma_{8,1}^{\circ 3}$ (conjugate to σ_8 and $\sigma_8^{\circ 3}$, respectively), $\sigma_{8,2}$ and $\sigma_{8,2}^{\circ 3}$. \square*

The automata and series, also for $\sigma_{8,2}$, may be found in [17].

Remark 7.3.2. We have constructed order-8 elements by using the Galois extension $K(\rho[z^5])/K$ with Galois group $\mathbf{Z}/8\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, and looking at its subextensions L_i/K with Galois group $\mathbf{Z}/8\mathbf{Z}$. We could instead look at an extension $K(\rho[z^5])/M$ with Galois group $\mathbf{Z}/8\mathbf{Z}$. This would work, but would produce a non-minimally ramified series generated by an automaton with many more states—the automaton corresponding to $\sigma(t) = \rho_{1+z}(t)$ with $t = \alpha$ has 136600 states.

8. Embedding the Klein four-group in $\mathcal{N}(\mathbf{F}_2)$ using automata

Since every p -group embeds in $\mathcal{N}(\mathbf{F}_p)$, we may ask for a representation for generators of a given p -group through automata. We show how to do this for the easiest case, that of the Klein four-group $V = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ for $p = 2$, by describing two automata that correspond to two commuting power series of order two in characteristic two (with minimal admissible break sequences), answering a question that Klopsch asked us.

8.1. Embedding with small conductor

For a general field \mathbf{F} , define the Nottingham group $\mathcal{N}(\mathbf{F})$ to be the group of power series $\sigma(t) \in \mathbf{F}[[t]]$ of the form $t + O(t^2)$ under composition. The following lemma shows that it is easy to embed V into the Nottingham group over any proper field extension \mathbf{F} of \mathbf{F}_2 such that all nontrivial elements of V have break sequence (1) (i.e. have depth 1), but one cannot do so over \mathbf{F}_2 .

Proposition 8.1.1. *There is an embedding of the Klein four-group $V = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ in the Nottingham group $\mathcal{N}(\mathbf{F})$ over a field \mathbf{F} of characteristic two with all nontrivial elements of V having break sequence (1) if and only if $\mathbf{F} \neq \mathbf{F}_2$.*

Note that all nontrivial elements having break sequence (1) means that the corresponding V -extension is *weakly ramified*, i.e. has trivial second ramification group. A much more general statement that implies Lemma 8.1.1 is given in [29, Korollar 3.2], but we give a short direct proof.

Proof. Assume $\mathbf{F} \neq \mathbf{F}_2$ and let U be a two-dimensional \mathbf{F}_2 -vector subspace of \mathbf{F} . Then the power series $t/(ut+1) = t + ut^2 + O(t^3)$ taken over $u \in U$ form a subgroup of $\mathcal{N}(\mathbf{F})$ isomorphic to the Klein four-group.

For the converse, assume we have an embedding of $V = \{\text{id}, \sigma, \tau, \sigma \circ \tau\}$ into $\mathcal{N}(\mathbf{F}_2)$ with nontrivial elements having break sequence (1). Then σ and τ are of the form $t + t^2 + O(t^3)$, implying that $\sigma \circ \tau = t + O(t^3)$, a contradiction. \square

There are further restrictions on possible depths of elements of the Klein four-group embedded in $\mathcal{N}(\mathbf{F}_2)$. In the next subsection, we will construct an embedding with non-

trivial elements having depths 1, 1 and 5. The next lemma shows that these are the minimal possible values.

Proposition 8.1.2. *For every embedding of the Klein four-group V in the Nottingham group $\mathcal{N}(\mathbf{F}_2)$ some nontrivial element of V has depth at least 5.*

Proof. Suppose the contrary. By Proposition 8.1.1 some nontrivial element has depth at least 2. Every element of finite order has odd depth: if σ had even depth, writing $\sigma = t + t^k + O(t^{k+1})$ with k odd, we would find by induction that $\sigma^{\circ 2^n} = t + t^{2^n(k-1)+1} + O(t^{2^n(k-1)+2})$ for all $n \geq 1$, so σ would not be of finite order. Also note that for every $k \geq 1$ the elements of depth at least k form a subgroup. Thus, the only possible sequences of depths < 5 of series in $\mathcal{N}(\mathbf{F}_2)$ representing nontrivial elements of V are 1, 1, 3 and 3, 3, 3. The latter is impossible, since the product of two elements of depth k has depth at least $k + 1$.

It remains to treat the case where the depths of the nontrivial elements are 1, 1, 3. By Klopsch's theorem [48] every element of order 2 and depth 1 is conjugate to $t/(t+1)$, so without loss of generality we may assume that $V = \{\text{id}, \sigma, \tau, \sigma \circ \tau\}$ with

$$\sigma(t) = \frac{t}{t+1} \quad \text{and} \quad \tau(t) = t + t^2 + \sum_{i \geq 3} a_i t^i.$$

We will reach a contradiction by computing up to order $O(t^9)$. We have

$$\begin{aligned} \tau^{\circ 2}(t) &= t + (1 + a_3)t^4 + (a_3a_4 + a_5)t^6 + \\ &\quad (a_3 + a_3a_4 + a_4a_5 + a_6 + a_3a_6 + a_7)t^8 + O(t^9). \end{aligned}$$

Since $\tau^{\circ 2} = \text{id}$, this gives $a_3 = 1$, $a_4 = a_5$, and $a_7 = 1$. Substituting these values allows us to compute

$$\begin{aligned} (\sigma \circ \tau)(t) &= t + (1 + a_4)t^4 + (1 + a_4)t^5 + (a_4 + a_6)t^6 + (1 + a_4)t^7 + \\ &\quad (1 + a_4 + a_6 + a_8)t^8 + O(t^9); \\ (\tau \circ \sigma)(t) &= t + (1 + a_4)t^4 + (1 + a_4)t^5 + (a_4 + a_6)t^6 + (1 + a_4)t^7 + \\ &\quad (a_6 + a_8)t^8 + O(t^9). \end{aligned}$$

Since $\sigma \circ \tau = \tau \circ \sigma$, this gives $a_4 = 1$, and shows that the depth of $\sigma \circ \tau$ is at least 5. \square

8.2. Using automata

We now show how to use automata to embed the Klein four-group V into $\mathcal{N}(\mathbf{F}_2)$. We start with the V -extension $\mathbf{F}_2((z))(x, y)$ of $\mathbf{F}_2((z))$ given by $x^2 + x = z^{-1}$ and $y^2 + y = z^{-3}$ with two generators $\sigma_{V,1}, \sigma_{V,2}$ of V acting as

$$\begin{cases} \sigma_{V,1}(x) = x + 1; \\ \sigma_{V,1}(y) = y \end{cases} \quad \text{and} \quad \begin{cases} \sigma_{V,2}(x) = x; \\ \sigma_{V,2}(y) = y + 1. \end{cases}$$

Since $\sigma_{V,1}, \sigma_{V,2}$ are different, of order two and commute, they generate the group V . Set $w = y + x^3 + x^2 + x$. We may regard $\mathbf{F}_2((z))(x, y)$ as the extension $\mathbf{F}_2((z))(x, y) = \mathbf{F}_2((z))(x, w)$ of $\mathbf{F}_2((z))$ given by

$$\begin{cases} x^2 + x = z^{-1}; \\ w^2 + w = x^5 + x \end{cases}$$

with the two generators $\sigma_{V,1}$ and $\sigma_{V,2}$ acting on x and w as

$$\begin{cases} \sigma_{V,1}(x) = x + 1; \\ \sigma_{V,1}(w) = w + x^2 + x + 1 \end{cases} \quad \text{and} \quad \begin{cases} \sigma_{V,2}(x) = x; \\ \sigma_{V,2}(w) = w + 1. \end{cases}$$

Writing $z_0 = z, z_1, z_2$ for uniformisers of the fields in the tower of field extensions

$$K_0 := \mathbf{F}_2((z)) \subsetneq K_1 = K_0(x) = \mathbf{F}_2((z_1)) \subsetneq K_2 = K_1(w) = \mathbf{F}_2((z_2)),$$

we have $v_{z_1}(x) = -1$, $v_{z_1}(x^5 + x) = -5$, and hence $v_{z_2}(w) = -5$ and $v_{z_2}(x) = -2$. Choosing a uniformiser $t = x^2w^{-1}$ (note that $v_{z_2}(t) = 1$), we find by elimination of the variables z, x, w that $\sigma_{V,1} = \sigma_{V,1}(t)$ and $\sigma_{V,2} = \sigma_{V,2}(t)$ satisfy, respectively,

$$\begin{aligned} t^4X^4 + t^3X^3 + X^2 + (t+1)X + t^2 + t &= 0; \\ (t^4 + 1)X^4 + tX^2 + t^2X + t^4 &= 0. \end{aligned}$$

This is solved with respective initial coefficients

$$\sigma_{V,1} = t + t^2 + O(t^3) \quad \text{and} \quad \sigma_{V,2} = t + t^6 + O(t^7).$$

The corresponding automata have 18 and 14 states, respectively.

Proposition 8.2.1. *The series $\sigma_{V,1}$ and $\sigma_{V,2}$ have break sequences (1) and (5) and satisfy $\sigma_{V,1}^{\circ 2} = \sigma_{V,2}^{\circ 2} = t$ and $\sigma_{V,1} \circ \sigma_{V,2} = \sigma_{V,2} \circ \sigma_{V,1}$, and hence exhibit an explicit embedding of the Klein four-group $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ into $\mathcal{N}(\mathbf{F}_2)$. The corresponding automata are depicted in Table 6. \square*

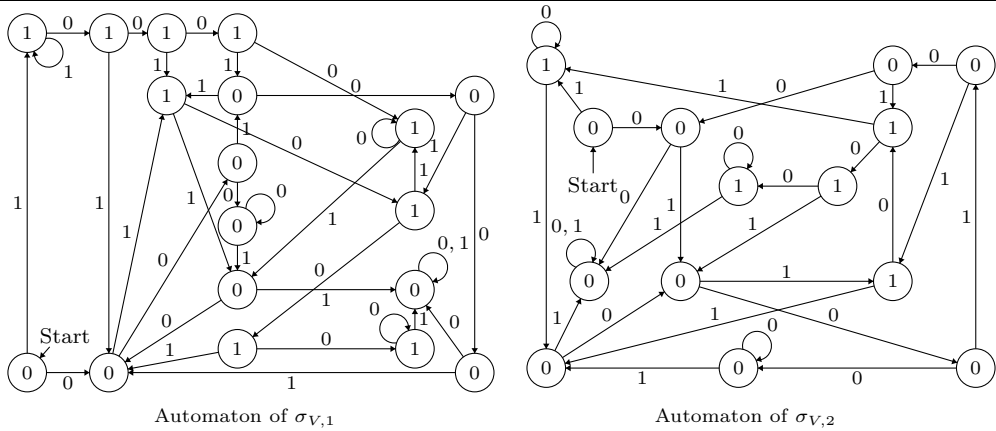
For completeness, writing $\sigma_{V,3} = \sigma_{V,1} \circ \sigma_{V,2}$ for the third nontrivial element of V , we find that $\sigma_{V,3}$ satisfies

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

with initial coefficients $\sigma_{V,3} = t + t^2 + t^3 + O(t^5)$, leading to an automaton with 25 states. The automaton is stored in standard Mathematica form in [17].

Table 6

Automata corresponding to the elements $\sigma_{V,1}$ and $\sigma_{V,2}$ that generate a copy of the Klein four-group in $\mathcal{N}(\mathbf{F}_2)$.



8.3. Other p -groups

In principle, since any finite p -group G can be realised explicitly as the Galois group of an extension of $\mathbf{F}_2((z))$ (this follows from Witt's work; see, e.g. the proof of [20, Theorem 3]), the Galois-theoretic method can be used to find equations satisfied by generators of an embedding of G into $\mathcal{N}(\mathbf{F}_p)$ by algebraic power series, and thus to represent them explicitly by automata. Recall from Remark 2.1.1 that any embedding of G into $\mathcal{N}(\mathbf{F}_p)$ can be conjugated into one in which the elements of G are represented by algebraic power series.

The examples in the current paper do not constitute the computational limit of the method. For example, we can give an embedding of $\mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ into $\mathcal{N}(\mathbf{F}_2)$ with two generators being produced by automata with 128 states, the order-4 element being minimally ramified and the order-2 element having depth 7; we can also obtain an order-9 element in $\mathcal{N}(\mathbf{F}_3)$ with break sequence $(1, 7) = \langle 1, 3 \rangle$ produced by an automaton with 3634 states, etc. However, we refrain from further expanding the catalogue of examples.

As pointed out by the reviewer, it would be interesting to provide explicit automata representing embeddings of generators of other (non-commutative) finite 2-groups in $\mathcal{N}(\mathbf{F}_2)$ (or $\mathcal{N}(\mathbf{F}_{2^m})$ for general m), such as the dihedral or quaternion group of order 8; for this, one again needs to explicitly find a Galois realisation of such groups over $\mathbf{F}_{2^m}((z))$, e.g. by constructing a corresponding Katz–Gabber cover of \mathbf{P}^1 . For the dihedral group, explicit realisations and a study of possible break sequences can be found in [65, §5, §4], at least for sufficiently large m . Also, the quaternion group Q acts by automorphisms (defined over \mathbf{F}_4) on the supersingular elliptic curve in characteristic 2 and stabilises the point at infinity. We did not pursue these lines of thought all the way up to an explicit automatic representation.

A challenge of a completely different level is to consider the question of embedding infinite groups in $\mathcal{N}(\mathbf{F}_p)$ using algebraic power series. Recall that Camina has proven that every countably based pro- p group embeds as subgroup in $\mathcal{N}(\mathbf{F}_p)$; thus, for example, the free pro- p group and the abstract free group on two generators embed. For the latter group, the question is whether we can find two algebraic (i.e. automatic) power series in $\mathcal{N}(\mathbf{F}_2)$ that generated a free group. Another example is the (first) Grigorchuk group, a 2-group with three generators, but without finite presentation (a countable set of relations was given by Lysënok); or other groups given by endomorphic presentations, see [6]. In all these cases, the Galois covering methods break down, and we do not know whether the Grigorchuk group may be realised inside $\mathcal{N}(\mathbf{F}_2)$ with all elements being described solely by algebraic (i.e. automatic) power series, i.e. whether the three generators can simultaneously be conjugated into a set of such algebraic power series (since it is residually finite, this property is true residually, but it is not clear how or whether the property lifts to the entire group).

9. State complexity of automata representing finite order elements in $\mathcal{N}(\mathbf{F}_p)$

9.1. General bounds on state complexity

How ‘complex’ is an automaton that computes a power series $\sigma \in \mathcal{N}(\mathbf{F}_p)$ of given order and break sequence? This is usually measured by ‘state complexity’, i.e. the minimal number of states in an automaton that computes the series.

This complexity can be bounded theoretically. The currently best results arise from the differential forms method described in Section 2: start with an algebraic equation (assumed irreducible) satisfied by $\sigma = \sigma(t)$ with coefficients from $\mathbf{F}_p[t]$, and consider it instead as a two-variable equation $F(t, X) = 0$ describing a (possibly singular) algebraic curve over \mathbf{F}_p . Consider the *degree*

$$d_\sigma := [\mathbf{F}_p(\sigma, t) : \mathbf{F}_p(t)] = \deg_X F$$

and the *height*

$$h_\sigma := [\mathbf{F}_p(\sigma, t) : \mathbf{F}_p(\sigma)] = \deg_t F$$

(the latter two equalities hold by the irreducibility of F), and let g_σ denote the genus of the normalisation \mathcal{X} of the projective curve defined by $F(t, X) = 0$. Bridy has proven that the series σ can be realised by an automaton with less than

$$p^{h_\sigma + 3d_\sigma + g_\sigma - 1}$$

states (see [13, Cor. 3.10], a result that assumes, like this paper, the leading zeros convention, see [13, Remark 2.1]). Concerning the optimality of the upper bound, Bridy has shown in [13, Prop. 3.14] for every $h \geq 1$, there are power series with

Table 7

For each series, we give: its compositional order, lower break sequence, the degree d_σ and genus g_σ of the algebraic equation it satisfies, the theoretical interval $[\lfloor \log_2(d_\sigma + 1) \rfloor, 2^{4d_\sigma + g_\sigma - 1}]$ for the number of states of a minimal automaton and the actual number of states ('?' means we conjecture this to be the correct answer, 'x' means we do not know the answer; see Remark 9.2.2).

series	order	breaks	$d_\sigma = h_\sigma$	g_σ	bounds	# of states
$\sigma_{S,1}$	2	(1)	2	1	$[1, 2^8]$	5
$\sigma_{S,m=2^\mu-1>1}$	2	$\frac{m+1}{2}$	$\frac{m-1}{2}$	$[\mu - 1, 2^{\frac{5m+1}{2}}]$	$\mu + 3$	
$\sigma_{S,m=2^\mu+1}$	2	$m - 1$	$\frac{(m-1)(m-2)}{2}$	$[\mu, 2^{\frac{m^2+5m-8}{2}}]$	$2^\mu + 3^\mu?$	
$\sigma_{K,3}$	2	(3)	3	1	$[2, 2^{12}]$	6
$\sigma_{K,m}$	2	(m)	m	$\frac{(m-1)(m-2)}{2}$	$[\lfloor \log_2(m+1) \rfloor, 2^{\frac{m(m+5)}{2}}]$	x
σ_{CS}^{o2}	2	(3)	2	1	$[1, 2^8]$	7
$\sigma_{V,1}$	2	(1)	4	2	$[2, 2^{17}]$	18
$\sigma_{V,2}$	2	(5)	4	2	$[2, 2^{17}]$	14
$\sigma_{V,3}$	2	(1)	4	2	$[2, 2^{17}]$	25
σ_{\min}	4	(1, 3)	3	1	$[2, 2^{12}]$	5
σ_{CS}	4	(1, 3)	2	1	$[1, 2^8]$	7
σ_{CS}^{o3}	4	(1, 3)	2	1	$[1, 2^8]$	7
σ_J	4	(1, 3)	2	1	$[1, 2^8]$	9
σ_J^{o3}	4	(1, 3)	2	1	$[1, 2^8]$	11
$\sigma_{T,1}$	4	(1, 3)	4	1	$[2, 2^{16}]$	9
$\sigma_{T,2}, \sigma_{T,3}, \sigma_{T,4}$	4	(1, 3)	4	1	$[2, 2^{16}]$	17
$\sigma_{(1,5)}$	4	(1, 5)	3	2	$[2, 2^{13}]$	13
$\sigma_{(1,9)}$	4	(1, 9)	7	4	$[3, 2^{31}]$	110
σ_8	8	(1, 3, 11)	6	7	$[2, 2^{30}]$	320

$d_\sigma = 1, h_\sigma = h, g_\sigma = 0$ that require at least $\geq p^h$ states. A lower bound for the minimal amount of states required to realise the given power series is given by $\log_p(d_\sigma + 1)$ [13, Prop. 2.13]; this bound appears optimal when running over all algebraic power series ([13]).

9.2. Degree equals height for series of finite order in $\mathcal{N}(\mathbf{F}_p)$

In our situation we have the following extra information.

Proposition 9.2.1. *Let $\sigma(t) \in \mathbf{F}_p((t))$ be an algebraic power series over $\mathbf{F}_p(t)$ of finite compositional order. Then $d_\sigma = h_\sigma$.*

Proof. Write n for the compositional order of $\sigma(t)$. The map σ , regarded as an automorphism of $\mathbf{F}_p((t))$, restricts to an automorphism of the field

$$K := \mathbf{F}_p(t, \sigma(t), \sigma^{o2}(t), \dots, \sigma^{o(n-1)}(t)).$$

Since $\sigma(t)$ is algebraic over $\mathbf{F}_p(t)$, successive application of the automorphism σ shows that $\mathbf{F}_p(\sigma^{ok}(t))$ is algebraic over $\mathbf{F}_p(\sigma^{o(k-1)}(t))$ for $k \geq 1$, and hence the extension $K/\mathbf{F}_p(t)$ is algebraic. Since the automorphism σ maps $\mathbf{F}_p(t)$ onto $\mathbf{F}_p(\sigma(t))$, we have $[K : \mathbf{F}_p(t)] = [K : \mathbf{F}_p(\sigma(t))]$, and hence

$$d_\sigma = [\mathbf{F}_p(t, \sigma(t)) : \mathbf{F}_p(t)] = \frac{[K : \mathbf{F}_p(t)]}{[K : \mathbf{F}_p(t, \sigma(t))]} = \frac{[K : \mathbf{F}_p(\sigma(t))]}{[K : \mathbf{F}_p(t, \sigma(t))]}$$

$$= [\mathbf{F}_p(t, \sigma(t)) : \mathbf{F}_p(\sigma(t))] = h_\sigma. \quad \square$$

In Table 7 we give the state complexity for the automata we constructed (where the first two rows refer to series that are considered in the next section), plus the theoretical upper and lower bounds (computed using SINGULAR [31] and MAGMA [11]). We observe that the required number of states is much lower than the (generically almost tight, at least in the genus zero case) upper bounds. The reader may be convinced of this non-generic behaviour by perturbing some of the coefficients in the equation for σ_8 and using [14] to compute the number of states required to solve those perturbed equations (which typically also have higher genus).

Remark 9.2.2. Table 7 lacks a general formula for the minimal number of states in a 2-automaton computing Klopsch's series $\sigma_{K,m}$ for general m . For $m = 1, 3, 5, \dots, 1023$ we computed this in [58] and [14] to be 2, 6, 14, 9, 28, 53, 67, 12, 54, 127, \dots , 30. One may show that for $m = 2^\mu - 1$ such an automaton has 3μ states. We conjecture that for $m = 2^\mu + 1$ it has $3 \cdot 2^\mu + 2\mu - 2$ states. For $m = 2^\mu + 3$, we find the sequence 14, 9, 53, 127, 90, 931, 2675, 770, \dots , which we could not fit into any mould.

10. A hierarchy of complexity of power series based on sparseness

Previously known examples of finite order elements of $\mathcal{N}(\mathbf{F}_2)$ were described as power series having as coefficients binomial coefficients modulo 2 (such as Klopsch's series) or by explicit formulas for the location of the nonzero coefficients (such as the Chinburg–Symonds series σ_{CS} and $\sigma_{CS}^{\circ 3}$). Our automatic description is somewhat different. In this section, we discuss the relation between the existence of ‘closed/explicit formulas’ and properties of the automaton.

10.1. Sparse power series

We propose a definition of a ‘closed formula’ for a power series based on the notion of sparseness (the concept occurs in the literature under various names such as ‘arid’, ‘poly-slender’, ‘polynomial growth’, and ‘bounded’; compare [18, §3]).

Definition 10.1.1. For a power series $\sigma = \sum a_k t^k \in \mathbf{F}_2[[t]]$ over \mathbf{F}_2 , let $E(\sigma)$ denote the *support* of σ , i.e. the set of integers k for which $a_k = 1$. A power series σ (as well as the corresponding automaton and automatic sequence, if they exist) is called *sparse* if

$$\#E(\sigma) \cap \{0, 1, \dots, N\} = O(\log(N)^r)$$

for some $r \geq 0$. The infimum of such r is called the *rank of sparseness* of σ . We say that σ is r -sparse if the rank of sparseness is at most r . If σ is automatic, then this infimum is attained and is an integer (this follows from Proposition 10.1.3 below).

Note that polynomials are sparse, sums of sparse series are sparse, and products of sparse series are sparse. More precisely, if σ is r -sparse and τ is s -sparse, then $\sigma + \tau$ is at most $\max(r, s)$ -sparse and $\sigma\tau$ is at most $(r + s)$ -sparse; this follows from the definition, since $E(\sigma + \tau) \subseteq E(\sigma) \cup E(\tau)$ and $E(\sigma\tau) \subseteq E(\sigma) + E(\tau)$. For automatic sequences, Cobham showed the following dichotomy for the word growth in the associated regular language.

Proposition 10.1.2 (Cobham [26]). *An automatic sequence $\sigma \in \mathbf{F}_2[[t]]$ is either sparse, or $\#E(\sigma) \cap \{0, 1, \dots, N\} \geq N^\alpha$ for some real $\alpha > 0$ and sufficiently large N . \square*

Define a *simple sparse set of rank at most r* to be a set of integers whose base-2 expansion is of the form $v_r w_r^{\ell_r} \cdots v_1 w_1^{\ell_1} v_0$ with $\ell_i \in \mathbf{Z}_{\geq 0}$ for some fixed binary words $v_0, \dots, v_r, w_1, \dots, w_r$.

Proposition 10.1.3 (Szilard, Yu, Zhang and Shallit [63]). *A series σ is automatic and sparse of rank at most r precisely if $E(\sigma)$ is a finite union of pairwise disjoint simple sparse sets of rank at most r .*

Proof. Except for the claim of ‘pairwise disjointness’, this is proven in [63]. The claim that the occurring simple sparse sets can be chosen pairwise disjoint is proven in detail in [18, Cor. 3.10]. \square

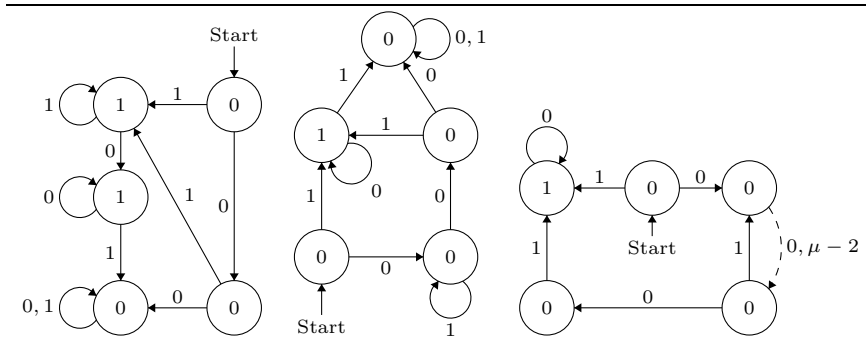
Remark 10.1.4. The proof in [18, Cor. 3.10] is a tedious combinatorial verification. Jason Bell pointed out to us that a much simpler argument is possible if one uses the structure of the corresponding automaton that results from Proposition 12.1.2 below.

Example 10.1.5. The support of σ_{CS}^3 is $E(\sigma_{\text{CS}}^3) = \{3 \cdot 2^k - 2 \mid k \geq 0\} \cup \{4 \cdot 2^k - 2 \mid k \geq 0\}$, and consists of the integers whose base-2 expansion is 1, $101^\ell 0$ or $1^\ell 10$ for some $\ell \in \mathbf{Z}_{\geq 0}$. Similarly, all power series in Table 3 are sparse. On the other hand, the description of the support of $\sigma_{K,3}$ in Example 1.3.1 in terms of the base-4 representation with only half the possible digits allowed shows that $\#E(\sigma_{K,3}) \cap \{1, \dots, N\}$ grows as $\sqrt{N}f(N)$ for a function f that is bounded away from both 0 and infinity, and so $\sigma_{K,3}$ is not sparse.

Remark 10.1.6. A sparse automatic series is ‘easy’ in the sense that the full set consisting of the first N terms of the series can be computed in ‘polylogarithmic time’, i.e. polynomial time in $\log(N)$, given the words v_i, w_i as in the definition of a simple sparse set, which allow one to output the nonzero exponents in the series. In contrast to this, computation of the n -th coefficient of a general automatic sequence can be done in time $O(\log(n))$ (by base-2 expansion and running through the automaton), so computing all first N coefficients would require $O(\log(N!)) = O(N \log N)$ time.

Table 8

Automata corresponding to the power series $\sigma_{S,1}$ (left), $\sigma_{S,2}$ (middle) and $\sigma_{S,2^\mu-1}$ ($\mu \geq 3$) (right) in Proposition 10.2.1. The dashed arrow replaces a path consisting of $\mu - 3$ vertices and $\mu - 2$ edges, all with label zero. The remaining missing edges (in the right automaton) all connect to a unique vertex with label 0, which has been omitted in order to simplify the graphical representation of the automaton.



10.2. Conjugating to a sparse representative

One may ask whether every series of finite order in $\mathcal{N}(\mathbf{F}_2)$ can be conjugated to a sparse series. We have no general answer to this question, not even for series of order 2, which form a unique conjugacy class for every value of the break sequence (m) , represented by Klopsch's series $\sigma_{K,m} = t/\sqrt[m]{1+t^m}$. Klopsch's series itself is not sparse, since its m -th power $\sigma_{K,m}^m = t^m/(1+t^m) = \sum_{k \geq 1} t^{km}$ is not. Nevertheless, for special values of the break sequence we can find a sparse representative.

Proposition 10.2.1. *Let m be an integer of the form $m = 2^\mu \pm 1$ for $\mu \geq 1$. Then any power series of order 2 and break sequence (m) is conjugate to a sparse power series. More precisely, we have the following:*

(i) *Any power series of order 2 and break sequence (1) is conjugate to the power series*

$$\sigma_{S,1} = t + \sum_{k \geq 2} \left(t^{2^k-2} + t^{2^k-1} \right), \quad (22)$$

which is sparse of rank 1. The corresponding automaton is displayed in Table 8.

(ii) *If $m = 2^\mu - 1 > 1$, then any power series of order 2 and break sequence (m) is conjugate to the power series*

$$\sigma_{S,m} = t + \sum_{k \geq 1} t^{\frac{m+1}{m-1} \left(m \cdot \left(\frac{m+1}{2} \right)^{k-1} - 1 \right)}, \quad (23)$$

which is sparse of rank 1. The set of exponents occurring in σ consists of the integers whose base-2 representation is either 1 or $10^{\mu-1}(10^{\mu-2})^\ell 0$ for some $\ell \in \mathbf{Z}_{\geq 0}$. The corresponding automata are displayed in Table 8.

(iii) If $m = 2^\mu + 1$, then any power series of order 2 and break sequence (m) is conjugate to the power series

$$\sigma_{S,m} = \sum_{\substack{\emptyset \neq J \subseteq \{0, \dots, \mu-1\} \\ k: J \rightarrow \mathbf{Z}_{\geq 0}}} t^{\left(\sum_{j \in J} 2^j (m-1)^{k(j)}\right)_{m-m+1}}, \quad (24)$$

which is sparse of rank μ : the support of $\sigma_{S,m}$ consists precisely of the integers $m(\ell-1) + 1$ with $\ell \geq 1$ an integer whose base-2 expansion contains at most μ occurrences of the digit 1 and all these occurrences are at distinct positions modulo μ .

The crucial observation used in the proof is stated in the following lemma.

Lemma 10.2.2. *If a polynomial $F(t, X) = 0 \in \mathbf{F}_2[t, X]$ is symmetric in t and X , i.e. $F(t, X) = F(X, t)$, and, when regarded as an algebraic equation in X over $\mathbf{F}_2((t))$, has, for some $m \geq 1$, a unique solution $\sigma \in \mathcal{N}(\mathbf{F}_2)$ of the form $\sigma = t + t^{m+1} + O(t^{m+2})$, then σ is of order 2.*

Proof. Composing the equality $F(t, \sigma) = 0$ on the right with $\sigma^{\circ-1}$ gives $F(\sigma^{\circ-1}, t) = 0$, and hence, by symmetry of F , $F(t, \sigma^{\circ-1}) = 0$. Now note that if $\sigma = t + t^{m+1} + O(t^{m+2})$, then also $\sigma^{\circ-1} = t + t^{m+1} + O(t^{m+2})$. By uniqueness, it follows that $\sigma^{\circ-1} = \sigma$, so σ is of order 2. \square

Proof of Proposition 10.2.1. We know that there is a unique conjugacy class of order-2 power series with a given break sequence (m) , so it suffices to construct such a sparse series. When $m = 2^\mu \pm 1$, we will construct a sparse representative by exhibiting a symmetric algebraic equation $F(t, X) = 0$ over \mathbf{F}_2 as in Lemma 10.2.2. Choose the polynomial as follows:

$$\begin{cases} F(t, X) = (tX)^2 + (tX) + X + t & \text{for } m = 1; \\ F(t, X) = (tX)^{2^{\mu-1}} + X + t & \text{for } m = 2^\mu - 1 > 1; \\ F(t, X) = (tX)^{2^\mu} + X^{2^\mu-1} + t^{2^\mu-1} & \text{for } m = 2^\mu + 1. \end{cases}$$

In all cases, Hensel's Lemma implies the existence and uniqueness of a solution $\sigma = t + t^{m+1} + O(t^{m+2})$, so Lemma 10.2.2 applies. We can find an explicit solution iteratively, as follows.

For $m = 1$ we have

$$\sigma = \frac{t}{t+1} + \frac{t^2}{t+1} \sigma^2 = \frac{t}{t+1} + \frac{t^4}{(t+1)^3} + \frac{t^6}{(t+1)^3} \sigma^4 = \cdots = \frac{t+1}{t^2} \sum_{k \geq 1} \frac{t^{3 \cdot 2^{k-1}}}{(t+1)^{2^k}}.$$

The latter sum is

$$\sum_{k \geq 1} \frac{t^{3 \cdot 2^{k-1}}}{(t+1)^{2^k}} = \sum_{k \geq 1} \sum_{m \geq 1} t^{(2m+1) \cdot 2^{k-1}} = \frac{t}{t+1} + \sum_{k \geq 1} t^{2^{k-1}},$$

leading to the stated formula for $\sigma = \sigma_{S,1}$.

For $m = 2^\mu - 1 > 1$, the same procedure leads to

$$\sigma_{S,m} = t + t^{2^{\mu-1}} \sigma^{2^{\mu-1}} = \dots = t + \sum_{k \geq 0} t^{2^{\mu-1} + 2^{2(\mu-1)} + \dots + 2^{k(\mu-1)} + 2 \cdot 2^{(k+1)(\mu-1)}},$$

which is equivalent to the stated formula.

Finally, for $m = 2^\mu + 1$, we let $\tau = t/\sigma$ and $q = 2^\mu = m - 1$. Then $\tau = 1 + O(t)$ satisfies

$$\tau = t^{q+1} + \tau^q \quad (25)$$

and hence

$$\tau = 1 + \sum_{k \geq 0} t^{q^k(q+1)}.$$

We find

$$t^q \sigma = 1 + \tau^{q-1} = 1 + \tau \cdot \tau^2 \cdot \tau^4 \dots \tau^{2^{\mu-1}} = 1 + \prod_{j=0}^{\mu-1} \left(1 + \sum_{k_j \geq 0} t^{(q+1)2^j q^{k_j}} \right),$$

which is equivalent to the stated formula. \square

Remark 10.2.3. For odd $m \geq 1$ consider the degree-2 extension $\mathbf{F}_2((z))(x)$ of $\mathbf{F}_2((z))$ with $x^2 + x = z^{-m}$. The element $t = xz^{\frac{m+1}{2}}$ is a uniformiser, and the generator σ of the Galois group acts by $\sigma(t) = (x+1)z^{\frac{m+1}{2}}$. We can eliminate the variables x and z by hand, obtaining the equation $(tX)^{\frac{m+1}{2}} + X + t = 0$. This equation always has a unique solution in $\mathcal{N}(\mathbf{F}_2)$, which has depth m , but is not sparse unless $m+1$ is a power of 2 and $m \neq 1$ (this follows from Proposition 11.1.2 below).

Remark 10.2.4. The power series $\sigma_{S,1}$ from Proposition 10.2.1(i) is conjugate to Klopsch's series $\sigma_{K,1} := t/(t+1)$. In this case, the conjugacy can be done using the simple *algebraic* power series $\chi = t/(t^2+1)$. Indeed, with $\psi := \sum_{k \geq 1} t^{2^k-1}$, we have

$$\chi \cdot (\psi \circ \chi) = (t \cdot \psi) \circ \chi = \chi^2 + \chi^4 + \chi^8 + \dots = t^2/(t^2+1),$$

since the support of $t^2/(t^2+1)$ consists of all even integers, and the support of χ^{2^k} consists of the odd multiples of 2^k . Hence $\chi \cdot (\psi \circ \chi) = \chi \cdot t$, so $\chi^{\circ-1} = \psi$. We have $\chi \circ \sigma_{K,1} = t + t^2$, and hence

$$\chi \circ \sigma_{K,1} \circ \chi^{\circ-1} = \chi^{\circ-1} + (\chi^{\circ-1})^2 = \sum_{k \geq 1} t^{2^k-1} + \sum_{k \geq 1} t^{2^{k+1}-2} = \sigma_{S,1}.$$

Remark 10.2.5. In Table 7, we have used that the genus of the smooth projective curve corresponding to $F(t, X) = (tX)^k + X + t$ is $k - 1$. This follows easily by the change of variables $t = y/x^k$, $X = x^{k-1}/y$, leading to the Artin–Schreier equation $y^2 + y = x^{2k-1}$, which has genus $k - 1$ (see e.g. [62, Thm. 6.4.1]). For the case $m = 2^\mu + 1$, we also used that the genus of the Artin–Schreier curve (25) is $2^{\mu-1}(2^\mu - 1)$.

Remark 10.2.6. We did not produce the general form of the automaton for $m = 2^\mu + 1$. Whereas the series for $m = 2^\mu - 1 > 1$ requires $\mu + 3 \approx \log(m)$ states and the rank of sparseness is 1, if $m = 2^\mu + 1$ an educated guess for the number of states of the minimal automaton is $2^\mu + 3^\mu \approx m^{\log(3)/\log(2)}$ and the rank of sparseness is (provably) μ . This looks somewhat similar to what happens with the Klopsch’s series $\sigma_{K,m}$ for such values of m , cf. Remark 9.2.2. In all these families, the number of states appears to be logarithmic or polynomial in the genus, and never exponential, as is theoretically possibly by Bridy’s bound discussed in Section 9.

10.3. Quasi-sparse series

Sparse series form an $\mathbf{F}_2[t]$ -algebra that we will denote by S . Consider the larger $\mathbf{F}_2[t]$ -algebra \widehat{S} consisting of power series in $\mathbf{F}_2[[t]]$ that can be written as products of sparse series and rational functions in $\mathbf{F}_2(t)$. Elements of this algebra can also be regarded as having nice ‘closed formulas’. We have the following characterisation:

Proposition 10.3.1. *Let $\sigma = \sum_{k \geq 0} a_k t^k \in \mathbf{F}_2[[t]]$ be a power series. The following conditions are equivalent:*

- (i) $\sigma \in \widehat{S}$;
- (ii) there exists an integer $m \geq 1$ such that $(t^m + 1)\sigma$ is sparse;
- (iii) there exists an integer $m \geq 1$ such that $\sum_{k \geq 0} (a_k + a_{k+m})t^k$ is sparse;
- (iv) there exists an integer $m \geq 1$ such that $\sum_{k \geq 0} (a_k + a_{k+2^q m})t^k$ is sparse for all integers $q \geq 0$.

Proof. Since sparse power series form a ring and include polynomials, $\sigma \in \widehat{S}$ if and only if there exists a nonzero $p \in \mathbf{F}_2[t]$ such that $p\sigma \in S$. Moreover, we may assume that p is not divisible by t since the class of sparse sequences is closed under shifts. The equivalence of (i) and (ii) then follows from the fact that every $p \in \mathbf{F}_2[t]$ that is not divisible by t divides the polynomial $t^m + 1$ for some $m \geq 1$: take $m = 2^k(2^r - 1)$ with r and k chosen so that the splitting field of p is \mathbf{F}_{2^r} and every root of p has multiplicity $\leq 2^k$. The equivalence of (ii) and (iii), with the same value of m , is easy. Finally, the equivalence of (ii) and (iv) follows from the fact that if $(t^m + 1)\sigma$ is sparse, then so is $(t^m + 1)^{2^q} \sigma = (t^{2^q m} + 1)\sigma$ for all $q \geq 0$. \square

A final operation that we allow without affecting our sense of ‘admitting a closed formula’ is for elements of \widehat{S} to be twisted by an automorphism of $\mathbf{F}_2(t)$, as follows. There is a unique nontrivial field automorphism of $\mathbf{F}_2(t)$ that is also an element of $\mathcal{N}(\mathbf{F}_2)$, given by the map

$$\varphi: t \mapsto t/(t+1).$$

The order of φ is two. It might happen that a power series $\sigma \in \mathcal{N}(\mathbf{F}_2)$ is not in S or \widehat{S} , but that $\sigma \circ \varphi$ is. This is equivalent with σ being in the algebra of sparse series in the variable $t/(t+1)$. Note that while composing with φ preserves the property of being an algebraic power series (if σ is a root of $F(t, X)$, then $\sigma \circ \varphi$ is a root of $F(\varphi(t), X)$), the property of being of finite order need not be preserved.

Definition 10.3.2. A series $\sigma = \sigma(t) \in \mathbf{F}_2[[t]]$ is called *quasi-sparse* if either $\sigma \in \widehat{S}$ or $\sigma \circ \varphi \in \widehat{S}$. We denote the collection of quasi-sparse series by $\widehat{\widehat{S}}$.

This leads to a *hierarchy of complexity* for power series

$$S \subset \widehat{S} \subset \widehat{\widehat{S}} \subset \mathbf{F}_2[[t]],$$

where every inclusion is strict. In the next two sections, we will study whether our series σ of finite order are in S, \widehat{S} or $\widehat{\widehat{S}}$. The next section will employ field-theoretic methods, whereas the following one will be based purely on characterisations in terms of automata. We believe both methods have their merits.

11. Detecting sparseness properties using field theory

11.1. Field-theoretic characterisation of sparseness

Recently, Albayrak and Bell [4, Thm. 1.1(b)] gave an exact field-theoretic characterisation of sparseness for generalized (Hahn) power series in arbitrary positive characteristic. We will use a special case of one direction of their characterisation, of which we include a short, self-contained proof.

The following result will be used without further reference.

Lemma 11.1.1. *For any algebraic power series $\tau \in \overline{\mathbf{F}}_2[[t]]$, the field extension $\mathbf{F}_2(t)(\tau)/\mathbf{F}_2(t)$ is separable.*

Proof. If the extension is not separable, the minimal polynomial $f \in \mathbf{F}_2(t)[X]$ of τ is of the form $f = \sum c_i(t)X^{2^i}$. Since the Cartier operator satisfies $\mathcal{C}_r(\psi\tau^2) = \tau\mathcal{C}_r(\psi)$, applying this to the equation $f(\tau) = 0$, we find that $\sum \mathcal{C}_r(c_i(t))\tau^i = 0$. This gives a polynomial of strictly smaller degree satisfied by τ and nonzero for at least one value of $r \in \{0, 1\}$. This contradiction shows the result. \square

Proposition 11.1.2 (Albayrak–Bell [4], special case). Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ denote a power series that is algebraic over $\mathbf{F}_2(t)$. Consider the field

$$\mathcal{F} = \bigcup_{\substack{\ell \geq 1, \\ \ell \text{ odd}}} \overline{\mathbf{F}}_2(t^{1/\ell}),$$

where $\overline{\mathbf{F}}_2$ is an algebraic closure of \mathbf{F}_2 . If σ is sparse, then the following conditions hold:

- (i) σ is integral over $\overline{\mathbf{F}}_2[t, t^{-1}]$;
- (ii) the extension $\overline{\mathbf{F}}_2(t)(\sigma)/\overline{\mathbf{F}}_2(t)$ is unramified outside of $0, \infty$;
- (iii) the splitting field of σ over \mathcal{F} has degree a power of two.

Proof. The essence of the proof is to show that for sparse power series the combinatorial structure of the support $E(\sigma)$ allows one to construct a tower of Artin–Schreier extensions of \mathcal{F} that contains σ .

By Proposition 10.1.3 a series σ is sparse precisely if $E(\sigma)$ is a finite union of pairwise disjoint simple sparse sets. Properties (i)–(iii) hold for the sum of several power series whenever they hold for the individual summands (for unramifiedness, use [62, Cor. 3.9.3]), and hence it is sufficient to prove that they hold for power series with simple sparse support. This will be done by induction on the rank of sparseness r .

Suppose that the support of σ is a simple sparse set, consisting of integers whose base-2 expansion is of the form $v_r w_r^{\ell_r} \cdots v_1 w_1^{\ell_1} v_0$ with $\ell_i \in \mathbf{Z}_{\geq 0}$ for some fixed binary words v_0, \dots, v_r , and w_1, \dots, w_r . If $r = 0$, then σ is a monomial, and properties (i)–(iii) hold. Suppose that $r \geq 1$ so w_1 is nontrivial. Let $k_0 = |v_0|$ and $k_1 = |w_1|$ be the lengths of the words v_0 and w_1 , and let m_0 and m_1 be the integers whose base-2 expansion is v_0 and w_1 . Let τ be the power series whose support consists of the integers with base-2 expansion of the form $v_r w_r^{\ell_r} \cdots w_2^{\ell_2} v_1 0^{k_0}$ with $\ell_i \in \mathbf{Z}_{\geq 0}$. By induction, we know that properties (i)–(iii) hold for τ . The relation between the supports of σ and τ leads directly to the formula

$$\sigma^{2^{k_1}} - t^{(2^{k_1}-1)m_0-2^{k_0}m_1} \sigma = t^{2^{k_1}m_0-2^{k_0}m_1} \tau. \quad (26)$$

This allows us to deduce the properties (i)–(iii) for σ from the corresponding properties of τ .

First of all, σ is integral over $\overline{\mathbf{F}}_2[t, t^{-1}][\tau]$, and hence also over $\overline{\mathbf{F}}_2[t, t^{-1}]$.

Secondly, the form of Equation (26) makes it very easy to compute the ramification of the extension $\overline{\mathbf{F}}_2(t)(\sigma)/\overline{\mathbf{F}}_2(t)(\tau)$. If f is the minimal polynomial of σ , then [62, Cor. 3.5.11] implies that the extension is unramified at all places P for which f is P -integral and $v_P(f'(\sigma)) = 0$. The same result then holds for any monic (not necessarily minimal) polynomial g satisfied by σ , since it is divisible by f . We apply this with g the polynomial in σ given in (26), and we find that the extension is unramified at all places P of $\overline{\mathbf{F}}_2(t)(\tau)$

with $v_P(t) = 0$ and $v_P(\tau) \geq 0$, and that for all places P' of $\overline{\mathbf{F}}_2(t)(\sigma)$ lying above such P we have $v_{P'}(\sigma) \geq 0$. This implies that $\overline{\mathbf{F}}_2(t)(\sigma)/\overline{\mathbf{F}}_2(t)$ is unramified outside of $0, \infty$.

Finally, multiplying Equation (26) by an appropriate (fractional) power of t leads to an equation of the form $(t^c \sigma)^{2^{k_1}} - (t^c \sigma) = t^d \tau$ for some c and d , which are rational numbers with odd denominators (more precisely, $c = -m_0 + \frac{2^{k_0} m_1}{2^{k_1} - 1}$ and $d = \frac{2^{k_0} m_1}{2^{k_1} - 1}$). This shows that the extension $\mathcal{F}(\sigma)/\mathcal{F}(\tau)$ is contained in a tower of Artin–Schreier extensions, and hence so is $\mathcal{F}(\sigma)/\mathcal{F}$. Thus, its Galois closure is of degree a power of two. \square

11.2. Field-theoretic test for membership in the hierarchy

From Proposition 11.1.2, we can deduce a method for establishing that a series is not in S or \widehat{S} . Since the properties (ii) and (iii) depend only on the field $\overline{\mathbf{F}}_2(t)(\sigma)$, and not on σ itself, any proof that uses them to show that $\sigma \notin S$ will establish the stronger property that $\sigma \notin \widehat{S}$. Actually, the method we will use to show that for a particular σ property (iii) does not hold will even show that $\sigma \notin \widehat{S}$. On the other hand, the integrality property (i) will be used to show that certain series are in \widehat{S} , but not in S .

The basic ingredient is the following field-theoretic result, restricting possible factorisations of polynomials after extension of the base field.

Lemma 11.2.1. *Let L/K be a (possibly infinite) Galois extension with Galois group G , let $f \in K[X]$ be a monic irreducible polynomial, and let $g \in L[X]$ be a monic irreducible factor of f in $L[X]$. Denote by H the stabiliser of g in G . Then*

$$f = \prod_{\phi \in G/H} g^\phi,$$

i.e. f is the product of all (pairwise distinct) Galois conjugates g^ϕ for ϕ running through the coset space G/H .

Proof. Let α denote a root of g in an algebraic closure of L ; then g is the minimal polynomial of α over L . Put $\tilde{f} := \prod g^\phi$, the product being taken over all ϕ running through the coset space G/H . By construction, \tilde{f} lies in $K[X]$ and has α as a root, hence f divides \tilde{f} . Conversely, g divides f in $L[X]$, and hence so does g^ϕ for all $\phi \in G$. Since the elements g^ϕ are irreducible and pairwise distinct for $\phi \in G/H$, the polynomial \tilde{f} divides f . Hence, $f = \tilde{f}$. \square

This implies the following valuation-theoretic result that can be used to check whether a polynomial stays irreducible under base field extension.

Lemma 11.2.2. *Let L/K be a (possibly infinite) Galois extension with Galois group G , and let $v: L \rightarrow \mathbf{R} \cup \{\infty\}$ be an (additive) valuation that is G -invariant, in the sense that $v \circ \phi = v$ for all $\phi \in G$. Let \overline{L} be an algebraic closure of L , and let \tilde{v} be an extension of the*

valuation v to \overline{L} . For a polynomial $f \in K[X]$, denote by $V_v(f)$ the multiset of valuations $\tilde{v}(\alpha)$ of all the roots α of f in \overline{L} . If f is irreducible over K , but becomes reducible over L , then the multiplicities of the elements of $V_v(f)$ have a nontrivial common divisor.

Elements of the set $V_v(f)$ are minus the slopes of the Newton polygon $\text{NP}(f)$ of $f = \sum_{i=0}^n a_i X^i$, where $\text{NP}(f)$ is given as the lower convex hull in \mathbf{R}^2 of the set of points $(i, v(a_i))$ for $0 \leq i \leq n$.

Proof. Since we assume that $v \circ \phi = v$, we have $\text{NP}(g^\phi) = \text{NP}(g)$. Since the multiset $V_v(h)$ of a polynomial h is determined by its Newton polygon (and hence by the valuations of its coefficients), it follows from the decomposition $f = \prod g^\phi$ as in Lemma 11.2.1 that $V_v(f)$ is the union of $[G : H] > 1$ copies of $V_v(g)$ (as multisets). \square

Proposition 11.2.3. *Let $f \in \mathbf{F}_2(t)[X]$ be a separable irreducible polynomial. If the multiplicities of the elements of the multiset $V_t(f)$ for the t -adic valuation have no nontrivial common divisor, then f remains irreducible over \mathcal{F} .*

Proof. The extension $\mathcal{F}/\mathbf{F}_2(t)$ is Galois. The t -adic valuation on $\mathbf{F}_2(t)$ has a unique extension to \mathcal{F} (which coincides on each $\overline{\mathbf{F}}_2(t^{1/j})$ with the $t^{1/j}$ -adic valuation v normalised so that $v(t^{1/j}) = 1/j$). By uniqueness, this extension is Galois invariant. The claim follows from Lemma 11.2.2. \square

Corollary 11.2.4. *Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ denote a power series that is algebraic over $\mathbf{F}_2(t)$ with minimal polynomial $F(t, X)$. Assume that F is of degree not a pure power of two, and that the multiplicities of the elements of the multiset $V_t(\sigma) := V_t(F)$ for the t -adic valuation have no nontrivial common divisor. Then $\sigma \notin \widehat{\widehat{S}}$.*

Proof. We conclude from Proposition 11.2.3 that F is the minimal polynomial of σ over \mathcal{F} , and so $[\mathcal{F}(\sigma) : \mathcal{F}]$ is not a pure power of two, contradicting Proposition 11.1.2(iii). Hence $\sigma \notin S$. Since the field $\mathcal{F}(\sigma)$ does not change after multiplying σ by a rational function, we get that $\sigma \notin \widehat{S}$.

For the final claim, observe that replacing σ by $\sigma \circ \varphi$ changes neither the degree of the minimal polynomial of σ over $\mathbf{F}_2(t) = \mathbf{F}_2(t/(t+1))$ nor the set $V_t(\sigma)$. Hence the same reasoning applied to $\sigma \circ \varphi$ shows that $\sigma \notin \widehat{\widehat{S}}$. \square

Corollary 11.2.5. *Let $\sigma \in \mathcal{N}(\mathbf{F}_2)$ denote a power series that is algebraic over $\mathbf{F}_2(t)$ with minimal polynomial $F(t, X)$ of degree 4*

$$F(t, X) = a_4 X^4 + a_3 X^3 + a_2 X^2 + a_1 X + a_0$$

and with cubic resolvent

$$R_3[F] := a_4^3 X^3 + a_2 a_4^2 X^2 + a_1 a_3 a_4 X + a_0 a_3^2 + a_1^2 a_4.$$

Assume that $R_3[F]$ is irreducible over $\mathbf{F}_2(t)$ and that the multiplicities of the elements of the multisets $V_t(F)$ and $V_t(R_3[F])$ for the t -adic valuation have no nontrivial common divisor. Then $\sigma \notin \widehat{\widehat{S}}$.

Proof. The possible Galois groups of an irreducible separable quartic are S_4 , A_4 , D_4 , $\mathbf{Z}/4\mathbf{Z}$ and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$. Only the last three of these are 2-groups, and those occur precisely when the cubic resolvent is reducible (see [27, Thm. 3.4]).

Since F is separable, so is its cubic resolvent $R_3[F]$. From the hypotheses and Proposition 11.2.3, we conclude that F and $R_3[F]$ are irreducible over \mathcal{F} . Therefore, the Galois group of σ over \mathcal{F} is not a 2-group, and $\sigma \notin S$ by Proposition 11.1.2(iii). Since this argument uses only the information about the field $\mathcal{F}(\sigma)$, we conclude that $\sigma \notin \widehat{\widehat{S}}$.

Finally, since changing σ to $\sigma \circ \varphi$ affects neither the irreducibility of F and $R_3[F]$ nor the sets $V_t(F)$ and $V_t(R_3[F])$, we find similarly that $\sigma \circ \varphi \notin \widehat{\widehat{S}}$, and so $\sigma \notin \widehat{\widehat{S}}$. \square

Theorem 11.2.6. We have the following membership properties (see also Table 9):

- (i) $\sigma_{S,2^\mu \pm 1} (\mu \geq 1), \sigma_{CS}^{\circ 3}, \sigma_{T,1}, \dots, \sigma_{T,4} \in S$;
- (ii) $\sigma_{CS}^{\circ 2}, \sigma_{CS} \in \widehat{S} \setminus S$;
- (iii) $\sigma_J, \sigma_J^{\circ 3} \in \widehat{\widehat{S}} \setminus \widehat{S}$;
- (iv) $\sigma_{K,m} (m \geq 3), \sigma_{V,1}, \sigma_{V,2}, \sigma_{V,3}, \sigma_{\min}, \sigma_{(1,5)}, \sigma_{(1,9)}, \sigma_8 \notin \widehat{\widehat{S}}$.

Proof. The series $\sigma_{S,2^\mu \pm 1}$ are sparse by Proposition 10.2.1. The sparseness of the series $\sigma_{CS}^{\circ 3}, \sigma_{T,1}, \dots, \sigma_{T,4}$ follows by representing $E(\sigma)$ in the same way as was done for $E(\sigma_{CS}^{\circ 3})$ in Example 10.1.5, using the closed formulas for the series in Table 3.

The series $\sigma_{CS}^{\circ 2}$ and σ_{CS} are not sparse by Proposition 11.1.2 since their minimal polynomials are not $\overline{\mathbf{F}}_2[t, t^{-1}]$ -integral. To show the series are in \widehat{S} , we have the following explicit relations, obtained from Remark 5.1.4 and Equation (12), with sparse right hand side:

$$(t+1)^2 \sigma_{CS}^{\circ 2} = t + t^3 + \sum_{k \geq 1} (t^{2 \cdot 2^k} + t^{3 \cdot 2^k}) \quad \text{and} \quad (t+1)^2 \sigma_{CS} = \sum_{k \geq 0} (t^{2^k} + t^{3 \cdot 2^k}).$$

If σ is any of the series σ_J and $\sigma_J^{\circ 3}$, then it is not in $\widehat{\widehat{S}}$. Indeed, from their minimal polynomial we can read out that the extension $\overline{\mathbf{F}}_2(t)(\sigma)/\overline{\mathbf{F}}_2(t)$ is ramified above $t+1$, and the conclusion follows from Proposition 11.1.2(ii). To prove the series are in $\widehat{\widehat{S}}$, we use the following explicit relations with sparse right hand side:

$$(t+1)\sigma_J(\varphi(t)) = (t+1)^2 \sigma_{CS}(t) \quad \text{and} \quad (t^2+t)\sigma_J^{\circ 3}(\varphi(t)) = \sum_{k \geq 0} (t^{3 \cdot 2^k} + t^{2 \cdot 2^k}).$$

Indeed, for the former equation, one verifies that σ_{CS} and $\sigma_J(\varphi(t))/(t+1)$ are equal, since they satisfy the same irreducible algebraic equation (12) having a unique solution

$t+O(t^2)$. For the latter equation, the left hand side is the unique solution to $\tau^2+\tau=t^3+t^2$ of the form $t^2+O(t^3)$. But this solution is clearly equal to the right hand side.

To prove that $\sigma_{K,m} \notin \widehat{S}$ for any odd $m \geq 3$, we use Proposition 11.1.2(iii). To this end, it suffices to check that $(t^m+1)X^m+t^m$ is irreducible over \mathcal{F} , which by [62, Prop. 3.7.3] is equivalent to showing that $t^m/(t^m+1)$ is not a d -th power in \mathcal{F} for any $d > 1$, $d|m$, or, equivalently, that $t^{mj}/(t^{mj}+1)$ is not a d -th power in $\mathbf{F}_2(t)$ for any odd j . This holds since $t^{mj}+1$ has only simple roots in $\overline{\mathbf{F}}_2$. Similarly, $\sigma_{K,m} \circ \varphi$ satisfies $(t^m+(t+1)^m)X^m+t^m=0$, and the polynomial $t^{mj}+(t+1)^{mj}$ has only simple roots in $\overline{\mathbf{F}}_2$ (as can be seen from computing its derivative); hence for the same reason $\sigma_{K,m} \circ \varphi \notin \widehat{S}$. We conclude that $\sigma_{K,m} \notin \widehat{S}$.

The multisets of slopes for the minimal polynomials of $\sigma_{V,1}$, $\sigma_{V,2}$ and $\sigma_{V,3}$ can be found in Table 9. The cubic resolvent for the minimal polynomial of $\sigma_{V,1}$ is $t^{12}X^3+t^8X^2+t^7(t+1)X+t^4(t^4+t^3+t^2+1)$, which is irreducible over $\mathbf{F}_2(t)$ with v_t -slopes $\{-4, (-2)^2\}$. (A convenient way to check irreducibility of the cubic resolvent over $\mathbf{F}_2(t)$ is to consider the v_{t-1} -slopes for the t^{-1} -adic valuation.) Similarly, the minimal polynomial for $\sigma_{V,2}$ has resolvent $(t+1)^{12}X^3+t(t+1)^8X^2+(t+1)^4t^4$, which is irreducible over $\mathbf{F}_2(t)$ and has v_t -slopes $\{1, (3/2)^2\}$, and the minimal polynomial for $\sigma_{V,3}$ has resolvent $t^{12}X^3+t^9(t^2+t+1)X^2+t^4(t+1)^6X+t(t+1)^6(t^3+t^2+1)$, which is irreducible over $\mathbf{F}_2(t)$ and has v_t -slopes $\{(-4)^2, -3\}$. By Corollary 11.2.5 we conclude that $\sigma_{V,1}, \sigma_{V,2}, \sigma_{V,3} \notin \widehat{S}$.

For all further series, $\deg F$ is not a pure power of 2 and $V_t(F)$ has no nontrivial common divisor of multiplicities (listed in Table 9), so we immediately conclude that $\sigma \notin \widehat{S}$ by Corollary 11.2.4. This finishes the proof. \square

12. Sparseness and automaton properties

12.1. Combinatorial characterisation of sparseness

We describe automaton-theoretic methods to verify whether a series σ is in S, \widehat{S} or $\widehat{\widehat{S}}$. In [63], it is shown that sparseness may be checked directly using a corresponding automaton (recall our convention that all states in the automaton are accessible, which is also part of the conditions below).

Definition 12.1.1. Call a vertex v of an automaton *tied* if the following two properties hold:

- (a) there exists a (possibly empty) path from v to a vertex with output 1 [v is co-accessible];
- (b) there exist two different walks of the same length from v to itself.

Proposition 12.1.2 ([63], [18, Prop. 3.4]). *An automatic series σ is not sparse if and only if there exists a tied vertex v in a corresponding automaton.* \square

This criterion can be used immediately to verify that the series $\sigma_{S, 2^\mu - 1} (\mu \geq 1)$, $\sigma_{CS}^{\circ 3}$, $\sigma_{T, 1}, \dots, \sigma_{T, 4}$ are sparse.

Example 12.1.3. The 2-automaton A corresponding to the expansion of the series $(1 + t)^{-1/m}$ can be succinctly described as follows. Let ϖ denote the multiplicative order of 2 modulo m and consider the base-2 expansion $(2^\varpi - 1)/m = \sum_{i=0}^{\varpi-1} x_i 2^i$. The set of vertices of A is $\{v_0, \dots, v_{\varpi-1}, w\}$. All v_j have vertex label 1, w has vertex label 0, and v_0 is the start vertex. For any j , v_j is connected to $v_{j+1 \bmod \varpi}$, always by an edge with label 0, and by an edge with label 1 exactly if $x_j = 1$. If $x_j = 0$, an edge with label 1 connects v_j to w . Finally, w has two self-loops labelled 0 and 1. The automaton A is not sparse since any vertex v_j with $x_j = 1$ is not tied: 0^ϖ and $0^{\varpi-1}1$ are two paths that satisfy condition (b). (This incidentally provides another proof of the non-sparseness of Klopsch's series $\sigma_{K,m}(t) = t/\sqrt[m]{1+t^m}$; however, we do not have a synthetic description for an automaton corresponding to $\sigma_{K,m}$ for general m and, in particular, do not have a formula for the minimal number of states as a function of m , cf. Table 7.)

A similar description of a minimal p -automaton for $(1+at)^{-1/m} \in \mathbf{F}_p[[t]]$ for any prime p , m coprime to p and $a \in \mathbf{F}_p^*$ is given in [64].

12.2. Combinatorial tests for membership in the hierarchy

We have not been able to find a necessary and sufficient condition for a series to be in \widehat{S} in terms of the automaton alone. We will however give a simple necessary criterion, from which one may deduce all statements in Theorem 11.2.6, *except* the facts that $\sigma_{\min} \notin \widehat{S}$ and $\sigma_{(1,9)} \circ \varphi \notin \widehat{S}$.

In applying the criterion, it is necessary to move the ‘start’ label to other vertices. This might produce non-accessible vertices, which should then be removed from the automaton; this does not affect the resulting automatic sequence.

Proposition 12.2.1. *Let $\sigma(t) = \sum_{k \geq 0} a_k t^k \in \mathbf{F}_2[[t]]$ be a power series generated by an automaton A . Then $\sigma(t) \notin \widehat{S}$ if there exists a vertex v in A satisfying the following two properties:*

- (i) *there exist arbitrarily long walks from the start vertex to v ;*
- (ii) *let v_0 and v_1 denote the vertices reached by following the edge starting at v and labelled 0 and 1, respectively, and let A_i be the automaton obtained from A by changing the start vertex to v_i . Then exactly one of the automata A_0 and A_1 is sparse (and the other one is not sparse).*

Remark 12.2.2. Since the automaton is finite, the existence of arbitrarily long walks from the start vertex to v is equivalent to the existence of paths w_0 , w_1 and w_2 such that w_1 is nontrivial and for every integer $\ell \geq 0$ the walk $w_2 w_1^\ell w_0$ goes from the start vertex to v .

Table 9

For each series, in column ‘ $\in S$ ’ the symbol ‘ \times ’ indicates the series is not sparse and ‘ $\checkmark(r)$ ’ indicates the series is r -sparse; the column ‘ $\in \widehat{S}$ ’ describes the property of being sparse up to multiplication with a rational function; the column ‘ $\in \widehat{\widehat{S}}$ ’ indicates whether or not the series itself or its composition with $t \mapsto t/(t+1)$ is in \widehat{S} ; ‘minimal polynomial F ’ is the minimal polynomial of the series over $\mathbf{F}_2(t)$; ‘method’ indicates the method of proof, where $V_t := V_t(F)$ is the multiset of t -adic valuations of the roots of F .

series	$\in S$	$\in \widehat{S}$	$\in \widehat{\widehat{S}}$	minimal polynomial F	method
$\sigma_{S,1}$	$\checkmark(1)$	\checkmark	\checkmark	$t^2X^2 + (t+1)X + t$	direct
$\sigma_{S,m=2^\mu-1>1}$	$\checkmark(1)$	\checkmark	\checkmark	$(tX)^{(m+1)/2} + X + t$	direct
$\sigma_{S,m=2^\mu+1}$	$\checkmark(\mu)$	\checkmark	\checkmark	$(tX)^{m-1} + X^{m-2} + t^{m-2}$	direct
σ_{CS}^3	$\checkmark(1)$	\checkmark	\checkmark	$t^2X^2 + X + t^2 + t$	direct
$\sigma_{T,1}$	$\checkmark(2)$	\checkmark	\checkmark	$t^2X^4 + (t^4 + t^2 + t + 1)X^2 + (t^3 + t^2 + t)X + t^3$	direct
$\sigma_{T,2}$	$\checkmark(3)$	\checkmark	\checkmark	$t^2X^4 + (t+1)X^3 + (t^4 + t^2 + t)X^2 + (t^2 + t)X + t^2$	direct
$\sigma_{T,3}, \sigma_{T,4}$	$\checkmark(3)$	\checkmark	\checkmark	$t^4X^4 + (t^2 + 1)X^3 + (t^3 + t)X^2 + t^2X + t^3$	direct
σ_{CS}^2	\times	\checkmark	\checkmark	$(t+1)^2X^2 + X + t^2 + t$	not integral
σ_{CS}	\times	\checkmark	\checkmark	$(t+1)^2X^2 + X + t$	not integral
σ_J	\times	\times	\checkmark	$(t+1)X^2 + (t^2 + 1)X + t$	not unramified
σ_J^3	\times	\times	\checkmark	$tX^2 + (t^2 + 1)X + t^2 + t$	not unramified
$\sigma_{K,m}$	\times	\times	\times	$(t^m + 1)X^m + t^m$	odd deg & direct
$\sigma_{V,1}$	\times	\times	\times	$t^4X^4 + t^3X^3 + X^2 + (t+1)X + t^2 + t$	R_3 & $V_t = \{(-2)^2, 0, 1\}$
$\sigma_{V,2}$	\times	\times	\times	$(t+1)^4X^4 + tX^2 + t^2X + t^4$	R_3 & $V_t = \{(\frac{1}{2})^2, 1, 2\}$
$\sigma_{V,3}$	\times	\times	\times	$t^4X^4 + (t+1)^3X^3 + t(t^2 + t + 1)X^2 + (t+1)^3X + t(t+1)^2$	R_3 & $V_t = \{-4, 0^2, 1\}$
σ_{\min}	\times	\times	\times	$(t+1)^3X^3 + (t^3 + t)X^2 + (t^3 + t + 1)X + t^3 + t$	odd deg & $V_t = \{0^2, 1\}$
$\sigma_{(1,5)}$	\times	\times	\times	$t^2X^3 + (t+1)^3X + t^3 + t$	odd deg & $V_t = \{(-1)^2, 1\}$
$\sigma_{(1,9)}$	\times	\times	\times	$t^2X^7 + t^3X^6 + (t^5 + t^4 + t^2)X^5 + (t^5 + t^3)X^4 +$ $+(t^7 + t^5 + t^4 + t^3 + t)X^3 + t^5X^2 + (t^3 + t + 1)X + t$	odd deg & $V_t = \{(-\frac{1}{3})^6, 1\}$
σ_8	\times	\times	\times	$t^6X^6 + (t^6 + t^2)X^4 + (t^6 + t^5 + t^4 + t^3 + t^2 + 1)X^2 +$ $+(t+1)^3X + t^6 + t^5 + t^2 + t$	deg not a power of 2 & $V_t = \{(-2)^2, (-1)^2, 0, 1\}$

Proof of Proposition 12.2.1. For the purpose of the proof, we let $(n)_2$ denote the base-2 expansion of an integer $n \geq 0$.

Consider a walk from the start vertex to v , say of length ℓ , and let w be the binary word given by the concatenation of its labels. Let c be the integer such that $(c)_2 = w$. It follows directly from the definition that the automatic sequences produced by A_0 and A_1 are $(a_{2^{\ell+1}n+c})_{n \geq 0}$ and $(a_{2^{\ell+1}n+2^\ell+c})_{n \geq 0}$, respectively. Let $i \in \{0, 1\}$ be such that the automaton A_i is not sparse; the automaton A_{1-i} is then sparse.

Let $m \geq 1$ be a fixed arbitrary odd integer. Consider integers k of the form $k = k(n) = 2^{\ell+1}n + 2^\ell i + c$ (where ℓ and $i = 0, 1$ are fixed while n runs through $\mathbf{Z}_{\geq 0}$). The base-2 expansion of $k + 2^\ell m$ is of the form $(k + 2^\ell m)_2 = u(1-i)w$ for some binary word u , and hence the walk given by it leads from the start vertex to a vertex in A_{1-i} . Since A_{1-i} is sparse, the number of $n \leq N$ such that $a_{k+2^\ell m} = 1$ grows as $O(\log(N)^r)$ for some $r \geq 0$. On the other hand, the base-2 expansion of k is $(k)_2 = (n)_2 iw$, the automaton A_i is not sparse, and hence the number of $n \leq N$ such that $a_k = 1$ grows faster than $\log(N)^r$ for any $r \geq 0$, and so does the number of n such that $a_k + a_{k+2^\ell m} = 1$. It follows that the power series

$$\sum_{n \geq 0} (a_{k(n)} + a_{k(n)+2^\ell m}) t^n$$

is not sparse, and hence neither is the series

$$\sum_{n \geq 0} (a_n + a_{n+2^\ell m}) t^n.$$

Since the integer $m \geq 1$ was arbitrary odd, and since the walk from the start vertex to v can be chosen with ℓ arbitrarily large, we conclude from Proposition 10.3.1(iv) that σ is not in \widehat{S} . \square

A heuristics to apply Proposition 12.2.1 is now as follows. To verify that one of the automata A_0, A_1 is non-sparse, we can use Proposition 12.1.2; for this one can use cycle-finding algorithms. The tricky part is to verify that the other automaton is sparse—to this end, we need to exclude the existence of appropriate walks in the graph. To simplify this problem one may insist that the sparse of the automata A_0, A_1 be very simple; in fact, in all the examples discussed below it is possible to find such an automaton consisting of only one state, with label 0, making the verification obvious. Inverting this logic, we can hope to apply the criterion by first finding a vertex w with label 0 and two self-loops (a so-called ‘absorbing state’, cf. Section 13 below), and then going through all the vertices v admitting an edge from v to w , and checking if any of them satisfies the conditions of Proposition 12.2.1.

Sketch of a second proof of (part of) Theorem 11.2.6. The verification that certain series belong to S, \widehat{S} or $\widehat{\widehat{S}}$ is direct and the same as in the first proof. The verification that

Table 10

‘path’ indicates a path from the start vertex to a tied vertex; ‘path to vertex with output 1’ indicates a path from the tied vertex to a vertex with output 1; p_1 and p_2 are walks of the same length that connect the tied vertex to itself, indicating that the series is non-sparse; ϵ indicates the empty path.

series	path	path to vertex with output 1	(p_1, p_2)
σ_{CS}	0	1	(101, 100)
$\sigma_{CS}^{\circ 2}$	0	10	(1101, 1110)
σ_{\min}	1	ϵ	(011, 100)

certain series do not belong to S , \widehat{S} or $\widehat{\widehat{S}}$ can be done by studying the corresponding automata and using Propositions 12.1.2 and 12.2.1. We have summarised some of the combinatorial data for this in Tables 10 & 11. For small automata, these data can be easily found just by looking at the graphical representation. This is the case for all the series in Tables 10 & 11 except for $\sigma_{(1,9)}$. To illustrate how one can use a computer algebra system to find these data for larger automata, we have written a MATHEMATICA notebook doing this for the series $\sigma_{(1,9)}$, generated by an automaton with 110 states, see [17].

To verify that a series is not in S , one indicates a path from the start vertex to a tied vertex v and two different walks of the same length from v to itself. One also checks that v is co-accessible by indicating a path from v to a vertex with output 1. These data are gathered in Table 10.

To verify that a series is not in \widehat{S} one indicates paths w_0, w_1, w_2 such that every walk $w_2 w_1^\ell w_0$ leads from the start vertex to the same vertex v ; a digit $i \in \{0, 1\}$ such that the automaton A_i (resp. A_{1-i}) obtained by moving the start vertex to the endpoint v_i of the edge starting at v and labelled i is non-sparse (resp. sparse); a path from v_i to a tied vertex; a path from that tied vertex to a vertex with output 1; and different walks of the same length from the tied vertex to itself, verifying that the automaton A_i is non-sparse. In all the cases listed in Table 11 the vertex v_{1-i} has label zero and two self-loops, implying that the automaton A_{1-i} is sparse, and providing the final step of the verification that the considered series is not in \widehat{S} .

We have listed the combinatorial data only for some of the series, but a similar procedure can be performed for all the series considered in Table 9 except σ_{\min} and $\sigma_{(1,9)} \circ \varphi$, which are not in \widehat{S} , but for which the criterion from Proposition 12.2.1 is not satisfied. \square

Remark 12.2.3. Since the class \widehat{S} contains all power series whose coefficients are ultimately periodic, an automaton-theoretic criterion for membership in \widehat{S} gives a necessary criterion for ultimate periodicity. It is known how to test for ultimate periodicity algorithmically, e.g. by work of Honkala [40] (this reference is phrased in a different, but equivalent language, where, for series over a binary alphabet, ‘ p -automatic’ is ‘ p -recognisable’ and ‘ultimately periodic’ is ‘recognisable’, or p -recognisable for all p , by

Table 11

The words w_i are the words needed to apply Remark 12.2.2; ‘edge to non-sparse’ has value $i \in \{0, 1\}$ if the automaton A_i considered in Proposition 12.2.1 is non-sparse; ‘path’ indicates a path from the vertex v_i from Proposition 12.2.1 to a tied vertex; ‘path to vertex with output 1’ indicates a path from the tied vertex to a vertex with output 1; p_1 and p_2 are walks of the same length that connect the tied vertex to itself, indicating that the series is not in \hat{S} ; ϵ indicates the empty path.

series	(w_2, w_1, w_0)	edge to non-sparse	path	path to vertex with output 1	(p_1, p_2)
σ_J	$(1, 0, 00)$	1	ϵ	ϵ	$(0, 1)$
σ_J^{o3}	$(1, 0, 001)$	1	ϵ	ϵ	$(0, 1)$
$\sigma_{V,1}$	$(1, 0, 000)$	0	ϵ	1	$(1001, 0100)$
$\sigma_{V,2}$	$(1, 0, 1)$	0	ϵ	1	$(1001, 0100)$
$\sigma_{K,3}$	$(\epsilon, 00, 0)$	0	01	ϵ	$(00, 11)$
$\sigma_{(1,5)}$	$(1, 0, 001)$	0	1	ϵ	$(11001, 01011)$
$\sigma_{(1,9)}$	$(0^21010, 1, 1^30^3)$	1	001	1	$(0^21^20^51^20^210^2101^20^2,$ $0^31^20^210^2101^20^41^20^2)$

Cobham’s theorem [25].) The algorithm involves constructing another non-deterministic automaton and determining it. It might be that one may find a similar algorithm for membership in \hat{S} . Nevertheless, this seems to indicate that ‘seeing’ membership in \hat{S} directly from the automaton might be hard.

13. ‘Non-randomness’ of the series and synchronisability of the automata

13.1. Synchronising automata

Recall that an automaton is called *synchronising* if there is an input string (a ‘synchronising word’ p_{sync}), which, when followed from an arbitrary vertex, always leads to the same end vertex; this means that the word resets the automaton—if the base-2 expansion of n contains the word p_{sync} , the corresponding coefficient a_n depends only on the part of the expansion that is to the left of the occurrence of p_{sync} .

Example 13.1.1. The word 1011 is synchronising for $\sigma_{K,3}$. Following this word (right to left) starting at any state of the automaton leads to the state in the middle of the bottom row of Fig. 1.

Synchronisation is particularly easy to check when there is an *absorbing state* v , meaning that both outgoing edges from v are loops.

Lemma 13.1.2. *If an automaton A has an absorbing state v , then A is synchronising if and only if for any vertex w in A there is a path from w to v (in particular, A is not synchronising if there is more than one absorbing state).*

Proof. Since v is mapped to itself by any word, the synchronising word should map any vertex to v . In particular, for A to be synchronising, any vertex needs to be connected by a path to v . If this holds, choose an input string p for which the number of end vertices

of all paths with label p and arbitrary beginning vertex is minimal. If the only such end vertex is the absorbing state v , p is a synchronising word. If not, let v_1 denote another such end vertex and choose a path p_1 from v_1 to v (which exists by assumption). Now, the number of end vertices of paths with label p_1p is strictly smaller than for p (since both v_1 and v are connected to v by a path with label p_1), contradicting the minimality. \square

As the number N tends to infinity, the fraction of synchronising automata with N states tends to 1 [7], but the fraction of automata with N states having an absorbing state tends to 0. The next lemma shows something very different happens for the class of minimal *sparse* automata.

Lemma 13.1.3. *If an automaton A is minimal and sparse, then A has a unique absorbing state v , and for any vertex w in A there is a path from w to v .*

Proof. Call any maximal subgraph of A that is connected as a directed graph a *strongly connected component*. For example, any absorbing state is a strongly connected component.

Let U denote the union of all strongly connected components. For any vertex v of A let $n(v)$ be the number of vertices that can be reached from v by following some directed path. It is easy to see that if for some vertex w there is a path from v to w , then $n(w) \leq n(v)$, and that equality holds for all such w exactly if v lies in U . Choosing w to be a vertex admitting a path from v to w for which the value of $n(w)$ is minimal, we see that for any vertex there is a path from that vertex to a vertex in U . An argument analogous as in the proof of Lemma 13.1.2 (but with U playing the role of the vertex v in that proof) shows that there is an input string p such that for every path with label p originating from any vertex, the end vertex lies in some strongly connected component.

We now assume that A is *sparse*, and we claim that then all vertices in U have vertex label 0. Indeed, by the combinatorial criterion in Proposition 12.1.2, A has no tied vertices, but any vertex v with label 1 lying in some strongly connected component is tied: by strong connectedness, two directed edges starting at v with different labels can each be continued to paths p and q leading back to v , and then pq and qp are two different paths of the same length connecting v to itself.

Thus, the automaton A' obtained by replacing every vertex in U with a single absorbing state with vertex label 0 produces the same output as A . We conclude that if A is *sparse and minimal*, it has only one strongly connected component, and this component is an absorbing state with label 0. \square

13.2. ‘Non-randomness’

A power series corresponding to a synchronising automaton with an absorbing state is not ‘random’ at all: if the binary expansion of n contains a synchronising word p_{sync}

leading to an absorbing state, the corresponding coefficient a_n will always be the same, namely the output value of the absorbing state. Since most integers have binary expansions containing p_{sync} , it follows that a_n is constant for ‘almost all’ n , i.e. there is some $c > 0$ such that a_n takes the same value for all except $O(N^{1-c})$ values $n < N$.

So far, we used the convention that our automata were leading-zero invariant, which we now drop. In order to produce automatic sequences from automata, we used the backwards-reading convention (starting from the least significant digit), and sequences obtained in this manner from synchronising automata may be more properly called *backwards synchronising* to distinguish them from the forwards-reading convention (starting from the most significant digit), which leads to the notion of a *forwards synchronising* automatic sequence. For a given sequence, these two notions are not equivalent (the sequence $(n \bmod 2)$ is forwards synchronising, but not backwards synchronising, and we will see below that the sequence of coefficients of the series σ_{\min} is backwards synchronising, but not forwards synchronising). With both of these notions at hand, we may now refer to the following precise result about structured versus random sequences. In [19, Thm. C] it was shown that any \mathbf{C} -valued automatic sequence (such as our sequences with the output alphabet \mathbf{F}_2 lifted to $\{0, 1\} \subset \mathbf{C}$) can be decomposed as a sum of a ‘structured sequence’, in which the n -th coefficient is a function of the n -th coefficients of a periodic sequence and forwards and backwards synchronising sequences, and a ‘random sequence’, meaning a highly Gowers uniform sequence. (Since in this sense sequences that are 0 almost everywhere are ‘random’, the terminology is somewhat loose.) The classical Thue–Morse sequence is an example of a highly Gowers uniform sequence [49]. By contrast, it turns out that our sequences are very structured and non-random in the sense of this decomposition. As an example, consider the series σ_{CS} : it follows from Equation (12) that the value of its n -th coefficient for $n \geq 3$ depends only on the two leading digits and the final digit of the base-2 expansion of n .

Proposition 13.2.1. *For all series $\sigma = \sum a_n t^n$ in Table 9 the sequence (a_n) is structured: there exists a backwards synchronising sequence (b_n) , a forwards synchronising sequence (f_n) and a function $F: \mathbf{F}_2^2 \rightarrow \mathbf{F}_2$ such that $a_n = F(b_n, f_n)$ for all n .*

Proof. All series in Table 9 except σ_{\min} , σ_{CS}^2 , σ_{CS} , σ_{J} and σ_{J}^3 are produced by automata that admit an absorbing state that is accessible from any other state of the automaton, and hence by Lemma 13.1.2, they are (backwards and forwards) synchronising. Indeed, for small automata, one may inspect the pictures; for the larger automata, the verification can be found in [17]; for the series $\sigma_{\text{S}, 2^\mu + 1}$, for which we have not given a representation of the corresponding automata, one may rely on their sparseness and invoke Lemma 13.1.3.

To treat the remaining cases, we observe the following. The minimal automaton corresponding (in backwards-reading convention) to σ_{\min} is synchronising with synchronising word 1^3 , and so the corresponding sequence is backwards synchronising (using [19, Lemma 3.2] it can be proven that it is not forwards synchronising). The automata

corresponding to σ_J and $\sigma_J^{\circ 3}$ have two absorbing states, and every state has a path to one of these two states; this is enough to conclude that these sequences are forwards synchronising (cf. [19, Lemma 3.2]). Finally, the automata for $\sigma_{CS}^{\circ 2}$ and σ_{CS} have two subgraphs that are synchronising and the start vertex is connected by an outgoing edge to these two subgraphs; it follows that the value a_n of the corresponding sequence depends on the value of a backwards synchronising sequence (the sequence produced by the product automaton for the subgraphs) and on the value of the sequence $(n \bmod 2)$, which is forwards synchronising. \square

Synchronisability is not invariant under conjugation of the corresponding power series, so one may wonder whether every conjugacy class of elements of finite order in $\mathcal{N}(\mathbf{F}_2)$ has a synchronising representative.

How computations and visualisations were done

- Equations and uniformisers were computed by hand. SINGULAR or MATHEMATICA were used for elimination of variables and checking irreducibility of equations.
- Automata were generated in MATHEMATICA by Rowland’s package [58]. Shapes of automata were verified using the MAGMA code in [14]. This code was also used to compute the number of states of certain automata that were not computed in further detail.
- Automata were redrawn using tikz and Evan Wallace’s Finite State Machine Design app (github.com/evanw/fsm), with the exception of the visualisation of the automaton for $\sigma_{(1,9)}$, which was drawn in MATHEMATICA, exported as eps and the ‘Start’-label was added in INKSCAPE.
- The genus of the curves in Table 7 were computed using SINGULAR, with the exception of $\sigma_{(1,9)}$, which was computed in MAGMA.
- All claimed automata and explicit series representations were verified in MATHEMATICA to $O(t^{200})$ at least.
- The file `LabelledDirectedGraph.txt` in [14] contains the MAGMA-routine to compute the labelled directed graph structure (without vertex output labels) from Algorithm 3.1.2 using the method of differential forms, in a form that can be parsed by Rowland’s MATHEMATICA package [58]. We give two examples of the running time using the online calculator for MAGMA V2.25-5: for σ_{\min} the labelled directed graph is computed in 0.090 seconds, and the computation of the number of states in Remark 7.3 being 668 required 2.74 seconds.

Description of supplementary material

- The file `automata-of-finite-order` contains, for each of the series occurring in this paper, an irreducible algebraic equation that it satisfies, initial coefficients that uniquely determine it as a solution to that algebraic equation, and the corresponding

automaton, stored in the format of [58] and visualised as a graph. The series occur by the name used in the current paper, and are ordered by compositional order, then by lexicographical order of the lower break sequence.

- The file **verification-of-non-sparseness** contains the material needed to verify combinatorially that $\sigma_{(1,9)} \notin \widehat{S}$.
- The file **verification-of-synchronisation** contains the material needed to verify that $\sigma_{V,3}$, $\sigma_{(1,9)}$ and σ_8 are synchronising.

Acknowledgments

JB was supported by National Science Center, Poland under grant no. 2016/23/-D/ST1/01124. DT was supported in part by the research training group *GRK 2240: Algebraic-geometric Methods in Algebra, Arithmetic and Topology*, funded by the DFG.

We thank Jeroen Sijsling for advice on various computations, Jonathan Lubin for sharing his unpublished work on conjugacy classes in the Nottingham group, Andrew Bridy and Eric Rowland for many interesting discussions about implementations, and Jason Bell for some insightful discussions. We also thank Ragnar Groot Koerkamp for setting up a computer search for small automata. We also thank the reviewer for some pertinent suggestions that are reflected in Subsection 8.3 and Remark 12.2.3.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jalgebra.2022.03.019>.

References

- [1] Boris Adamczewski, Jason P. Bell, On vanishing coefficients of algebraic power series over fields of positive characteristic, *Invent. Math.* 187 (2) (2012) 343–393.
- [2] Boris Adamczewski, Jason P. Bell, Diagonalization and rationalization of algebraic Laurent series, *Ann. Sci. Éc. Norm. Supér.* (4) 46 (6) (2013) 963–1004.
- [3] Boris Adamczewski, Reem Yassawi, A note on Christol’s theorem, arxiv:1906.08703, 2019 (14 pp.), not for publication.
- [4] Seda Albayrak, Jason P. Bell, A refinement of Christol’s theorem for algebraic power series, *Math. Z.* 300 (3) (2022) 2265–2288.
- [5] Jean-Paul Allouche, Jeffrey Shallit, *Automatic Sequences*, Cambridge University Press, Cambridge, 2003.
- [6] Laurent Bartholdi, Endomorphic presentations of branch groups, *J. Algebra* 268 (2) (2003) 419–443.
- [7] Mikhail V. Berlinkov, On the probability of being synchronizable, in: *Algorithms and Discrete Applied Mathematics*, in: *Lecture Notes in Comput. Sci.*, vol. 9602, Springer, Cham, 2016, pp. 73–84.
- [8] José Bertin, Ariane Mézard, Déformations formelles des revêtements sauvagement ramifiés de courbes algébriques, *Invent. Math.* 141 (1) (2000) 195–238.
- [9] Frauke M. Bleher, Ted Chinburg, Bjorn Poonen, Peter Symonds, Automorphisms of Harbater-Katz-Gabber curves, *Math. Ann.* 368 (1–2) (2017) 811–836.
- [10] Svetlana I. Bogataya, Semen A. Bogaty, Denis D. Kiselev, Powers of elements of the series substitution group $\mathcal{J}(\mathbb{Z}_2)$, *Topol. Appl.* 201 (2016) 29–56.
- [11] Wieb Bosma, John Cannon, Catherine Playoust, The Magma algebra system. I. The user language, in: *Computational Algebra and Number Theory* (London, 1993), *J. Symb. Comput.* 24 (3–4) (1997) 235–265, <http://magma.maths.usyd.edu.au/magma/>.

- [12] Alin Bostan, Xavier Caruso, Gilles Christol, Philippe Dumas, Fast coefficient computation for algebraic power series in positive characteristic, in: R. Scheidler, J. Sorenson (Eds.), ANTS XIII – Proceedings of the Thirteenth Algorithmic Number Theory Symposium, U Wisconsin, Madison, in: The Open Book Series, vol. 2, Mathematical Sciences Publishers, Berkeley, 2019, pp. 119–135, <http://msp.org/obs/2>.
- [13] Andrew Bridy, Automatic sequences and curves over finite fields, *Algebra Number Theory* 11 (3) (2017) 685–712.
- [14] Andrew Bridy, Gunther Cornelissen, `LabelledDirectedGraph.txt`, MAGMA routine, supplementary material, <https://arxiv.org/abs/2008.04971>, 2020.
- [15] Jakub Byszewski, Gunther Cornelissen, Which weakly ramified group actions admit a universal formal deformation?, *Ann. Inst. Fourier (Grenoble)* 59 (3) (2009) 877–902.
- [16] Jakub Byszewski, Gunther Cornelissen, Fumiharu Kato, Un anneau de déformation universel en conducteur supérieur, *Proc. Jpn. Acad., Ser. A, Math. Sci.* 88 (2) (2012) 25–27.
- [17] Jakub Byszewski, Gunther Cornelissen, Djurre Tijsma, `automata-of-finite-order; verification-of-non-sparseness; verification-of-synchronisation`, MATHEMATICA® (v12) notebooks (nb) and pdf printout (pdf), supplementary material, <https://arxiv.org/abs/2008.04971>, 2020.
- [18] Jakub Byszewski, Jakub Konieczny, Automatic sequences and generalised polynomials, *Can. J. Math.* 72 (2) (2020) 392–426.
- [19] Jakub Byszewski, Jakub Konieczny, Clemens Müllner, Gowers norms for automatic sequences, preprint, arXiv:2002.09509, 2020.
- [20] Rachel Camina, Subgroups of the Nottingham group, *J. Algebra* 196 (1) (1997) 101–113.
- [21] Rachel Camina, The Nottingham group, in: Marcus du Sautoy, Dan Segal, Aner Shalev (Eds.), *New Horizons in Pro- p Groups*, in: *Progr. Math.*, vol. 184, Birkhäuser Boston, Boston, MA, 2000, pp. 205–221.
- [22] Ted Chinburg, Peter Symonds, An element of order 4 in the Nottingham group at the prime 2, preprint, arXiv:1009.5135, 2010, 3 pp.
- [23] Gilles Christol, Ensembles presque periodiques k -reconnaissables, *Theor. Comput. Sci.* 9 (1) (1979) 141–145.
- [24] Gilles Christol, Teturo Kamae, Michel Mendès France, Gérard Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. Math. Fr.* 108 (4) (1980) 401–419.
- [25] Alan Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Syst. Theory* 3 (1969) 186–192.
- [26] Alan Cobham, Uniform tag sequences, *Math. Syst. Theory* 6 (1972) 164–192.
- [27] Keith Conrad, Galois groups of cubics and quartics in all characteristics, <https://kconrad.math.uconn.edu/blurbs/galoistheory/cubicquarticallchar.pdf>. (Accessed 4 November 2019), undated expository note (21 pp.).
- [28] Gunther Cornelissen, Fumiharu Kato, Equivariant deformation of Mumford curves and of ordinary curves in positive characteristic, *Duke Math. J.* 116 (3) (2003) 431–470.
- [29] Gunther Cornelissen, Fumiharu Kato, Zur Entartung schwach verzweigter Gruppenoperationen auf Kurven, *J. Reine Angew. Math.* 589 (2005) 201–236.
- [30] Gunther Cornelissen, Ariane Mézard, Relèvements des revêtements de courbes faiblement ramifiés, *Math. Z.* 254 (2) (2006) 239–255.
- [31] Wolfram Decker, Gert-Martin Greuel, Gerhard Pfister, Hans Schönemann, SINGULAR 4-1-2 — A computer algebra system for polynomial computations, <http://www.singular.uni-kl.de>, 2019 (used through Singular Online, 11 July 2019).
- [32] Mikhail Ershov, On the commensurator of the Nottingham group, *Trans. Am. Math. Soc.* 362 (12) (2010) 6663–6678.
- [33] Ivan B. Fesenko, On just infinite pro- p -groups and arithmetically profinite extensions of local fields, *J. Reine Angew. Math.* 517 (1999) 61–80.
- [34] Ivan B. Fesenko, Sergei V. Vostokov, *Local Fields and Their Extensions*, second ed., *Translations of Mathematical Monographs*, vol. 121, American Mathematical Society, Providence, RI, 2002.
- [35] Harry (Hillel) Furstenberg, Algebraic functions over finite fields, *J. Algebra* 7 (1967) 271–277.
- [36] Barry Green, Realizing deformations of curves using Lubin-Tate formal groups, *Isr. J. Math.* 139 (2004) 139–148.
- [37] Ragnar Groot Koerkamp, C++ program for searching small automata for algebraic power series over \mathbf{F}_2 , <https://github.com/RagnarGrootKoerkamp/automata/releases/tag/v1.0> (version released 5 Aug 2020).
- [38] David Harbater, Moduli of p -covers of curves, *Commun. Algebra* 8 (12) (1980) 1095–1122.

- [39] David R. Hayes, Explicit class field theory for rational function fields, *Trans. Am. Math. Soc.* 189 (1974) 77–91.
- [40] Juka Honkala, A decision method for the recognizability of sets defined by number systems, *RAIRO Inform. Théor. Appl.* 20 (4) (1986) 395–403.
- [41] Chun Yin Hui, Krishna Kishore, Torsion elements of order p^2 in the Nottingham group, *J. Group Theory* 23 (3) (2020) 489–502.
- [42] Sandrine Jean, Classification à conjugaison près des séries de p -torsion, Doctoral Thesis (108 pp.), Université de Limoges, <https://www.theses.fr/2008LIMO4011>, 2008. (Accessed 6 April 2018).
- [43] Sandrine Jean, Conjugacy classes of series in positive characteristic and Witt vectors, *J. Théor. Nr. Bordx.* 21 (2) (2009) 263–284.
- [44] Stephen A. Jennings, Substitution groups of formal power series, *Can. J. Math.* 6 (1954) 325–340.
- [45] Kiyomi Kanesaka, Koji Sekiguchi, Representation of Witt vectors by formal power series and its applications, *Tokyo J. Math.* 2 (2) (1979) 349–370.
- [46] Nicholas M. Katz, Local-to-global extensions of representations of fundamental groups, *Ann. Inst. Fourier* 36 (4) (1986) 69–106.
- [47] Denis D. Kiselev, Explicit embeddings of finite abelian p -groups in the group $\mathcal{J}(\mathbb{F}_p)$, *Mat. Zametki* 97 (1) (2015) 74–79.
- [48] Benjamin Klopsch, Automorphisms of the Nottingham group, *J. Algebra* 223 (1) (2000) 37–56.
- [49] Jakub Konieczny, Gowers norms for the Thue-Morse and Rudin-Shapiro sequences, *Ann. Inst. Fourier (Grenoble)* 69 (4) (2019) 1897–1913.
- [50] Aristides Kontogeorgis, Ioannis Tsouknidas, A cohomological treatise of HKG-covers with applications to the Nottingham group, *J. Algebra* 555 (2020) 325–345.
- [51] Peter Linz, *An Introduction to Formal Languages and Automata*, 6th edition, Jones and Bartlett Publishers, 2016.
- [52] Falko Lorenz, *Algebra. Vol. II (Fields with Structure, Algebras and Advanced Topics)*, Universitext, Springer, New York, 2008.
- [53] Jonathan Lubin, Classifying torsion elements of the Nottingham group of period p^2 and type $\langle 1, m \rangle$, unpublished manuscript (9 pp.), 23 Jan 2016.
- [54] Jonathan Lubin, Torsion in the Nottingham group, *Bull. Lond. Math. Soc.* 43 (3) (2011) 547–560.
- [55] Jonathan Lubin, John Tate, Formal complex multiplication in local fields, *Ann. Math.* (2) 81 (1965) 380–387.
- [56] Barry Mazur, An introduction to the deformation theory of Galois representations, in: *Modular Forms and Fermat’s Last Theorem*, Boston, MA, 1995, Springer, New York, 1997, pp. 243–311.
- [57] R. Hjalmar Mellin, Résolution de l’équation algébrique générale à l’aide de la fonction gamma, *C. R. Acad. Sci., Paris* 172 (1921) 658–661.
- [58] Eric Rowland, INTEGERSEQUENCES, a MATHEMATICA® package for identifying and analyzing a variety of classes of integer sequences, <https://github.com/ericrowland/IntegerSequences> (version 1.53 dated 30 May 2020).
- [59] Eric Rowland, IntegerSequences: a package for computing with k -regular sequences, in: J.H. Davenport, M. Kauers, G. Labahn, J. Urban (Eds.), *6th International Congress on Mathematical Software – ICMS 2018*, South Bend, in: *Lecture Notes in Computer Science*, vol. 10931, Springer, Cham, 2018, pp. 414–421.
- [60] Eric Rowland, Reem Yassawi, Automatic congruences for diagonals of rational functions, *J. Théor. Nr. Bordx.* 27 (1) (2015) 245–288.
- [61] Imke Rust, Ortwin Scheja, A guide to explicit class field theory in global function fields, in: *Drinfeld Modules, Modular Schemes and Applications*, Alden-Biesen, 1996, World Sci. Publ., River Edge, NJ, 1997, pp. 44–65.
- [62] Henning Stichtenoth, *Algebraic Function Fields and Codes*, second ed., Graduate Texts in Mathematics, vol. 254, Springer-Verlag, Berlin, 2009.
- [63] Andrew Szilard, Sheng Yu, Kaizhong Zhang, Jeffrey Shallit, Characterizing regular languages with polynomial densities, in: *Mathematical Foundations of Computer Science 1992*, Prague, 1992, in: *Lecture Notes in Comput. Sci.*, vol. 629, Springer, Berlin, 1992, pp. 494–503.
- [64] Djurre Tijsma, Automata and finite order elements in the Nottingham group, Master thesis, Utrecht University, 2018.
- [65] Bradley Weaver, The local lifting problem for D_4 , *Isr. J. Math.* 228 (2) (2018) 587–626.