AN ELEMENTARY PROOF OF BRIDY'S THEOREM

ERIC ROWLAND, MANON STIPULANTI, AND REEM YASSAWI

Abstract. Christol's theorem states that a power series with coefficients in a finite field is algebraic if and only if its coefficient sequence is automatic. A natural question is how the size of a polynomial describing such a sequence relates to the size of an automaton describing the same sequence. Bridy used tools from algebraic geometry to bound the size of the minimal automaton for a sequence, given its minimal polynomial. We produce a new proof of Bridy's bound by embedding algebraic sequences as diagonals of rational functions.

1. Introduction

A well-known result of Christol [9, 11] states that a sequence $a(n)_{n\geq 0}$ of elements in the finite field $\mathbb{F}_q$ is algebraic if and only if it is q-automatic. That is, its generating series $F = \sum_{n\geq 0} a(n)x^n$ satisfies $P(x, F) = 0$ for some nonzero polynomial $P \in \mathbb{F}_q[x, y]$ precisely when there exists a finite automaton that outputs $a(n)$ when fed the standard base-q representation of n. Such sequences can therefore be represented both by polynomials and by automata. A natural question is how the size of the automaton, measured by the number of states, depends on the size of the polynomial, measured by its *height* $h := \deg_x P$ and *degree* $d := \deg_y P$. Using tools from algebraic geometry, Bridy [7] showed that the number of states is in $(1 + o(1))q^{hd}$ as q, h, or d tends to infinity.

In this paper, we give a new proof of Bridy's theorem using tools from linear algebra and results about constant-recursive sequences. Quite apart from the interest of providing an elementary proof of Bridy's result, our approach generalizes to settings that are not accessible to algebraic geometry. An analogue of Christol's theorem has been established for sequences of p-adic integers [10, 13]. In a subsequent paper [22], we use our approach to bound the number of states in the minimal automaton for an algebraic sequence of p-adic integers reduced modulo p^α.

All automata in this article read representations of integers starting with the least significant digit; see Section 3. We will be interested in sequences with polynomial representations as follows.

Definition. Let $P \in \mathbb{F}_q[x, y]$ such that $P(0, 0) = 0$ and $\frac{\partial P}{\partial y}(0, 0) \neq 0$. The *Furstenberg series* associated with P is the unique power series $F \in \mathbb{F}_q[\![x]\!]$ satisfying $F(0) = 0$ and $P(x, F) = 0$.

The condition $\frac{\partial P}{\partial y}(0, 0) \neq 0$ is a statement about the coefficient of x^0y^1. It guarantees that $d \geq 1$. If $h = 0$, then F is the trivial 0 series, so we may assume $h \geq 1$. Along with the condition $P(0, 0) = 0$, a version of the implicit function

Date: November 18, 2024.

1991 *Mathematics Subject Classification.* 11B85, 13F25.

The second author is an FNRS Research Associate supported by the Research grant 1.C.104.24F. The third author was supported by the EPSRC, grant number EP/V007459/2.

theorem guarantees the uniqueness of F [16, Theorem 2.9]. Given a polynomial P which does not satisfy the conditions $P(0,0) = 0$ and $\frac{\partial P}{\partial y}(0,0) \neq 0$, and a power series F satisfying $P(x,F) = 0$, there is a technique to obtain a polynomial $\bar{P}$ and a "shift" $\bar{F}$ of F such that $\bar{F}$ is the Furstenberg series associated with $\bar{P}$. For example, see [1, Lemma 6.2] for details. The results in this article are stated for Furstenberg series, but this technique can be used to extend them to general algebraic series.

Our main result is Theorem 1, whose statement needs a few definitions. Define $\mathrm{parts}(n)$ to be the set of all integer partitions of n. We are interested in the lcm of an integer partition, since it will arise as $\mathrm{lcm}(\deg R_1, \ldots, \deg R_k)$ where $R_1, \ldots, R_k$ are the irreducible factors of a polynomial of fixed degree n. The *Landau function* $g(n)$ outputs the maximum value of $\mathrm{lcm}(\sigma)$ over all integer partitions $\sigma \in \mathrm{parts}(n)$ [23, A000793]. For example, $g(5)$ is the maximum value among $\mathrm{lcm}(5)$, $\mathrm{lcm}(4,1)$, $\mathrm{lcm}(3,2)$, $\mathrm{lcm}(3,1,1)$, $\mathrm{lcm}(2,2,1)$, $\mathrm{lcm}(2,1,1,1)$, and $\mathrm{lcm}(1,1,1,1,1)$, so we have $g(5) = 6$. The Landau function also appeared in Bridy's analysis. We will use a variant of the Landau function that gives a better bound. Define

$$\mathcal{L}(l,m,n) := \max_{\substack{1\le i\le l \\ 1\le j\le m \\ 1\le k\le n}} \max_{\substack{\sigma_1\in\mathrm{parts}(i) \\ \sigma_2\in\mathrm{parts}(j) \\ \sigma_3\in\mathrm{parts}(k)}} \mathrm{lcm}(\mathrm{lcm}(\sigma_1), \mathrm{lcm}(\sigma_2), \mathrm{lcm}(\sigma_3)).$$

Theorem 1. *Let $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!] \setminus \{0\}$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$ of height h and degree d. Then the minimal q-automaton that generates $a(n)_{n\ge 0}$ has size at most*

$$q^{hd} + q^{(h-1)(d-1)}\mathcal{L}(h,d,d) + \lfloor \log_q h \rfloor + \lceil \log_q \max(h, d-1) \rceil + 3.$$

In Section 2, we give numeric evidence that the bound in Theorem 1 is asymptotically sharp. As a corollary of Theorem 1, we obtain Bridy's theorem [7].

Theorem 2. *Let $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!]$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$ of height h and degree d. Then the size of the minimal q-automaton generating $a(n)_{n\ge 0}$ is in $(1+o(1))q^{hd}$ as any of q, h, or d tends to infinity and the others remain constant.*

Bridy [7] also showed that the number of states is in $(1+o(1))q^{h+d+g-1}$ as any of q, h, d, g tends to infinity, where g is the genus of P. Since the genus satisfies $g \le (h-1)(d-1)$, Bridy obtains the bound $(1+o(1))q^{hd}$ for the number of states. Let G be the number of interior points in the Newton polygon of P. We have $g \le G$ by Baker's theorem [5], with equality generically. In our setting, one could use G to obtain more refined bounds than in Theorem 1, analogous to Bridy's bound. This approach is discussed briefly in [3, Section 6].

Broadly, the proof of Theorem 1 consists of two steps. First, in Section 3, we represent states in the automaton with bivariate polynomials, and we establish basic properties of a space W of bivariate polynomials containing most of the automaton's states. Namely, W contains all states except those in the orbit $\mathrm{orb}_{\lambda_{0,0}}(S_0)$ of the initial state S_0 under the linear transformation $\lambda_{0,0}(S) := \Lambda_{0,0}\big(SQ^{q-1}\big)$, where $\Lambda_{0,0}$ is a Cartier operator and $Q = P/y$. The space W has size q^{hd}, giving the main term in Theorem 1. This first step is elementary and yields an initial upper bound of $q^{hd} + |\mathrm{orb}_{\lambda_{0,0}}(S_0)|$ for the number of states.

The second step is considerably more involved. We show that the size of an orbit under $\lambda_{0,0}$ is small, giving the lower-order terms in Theorem 1. The key idea is that one can bound the orbit size under $\lambda_{0,0}$ in terms of the orbit sizes under restrictions

of $\lambda_{0,0}$ to four subspaces. One subspace has size $q^{(h-1)(d-1)}$. On the other three, the operator $\lambda_{0,0}$ behaves like linear transformations $\lambda_0(S) := \Lambda_0\big(SR^{q-1}\big)$ on univariate polynomials for certain Laurent polynomials R, where Λ_0 is a Cartier operator. We show how to bound the orbit size under λ_0 in terms of the factorization of R, using the period length of the coefficient sequence of the series $\frac{1}{R}$. Surprisingly, this period length is not dependent on q; this appears starting in Theorem 23.

Our proof of Theorem 1 begins by converting the representation of the series F by a polynomial P to a representation as the diagonal of a rational function. More generally, in Theorem 35 we use the same two steps to bound the automaton size for the diagonal of a rational function in two variables. For more than two variables, new techniques would be needed to further extend the second step. The current techniques would only give an analogue of Corollary 10.

This analogue is already included in recent work by Adamczewski, Bostan, and Caruso [2], who bound the dimension of a vector space containing the kernel (see Section 3) of a multidimensional algebraic sequence, generalizing a result of Bostan, Caruso, Christol, and Dumas [6] for one-dimensional algebraic sequences. These papers also use diagonals, and the argument fundamentally follows the lines of a multivariate version of Section 3 below. However, like Bridy, the authors of [2] give a more refined bound in terms of the genus of the associated surface. They also give several applications of their bound, establishing a polynomial bound on the algebraic degree of reductions modulo p of diagonals of multivariate algebraic power series, answering a question of Deligne [12], and improving Harase's bound [15] on the degree of the Hadamard product of two algebraic power series.

In Section 3, we lay the groundwork and obtain a preliminary, coarser bound on the size of the automaton, in Corollary 11. In Section 4, we study the linear structure of the operator $\lambda_{0,0}$ and show in Proposition 13 that it can be emulated by univariate operators λ_0 on certain subspaces of $\mathbb{F}_q[z]$. In Section 5, we bound the orbit size of a polynomial under λ_0, leading to Theorem 30. Finally in Section 6, we tie these results together to obtain Theorem 1 and an analogous result for diagonals of rational functions. In Section 7, we give some intriguing conjectures about orbits under λ_0 that were discovered in the process of proving Theorem 1 and ultimately not used.

2. Numeric evidence for sharpness

In this section, we systematically find Furstenberg series, represented by polynomials P, for which the corresponding automata are large. The computations are performed with the Mathematica package IntegerSequences [18, 19].

For fixed values of q, h, and d, we generate all polynomials $P \in \mathbb{F}_q[x, y]$ with height h and degree d that satisfy the conditions in the definition of a Furstenberg series. We also require that the coefficient of x^0y^1 in P is 1, since P and cP (where $c \neq 0$) define the same series F and produce the same automaton. Then, for each P, we use the construction described in Section 3 below to compute an automaton generating the coefficient sequence of its associated Furstenberg series. In general, this construction does not produce a minimal automaton. Minimizing is costly, so to expand the feasible search space we do not minimize automata at this step. Instead, we determine the size of each unminimized automaton, select one of the polynomials P that maximizes this size, and minimize its automaton.

Table 1 in the appendix lists the maximum unminimized automaton size for several values of q, h, and d, along with one polynomial that achieves this size and the value of the bound in Theorem 1. For each polynomial in Table 1, the automaton size drops by at most 1 during minimization. This justifies the decision to not minimize all automata initially. For $d = 1$ (that is, rational series), Bridy showed that the bound $(1 + o(1))q^{hd}$ is sharp by constructing polynomials P from univariate primitive polynomials [7, Proposition 3.14]. Table 1 suggests this bound is also sharp for $d \geq 2$. For $d = 2$, Figure 2 in the appendix shows the distribution of unminimized automaton sizes for some values of q and h by plotting the number of polynomials with each size.

Most of the article is concerned with bounding the orbit sizes of polynomials under the operator $\lambda_{0,0}$. This will yield the terms other than q^{hd} in Theorem 1. Table 2 lists the maximum orbit size under $\lambda_{0,0}$ for several values of q, h, and d.

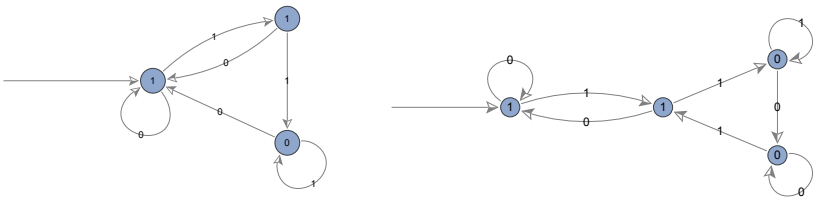
Whereas the polynomials in Table 1 produce automata close to the upper bound, some algebraic sequences that arise in combinatorics, when reduced modulo p, are generated by rather small automata. For example, let $C(n)$ be the nth Catalan number [23, A000108]. Its generating series $F = 1 + x + 2x^2 + 5x^3 + \cdots$ satisfies $xF^2 - F + 1 = 0$, so $h = 1$ and $d = 2$. Burns [8, Section 4] gave an explicit construction for an automaton that generates $(C(n) \bmod p)_{n\geq 0}$. This automaton has only $p+3$ states, compared to the bound $p^2+\mathcal{L}(1, 2, 2)+3 = p^2+5$ in Theorem 1.

3. The vector space of possible states

Christol's theorem implies that an algebraic sequence of elements in $\mathbb{F}_q$ is q-automatic. In this section, we establish a correspondence between states of an automaton generating such a sequence and polynomials in a finite-dimensional $\mathbb{F}_q$-vector space. We do this by converting states in the automaton first to sequences, then to power series, and finally to polynomials. This correspondence provides the foundation for the rest of the article, and we use it to give a preliminary upper bound on the number of states in Corollary 11.

We assume the reader is familiar with deterministic finite automata with output. See [4] for a comprehensive treatment and [20] for a short introduction. An automaton with input alphabet $\{0, 1, \ldots, q-1\}$ generates the q-automatic sequence $a(n)_{n\geq 0}$, where $a(n)$ is the output of the automaton when fed the standard base-q representation of n, starting with the least significant digit. In general, automata are sensitive to leading 0s; that is, the output changes when fed a nonstandard representation of n. One can always produce an automaton without this drawback [4, Theorem 5.2.3], although the number of states may increase.

Example 3. The two automata



generate the same 2-automatic sequence $1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$. The behavior of the first automaton is affected by leading 0s; for example, feeding 11

into this automaton produces the output 0, whereas the input 011 produces the output 1. The behavior of the second automaton is not affected by leading 0s, and in fact this is the smallest automaton with this property for this sequence.

Given a q-automatic sequence $a(n)_{n\geq 0}$, we refer to the smallest automaton that generates $a(n)_{n\geq 0}$ and that is not affected by leading 0s as its *minimal automaton*. Theorem 1 gives an upper bound on the number of states in the minimal automaton. Theorem 1 also gives an upper bound on the size of the *q-kernel* of $a(n)_{n\geq 0}$, defined as

$$\ker_q(a(n)_{n\geq 0}) := \{a(q^e n + r)_{n\geq 0} : e \geq 0 \text{ and } 0 \leq r \leq q^e - 1\}.$$

A sequence is q-automatic if and only if its q-kernel is finite; this is known as Eilenberg's theorem. Moreover, the states of the minimal automaton are in bijection with the elements of the q-kernel.

We then represent kernel sequences $a(q^e n + r)_{n\geq 0}$ by their generating series $\sum_{n\geq 0} a(q^e n + r)x^n$. Let $\mathbb{F}_q[\![x]\!]$ and $\mathbb{F}_q[\![x, y]\!]$ denote the sets of univariate and bivariate power series with coefficients in $\mathbb{F}_q$. Analogously, $\mathbb{F}_q[x]$ and $\mathbb{F}_q[x, y]$ denote sets of polynomials. Elements of the q-kernel (and therefore states in the minimal automaton) can be accessed by applying the following operators.

Definition. Let $n \in \mathbb{Z}$. For each $r \in \{0, 1, \ldots, q-1\}$, define the *Cartier operator* Λ_r on the monomial x^n by

$$\Lambda_r(x^n) = \begin{cases} x^{\frac{n-r}{q}} & \text{if } n \equiv r \mod q \\ 0 & \text{otherwise.} \end{cases}$$

Then extend Λ_r linearly to polynomials (as well as to Laurent polynomials and Laurent series) in x with coefficients in $\mathbb{F}_q$. In particular, for polynomials we have

$$\Lambda_r\left(\sum_{n=0}^{N} a(n)x^n\right) = \sum_{n=0}^{\lfloor N/q \rfloor} a(qn + r)x^n.$$

Similarly, for $m, n \in \mathbb{Z}$ and $r, s \in \{0, 1, \ldots, q-1\}$, define the bivariate Cartier operator

$$\Lambda_{r,s}(x^m y^n) = \begin{cases} x^{\frac{m-r}{q}} y^{\frac{n-s}{q}} & \text{if } m \equiv r \mod q \text{ and } n \equiv s \mod q \\ 0 & \text{otherwise,} \end{cases}$$

and extend $\Lambda_{r,s}$ linearly to bivariate polynomials (as well as to Laurent polynomials and Laurent series).

The map Λ_r on $\mathbb{F}_q[\![x]\!]$ corresponds to the map $a(n)_{n\geq 0} \mapsto a(qn + r)_{n\geq 0}$. An advantage of representing sequences by power series is that a factor of the form F^q can be pulled out of a Cartier operator, as in the following proposition. We will use this repeatedly. The univariate case is proved in [4, Lemma 12.2.2] for power series; the Laurent series and bivariate cases are similar.

Proposition 4. *If F and G are Laurent series in x with coefficients in $\mathbb{F}_q$, then $\Lambda_r(GF^q) = \Lambda_r(G)F$. Similarly, if $F, G \in \mathbb{F}_q[\![x, y]\!]$, then $\Lambda_{r,s}(GF^q) = \Lambda_{r,s}(G)F$.*

The final step is to use a theorem of Furstenberg [14] to convert each algebraic power series $\sum_{n\geq 0} a(q^e n + r)x^n$ corresponding to a kernel sequence to the diagonal of a rational function. Since different rational functions can have the same diagonal, a given kernel sequence is potentially the diagonal of several rational functions that

arise, so the resulting automaton is not necessarily minimal. However, the number of distinct rational functions that arise is an upper bound on the size of the kernel. Furstenberg's theorem holds more generally over every field, but we state it for $\mathbb{F}_q$. The *diagonal operator* $\mathcal{D}\colon \mathbb{F}_q[\![x,y]\!] \to \mathbb{F}_q[\![x]\!]$ is defined by

$$\mathcal{D}\left(\sum_{m\geq 0}\sum_{n\geq 0} a(m,n)x^m y^n\right) = \sum_{n\geq 0} a(n,n)x^n.$$

For a bivariate series or polynomial P, define $P(a,b)$ to be $P|_{x\to a,y\to b}$, and similarly for univariate series; for example, if $P = 3x+2y+xy$, then $P(xy,y) = 3xy+2y+xy^2$, $\frac{\partial P}{\partial y} = 2+x$, $\frac{\partial P}{\partial y}(xy,y) = 2+xy$, and $\frac{\partial P}{\partial y}(0,0) = 2$.

Recall the definition of a Furstenberg series from the introduction.

Theorem 5 (Furstenberg)**.** *Let $F \in \mathbb{F}_q[\![x]\!]$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$. Then*

$$F = \mathcal{D}\left(\frac{y\frac{\partial P}{\partial y}(xy,y)}{P(xy,y)/y}\right).$$

The conditions $P(0,0) = 0$ and $\frac{\partial P}{\partial y}(0,0) \neq 0$ guarantee that every monomial in $P(xy,y)$ is divisible by y and that

$$\frac{y\frac{\partial P}{\partial y}(xy,y)}{P(xy,y)/y} \tag{1}$$

has a unique power series expansion.

Now applying a Cartier operator to the diagonal of a rational power series produces another diagonal of a rational power series, namely

$$\Lambda_r\mathcal{D}\left(\frac{S}{Q}\right) = \mathcal{D}\Lambda_{r,r}\left(\frac{S}{Q}\right) = \mathcal{D}\Lambda_{r,r}\left(\frac{SQ^{q-1}}{Q^q}\right) = \mathcal{D}\left(\frac{\Lambda_{r,r}\left(SQ^{q-1}\right)}{Q}\right), \tag{2}$$

where the last equality follows from Proposition 4. Since the initial and final rational series in Equation (2) have the same denominator Q, every sequence in the q-kernel of $a(n)_{n\geq 0}$, and hence every state in the automaton, is the diagonal of a rational function with denominator Q. Therefore we can represent each state simply by its numerator, and the map $S \mapsto \Lambda_{r,r}\left(SQ^{q-1}\right)$ on $\mathbb{F}_q[x,y]$ emulates the Cartier operator Λ_r on $\mathbb{F}_q[\![x]\!]$. Moreover, the common denominator Q is the denominator of the rational expression corresponding to $a(n)_{n\geq 0}$ itself, which is $P(xy,y)/y$ by Theorem 5. This is the approach taken elsewhere [13, 1, 21].

However, in this article we shear the bivariate series (1) by replacing x with xy^{-1}, obtaining

$$\frac{y\frac{\partial P}{\partial y}}{P/y} \tag{3}$$

instead. The diagonal of (1) is the y^0 row of (3). The latter is significantly more convenient notationally for obtaining the desired bound. Let $Q := P/y$ be the denominator. Note that Q is a polynomial in x, but it may be a Laurent polynomial in y and not a polynomial. This will not cause us trouble, but we mention that, to expand (3) as a series and get the intended row sequence, we should expand using the constant term of the denominator Q (since it is the same as the constant term

of $P(xy, y)/y$) and not a monomial involving y^{-1} if present. The series expansion of (3) is a power series in x but may have terms involving y^n with negative n.

In this formulation, the diagonal operator is replaced by the *center row operator* $\mathcal{C}$, defined by

$$\mathcal{C}\left(\sum_{m\geq 0}\sum_{n\in\mathbb{Z}} a(m,n)x^m y^n\right) = \sum_{m\geq 0} a(m,0)x^m.$$

Equation (2) becomes

$$\Lambda_r\mathcal{C}\left(\frac{S}{Q}\right) = \mathcal{C}\Lambda_{r,0}\left(\frac{S}{Q}\right) = \mathcal{C}\Lambda_{r,0}\left(\frac{SQ^{q-1}}{Q^q}\right) = \mathcal{C}\left(\frac{\Lambda_{r,0}\left(SQ^{q-1}\right)}{Q}\right).$$

Therefore the map $S \mapsto \Lambda_{r,0}\left(SQ^{q-1}\right)$ on $\mathbb{F}_q[x, y]$ emulates the Cartier operator Λ_r on $\mathbb{F}_q[\![x]\!]$. We represent each state in the automaton by a polynomial $S \in \mathbb{F}_q[x, y]$. The initial state is $S_0 := y\frac{\partial P}{\partial y}$, since this is the numerator of the rational series corresponding to $a(n)_{n\geq 0}$. From each state S, upon reading $r \in \{0, 1, \ldots, q-1\}$ we transition to $\Lambda_{r,0}\left(SQ^{q-1}\right)$. The output assigned to each state S is $\frac{S(0,0)}{Q(0,0)}$.

Remark 6. If $r = 0$, then the output assigned to the state S is the same as the output assigned to $\Lambda_{0,0}\left(SQ^{q-1}\right)$ since the constant term of Q^{q-1} is 1 by the assumption $\frac{\partial P}{\partial y}(0, 0) \neq 0$. Therefore, the constructed automaton is not sensitive to leading 0s.

We solidify our notation as follows.

Notation. For the remainder of the article, we fix a prime power q and a polynomial $P \in \mathbb{F}_q[x, y]$ with height $h \geq 1$ and degree $d \geq 1$. We assume that $P(0, 0) = 0$ and $\frac{\partial P}{\partial y}(0, 0) \neq 0$ so that we obtain the Furstenberg series $F \in \mathbb{F}_q[\![x]\!]$ given by Theorem 5. Let $Q = P/y$. For each $r \in \{0, 1, \ldots, q-1\}$ and each $S \in \mathbb{F}_q[x, y]$, define

$$\lambda_{r,0}(S) := \Lambda_{r,0}\left(SQ^{q-1}\right).$$

Note that $\lambda_{r,0}$ depends on Q, even though the notation does not reflect this. Let W be the $\mathbb{F}_q$-vector space defined by

$$W := \left\langle x^i y^j : 0 \leq i \leq h-1 \text{ and } 0 \leq j \leq d-1 \right\rangle.$$

We will always use this basis of W. We have $\dim W = hd$, so $|W| = q^{hd}$.

Example 7. Let $q = 3$, and consider the polynomial

$$P = (x^2 + x + 2)y^4 + xy^3 + (2x + 1)y^2 + (x^2 + 1)y + 2x^2 + x \in \mathbb{F}_3[x, y]$$

with height $h = 2$ and degree $d = 4$. We will use this polynomial as a running example throughout the paper. The coefficient sequence $a(n)_{n\geq 0}$ of the series $F \in \mathbb{F}_3[\![x]\!]$ satisfying $P(x, F) = 0$ is

$$0, 2, 0, 2, 0, 2, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 2, 1, 1, 2, 0, 0, 2, 2, 1, 0, 1, \ldots.$$

We have

$$Q = P/y = (x^2 + x + 2)y^3 + xy^2 + (2x + 1)y + x^2 + 1 + (2x^2 + x)y^{-1}.$$

The initial state is

$$S_0 = y\frac{\partial P}{\partial y} = (x^2 + x + 2)y^4 + (x + 2)y^2 + (x^2 + 1)y.$$

The space W consists of all bivariate polynomials with height at most 1 and degree at most 3.

In the remainder of this section, we use elementary methods to highlight the relevance of W, leading us to a preliminary bound on the size of the kernel in Corollary 11.

Proposition 8 shows that W is closed under $\lambda_{r,0}$. In particular, even though Q is possibly a Laurent polynomial, $\lambda_{r,0}(S)$ is a polynomial for each $S \in W$.

Proposition 8. *For each $r \in \{0, 1, \ldots, q-1\}$, we have $\lambda_{r,0}(W) \subseteq W$.*

Proof. Let $c\,x^I y^J$ be a nonzero monomial in the Laurent polynomial Q^{q-1}. The height I of this monomial satisfies $0 \leq I \leq (q-1)h$, and the degree J satisfies $-(q-1) \leq J \leq (q-1)(d-1)$. To show that $\lambda_{r,0}(W) \subseteq W$, by linearity it suffices to show that if $x^i y^j$ is an element of the basis of W then $\Lambda_{r,0}\big(x^i y^j \cdot x^I y^J\big) \in W$. One computes

$$\Lambda_{r,0}\big(x^i y^j \cdot x^I y^J\big) = \begin{cases} x^{\frac{i+I-r}{q}} y^{\frac{j+J}{q}} & \text{if } i+I \equiv r \bmod q \text{ and } j+J \equiv 0 \bmod q \\ 0 & \text{otherwise.} \end{cases}$$

In the second case, clearly the monomial 0 belongs to W. In the first case, the height of this monomial satisfies $\frac{i+I-r}{q} \geq 0$. We use the fact that $\frac{i+I-r}{q}$ and $\frac{j+J}{q}$ are integers. Then

$$\frac{i+I-r}{q} = \left\lfloor \frac{i+I-r}{q} \right\rfloor \leq \left\lfloor \frac{(h-1)+(q-1)h-r}{q} \right\rfloor = \left\lfloor \frac{qh-r-1}{q} \right\rfloor \leq h-1. \tag{4}$$

The degree of $\Lambda_{r,0}\big(x^i y^j \cdot x^I y^J\big)$ satisfies $\frac{j+J}{q} \leq \frac{(d-1)+(q-1)(d-1)}{q} = d-1$ and

$$\frac{j+J}{q} = \left\lceil \frac{j+J}{q} \right\rceil \geq \left\lceil \frac{0-(q-1)}{q} \right\rceil = \left\lceil \frac{1}{q} - 1 \right\rceil = 0.$$

Consequently $\Lambda_{r,0}\big(x^i y^j \cdot x^I y^J\big) \in W$. □

The initial state $y\frac{\partial P}{\partial y}$ has degree at most d. Since elements of W have degree at most $d-1$, the initial state is not necessarily an element of W. However, the following result shows that most of its images under compositions of $\lambda_{r,0}$ are elements of W. A similar result was obtained in [6, Remark 2.6].

Proposition 9. *Let $S \in \mathbb{F}_q[x,y]$ such that $\deg_x S \leq h$ and $\deg_y S \leq d$.*

- *We have $\deg_x \lambda_{0,0}(S) \leq h$ and $\deg_y \lambda_{0,0}(S) \leq d-1$.*
- *For each $r \in \{1, \ldots, q-1\}$, we have $\lambda_{r,0}(S) \in W$.*

In particular, every polynomial $(\lambda_{r_n,0} \circ \cdots \circ \lambda_{r_2,0} \circ \lambda_{r_1,0})(S_0)$, where at least one r_i is not 0, is an element of W.

Proof. For the first statement, we follow the proof of Proposition 8. After setting $r = 0$, Equation (4) is replaced with

$$\frac{i+I}{q} = \left\lfloor \frac{i+I}{q} \right\rfloor \leq \left\lfloor \frac{h+(q-1)h}{q} \right\rfloor = \left\lfloor \frac{qh}{q} \right\rfloor = h.$$

Therefore $\deg_x S \leq h$. Similarly, in the case that $\Lambda_{0,0}\big(x^i y^j \cdot x^I y^J\big)$ is not 0, its degree satisfies

$$\frac{j+J}{q} = \left\lfloor \frac{j+J}{q} \right\rfloor \leq \left\lfloor \frac{d+(q-1)(d-1)}{q} \right\rfloor = \left\lfloor \frac{q(d-1)+1}{q} \right\rfloor = d-1.$$

For the second statement, we also follow the proof of Proposition 8; Equation (4) is replaced with

$$\tfrac{i+I-r}{q} = \left\lfloor \tfrac{i+I-r}{q} \right\rfloor \le \left\lfloor \tfrac{h+(q-1)h-r}{q} \right\rfloor = \left\lfloor \tfrac{qh-r}{q} \right\rfloor \le h-1,$$

and analogously

$$\tfrac{j+J}{q} = \left\lfloor \tfrac{j+J}{q} \right\rfloor \le \left\lfloor \tfrac{d+(q-1)(d-1)-r}{q} \right\rfloor = \left\lfloor \tfrac{q(d-1)+1-r}{q} \right\rfloor \le d-1.$$

The initial state $S_0 = y\frac{\partial P}{\partial y}$ satisfies $\deg_x S_0 \le h$ and $\deg_y S_0 \le d$. Let $r \in \{1, \dots, q-1\}$. Applying both statements, we obtain $(\lambda_{r,0} \circ \lambda_{0,0}^n)(S_0) \in W$ for all $n \ge 1$. Therefore, by Proposition 8, the final statement follows. □

An immediate corollary of Propositions 8 and 9 is the following, since all states S except the initial state satisfy $\deg_x S \le h$ and $\deg_y S \le d-1$.

Corollary 10. *Let $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!]$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$ of height h and degree d. Then*

$$|\ker_q(a(n)_{n\ge 0})| \le q^{(h+1)d} + 1.$$

Proposition 9 indicates that we must further study $\lambda_{0,0}$ to lower the bound in Corollary 10. We start to do this next. For a function $f\colon X \to X$, define the *orbit* of $S \in X$ under f to be the sequence $S, f(S), f^2(S), \dots$, and let $|\mathrm{orb}_f(S)|$ be the number of distinct terms in the orbit.

Corollary 11. *Let $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!]$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$ of height h and degree d. Then*

$$|\ker_q(a(n)_{n\ge 0})| \le q^{hd} + |\mathrm{orb}_{\Lambda_0}(F)|.$$

Proof. Recall that each sequence in $\ker_q(a(n)_{n\ge 0})$ is represented by at least one polynomial obtained by iteratively applying some sequence of the operators $\lambda_{r,0}$ to the initial state $S_0 = y\frac{\partial P}{\partial y}$. Applying $\lambda_{0,0}$ iteratively to S_0 produces $|\mathrm{orb}_{\lambda_{0,0}}(S_0)|$ states, and $|\mathrm{orb}_{\lambda_{0,0}}(S_0)| = |\mathrm{orb}_{\Lambda_0}(F)|$ by definition. By Propositions 8 and 9, all states that are not in $\mathrm{orb}_{\lambda_{0,0}}(S_0)$ are in W, which has size q^{hd}. □

4. Structure of the linear transformation $\lambda_{0,0}$

By Corollary 11, it remains to bound $|\mathrm{orb}_{\Lambda_0}(F)|$. In this section, we take the first step toward this goal by identifying univariate operators λ_0 that emulate $\lambda_{0,0}$ on three subspaces. The main result is Proposition 13.

We continue to use the notation h, d, P, Q, W established in the previous section. As we saw with regard to Proposition 9, the elements of $\mathrm{orb}_{\lambda_{0,0}}(S_0)$ do not necessarily belong to W. However, they do belong to the slightly larger space

$$V := \left\langle x^i y^j : 0 \le i \le h \text{ and } 0 \le j \le d-1 \right\rangle.$$

We define three subspaces of V, which we label suggestively using ℓ (left), r (right), and t (top):

$$V_\ell = \left\langle x^0 y^j : 0 \le j \le d-1 \right\rangle$$
$$V_{\mathrm{r}} = \left\langle x^h y^j : 0 \le j \le d-1 \right\rangle$$
$$V_{\mathrm{t}} = \left\langle x^i y^{d-1} : 0 \le i \le h \right\rangle.$$

We also define the *interior* of V to be

$$V^\circ = \left\langle x^i y^j : 1 \le i \le h-1 \text{ and } 0 \le j \le d-2 \right\rangle. \tag{5}$$

Note that, despite the name "interior", the basis of V° contains monomials $x^i y^0$ along the bottom edge of the rectangle. We have $V^\circ \cap V_\ell = V^\circ \cap V_{\mathrm{r}} = V^\circ \cap V_{\mathrm{t}} = \{0\}$. We will see that the factor $q^{(h-1)(d-1)}$ in Theorem 1 comes from the size of V°.

To establish the structure of $\lambda_{0,0}$, we introduce three projection-like maps.

Notation. Let $\pi_\ell \colon \mathbb{F}_q[x,y] \to \mathbb{F}_q[y]$ denote the projection map from $\mathbb{F}_q[x,y]$ to V_ℓ. We define π_{r} slightly differently. We have $V_{\mathrm{r}} \subset x^h \mathbb{F}_q[y]$, so rather than projecting to V_{r} we will dispense with the factor x^h. Namely, define $\pi_{\mathrm{r}} \colon \mathbb{F}_q[x,y] \to \mathbb{F}_q[y]$ by $\pi_{\mathrm{r}}(S) = \frac{1}{x^h}\rho(S)$, where ρ projects from $\mathbb{F}_q[x,y]$ to V_{r}. Similarly, define $\pi_{\mathrm{t}} \colon \mathbb{F}_q[x,y] \to \mathbb{F}_q[x]$ by $\pi_{\mathrm{t}}(S) = \frac{1}{y^{d-1}}\rho(S)$, where ρ projects from $\mathbb{F}_q[x,y]$ to V_{t}.

Example 12. As in Example 7, let $q = 3$ and

$$Q = (x^2+x+2)y^3 + xy^2 + (2x+1)y + x^2 + 1 + (2x^2+x)y^{-1}$$
$$S_0 = (x^2+x+2)y^4 + (x+2)y^2 + (x^2+1)y.$$

The second state in the orbit of S_0 under $\lambda_{0,0}$ is

$$S_1 := \lambda_{0,0}(S_0) = \Lambda_{0,0}(S_0 Q^{3-1}) = xy^3 + (x^2+x+1)y^2 + (2x^2+2)y + x^2 + x.$$

We have $S_1 \in V$, which is consistent with Proposition 9. Since $h = 2$ and $d = 4$, the images of S_1 under $\pi_\ell, \pi_{\mathrm{r}}, \pi_{\mathrm{t}}$ are

$$\pi_\ell(S_1) = y^2 + 2y$$
$$\pi_{\mathrm{r}}(S_1) = y^2 + 2y + 1$$
$$\pi_{\mathrm{t}}(S_1) = x.$$

We use these projections in Example 14 below.

Next we define univariate versions of $\lambda_{0,0}$. We will use the symbol z to denote either x or y, depending on which subspace we are considering.

Notation. Let $R \in z^{-1}\mathbb{F}_q[z]$. Define $\lambda_0 \colon \mathbb{F}_q[z] \to \mathbb{F}_q[z]$ by

$$\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big). \tag{6}$$

The next proposition shows that λ_0 emulates $\lambda_{0,0}$ on the subspaces V_ℓ, V_{r}, and V_{t}. Each of the three statements describes a commuting diagram. For example, the first statement says that the diagram

$$\begin{array}{ccc} \mathbb{F}_q[x,y] & \xrightarrow{\lambda_{0,0}} & \mathbb{F}_q[x,y] \\ {\scriptstyle \pi_\ell}\downarrow & & \downarrow{\scriptstyle \pi_\ell} \\ \mathbb{F}_q[y] & \xrightarrow{\lambda_0} & \mathbb{F}_q[y] \end{array}$$

commutes. Write

$$P(x,y) = \sum_{i=0}^{h} x^i A_i(y) = \sum_{j=0}^{d} B_j(x) y^j. \tag{7}$$

Note that A_0/y is a polynomial since we assume $P(0,0) = 0$ for a Furstenberg series.

Proposition 13. *We have the following.*

(1) *Let $R = A_0/y$. For all $S \in \mathbb{F}_q[x,y]$,*

$$\pi_\ell(\lambda_{0,0}(S)) = \lambda_0(\pi_\ell(S)).$$

(2) *Let $R = A_h/y$. For all $S \in \mathbb{F}_q[x,y]$ with height at most h,*

$$\pi_{\mathrm{r}}(\lambda_{0,0}(S)) = \lambda_0(\pi_{\mathrm{r}}(S)).$$

In particular, $\lambda_0(\pi_{\mathrm{r}}(S))$ is a polynomial despite R not necessarily being a polynomial.

(3) *Let $R = B_d$. For all $S \in \mathbb{F}_q[x,y]$ with degree at most $d-1$,*

$$\pi_{\mathrm{t}}(\lambda_{0,0}(S)) = \lambda_0(\pi_{\mathrm{t}}(S)).$$

In particular, Proposition 13 implies that the V_ℓ, V_{r}, and V_{t} components of $\lambda_{0,0}(S)$ depend only on the respective V_ℓ, V_{r}, and V_{t} components of S.

Example 14. For the polynomial P in Example 7, we have

$$\begin{aligned} A_0/y &= 2y^3 + y + 1 \\ A_h/y &= y^3 + 1 + 2y^{-1} \\ B_d &= x^2 + x + 2. \end{aligned}$$

With these respective values of R, the second state S_1 in the orbit of S_0 under $\lambda_{0,0}$, computed in Example 12, satisfies

$$\begin{aligned} \pi_\ell(\lambda_{0,0}(S_1)) &= y^2 + y = \lambda_0(\pi_\ell(S_1)) \\ \pi_{\mathrm{r}}(\lambda_{0,0}(S_1)) &= y^2 + y + 1 = \lambda_0(\pi_{\mathrm{r}}(S_1)) \\ \pi_{\mathrm{t}}(\lambda_{0,0}(S_1)) &= 2x = \lambda_0(\pi_{\mathrm{t}}(S_1)), \end{aligned}$$

confirming the statement of Proposition 13. That is, Proposition 13 reduces the computation of $\pi_\ell(\lambda_{0,0}(S_1))$ to the univariate computation of $\lambda_0(\pi_\ell(S_1))$, and similarly for π_{r} and π_{t}.

Proof of Proposition 13. First we consider $\pi_\ell(\lambda_{0,0}(S))$ for $S \in \mathbb{F}_q[x,y]$. Since π_ℓ projects onto polynomials in y, we are interested in monomials with height 0 in $\lambda_{0,0}(S) = \Lambda_{0,0}\big(SQ^{q-1}\big)$. A monomial $c\,x^0y^J$ in SQ^{q-1} arises only from the product of a monomial in S with height 0 together with a monomial in Q^{q-1} with height 0, that is, only from the product of a monomial in $\pi_\ell(S)$ together with a monomial in Q^{q-1} with height 0. Therefore

$$\pi_\ell(\lambda_{0,0}(S)) = \pi_\ell(\lambda_{0,0}(\pi_\ell(S))).$$

Additionally, the only way to get a monomial in Q^{q-1} with height 0 is to take a product of $q-1$ monomials in $Q = P/y$ with height 0, namely, monomials in A_0/y. Therefore,

$$\pi_\ell(\lambda_{0,0}(S)) = \pi_\ell\big(\Lambda_{0,0}\big(\pi_\ell(S) \cdot (A_0/y)^{q-1}\big)\big)\,.$$

Since $\pi_\ell(S) \cdot (A_0/y)^{q-1}$ is a univariate polynomial in y, we obtain

$$\pi_\ell(\lambda_{0,0}(S)) = \Lambda_0\big(\pi_\ell(S)(A_0/y)^{q-1}\big) = \lambda_0(\pi_\ell(S)).$$

The argument is similar for $\pi_{\mathrm{r}}(\lambda_{0,0}(S))$. Let $\deg_x S \le h$. We have $\pi_{\mathrm{r}}(x^Iy^J) = 0$ if $I \ne h$. Since $\deg_x Q = h$, each monomial $c\,x^{qh}y^J$ in each of $S \cdot Q^{q-1}$ and $x^h\pi_{\mathrm{r}}(S) \cdot Q^{q-1}$ arises only from the product of a monomial in $x^h\pi_{\mathrm{r}}(S)$ together

with a product of $q-1$ monomials in Q with height h, namely, monomials in $x^h A_h/y$. Therefore

$$\begin{aligned}\pi_{\mathrm{r}}(\lambda_{0,0}(S)) = \pi_{\mathrm{r}}(\lambda_{0,0}(x^h\pi_{\mathrm{r}}(S))) &= \pi_{\mathrm{r}}\big(\Lambda_{0,0}\big(x^h\pi_{\mathrm{r}}(S)\cdot(x^hA_h/y)^{q-1}\big)\big)\\ &= \pi_{\mathrm{r}}\big(x^h\Lambda_{0,0}\big(\pi_{\mathrm{r}}(S)(A_h/y)^{q-1}\big)\big)\\ &= \Lambda_0\big(\pi_{\mathrm{r}}(S)(A_h/y)^{q-1}\big)\\ &= \lambda_0(\pi_{\mathrm{r}}(S)),\end{aligned}$$

where in the third equality we use Proposition 4 to rewrite $\Lambda_{0,0}(Gx^{hq}) = x^h\Lambda_{0,0}(G)$. Moreover, $\lambda_0(\pi_{\mathrm{r}}(S))$ is a polynomial since monomials in $\pi_{\mathrm{r}}(S)$ have degree at least 0 and monomials in $(A_h/y)^{q-1}$ have degree at least $-(q-1)$.

Finally, we consider $\pi_{\mathrm{t}}(\lambda_{0,0}(S))$ for $\deg_y S \le d-1$. We have $\pi_{\mathrm{t}}(x^Iy^J)=0$ if $J\neq d-1$. Since $\deg_y Q = d-1$, each monomial $c\,x^Iy^{q(d-1)}$ in each of $S\cdot Q^{q-1}$ and $\pi_{\mathrm{t}}(S)y^{d-1}\cdot Q^{q-1}$ arises only from the product of a monomial in $\pi_{\mathrm{t}}(S)y^{d-1}$ together with a product of $q-1$ monomials in Q with degree $d-1$, namely, monomials in B_dy^{d-1}. Therefore

$$\begin{aligned}\pi_{\mathrm{t}}(\lambda_{0,0}(S)) = \pi_{\mathrm{t}}(\lambda_{0,0}(\pi_{\mathrm{t}}(S)y^{d-1})) &= \pi_{\mathrm{t}}\big(\Lambda_{0,0}\big(\pi_{\mathrm{t}}(S)y^{d-1}\cdot(B_dy^{d-1})^{q-1}\big)\big)\\ &= \pi_{\mathrm{t}}\Big(y^{d-1}\Lambda_{0,0}\Big(\pi_{\mathrm{t}}(S)B_d^{q-1}\Big)\Big)\\ &= \Lambda_0\Big(\pi_{\mathrm{t}}(S)B_d^{q-1}\Big)\\ &= \lambda_0(\pi_{\mathrm{t}}(S)).\end{aligned}$$

□

4.1. **The linear structure of $\lambda_{0,0}$.** Proposition 13 identifies three subspaces on which $\lambda_{0,0}$ is equivalent to a univariate operator λ_0. This proposition is sufficient for the proof of Theorem 1, which we resume in Section 5. However, in the remainder of this section we develop additional intuition by using Proposition 13 to refine, in two steps, the standard basis of V to reveal additional structure of the linear transformation $\lambda_{0,0}$ and its corresponding matrix.

Define

$$\begin{aligned}V_\ell^\circ &= \langle x^0y^j : 1\le j\le d-2\rangle\\ V_{\mathrm{r}}^\circ &= \langle x^hy^j : 0\le j\le d-2\rangle\\ V_{\mathrm{t}}^\circ &= \langle x^iy^{d-1} : 1\le i\le h-1\rangle\end{aligned}$$

so that

$$\begin{aligned}V_\ell &= \langle x^0y^0\rangle \oplus V_\ell^\circ \oplus \langle x^0y^{d-1}\rangle\\ V_{\mathrm{r}} &= V_{\mathrm{r}}^\circ \oplus \langle x^hy^{d-1}\rangle\\ V_{\mathrm{t}} &= \langle x^0y^{d-1}\rangle \oplus V_{\mathrm{t}}^\circ \oplus \langle x^hy^{d-1}\rangle\,.\end{aligned}$$

The bases of the seven subspaces

$$V^\circ,\quad V_\ell^\circ,\quad \langle x^0y^0\rangle,\quad V_{\mathrm{t}}^\circ,\quad \langle x^0y^{d-1}\rangle,\quad V_{\mathrm{r}}^\circ,\quad \langle x^hy^{d-1}\rangle \tag{8}$$

are disjoint and form a set partition of the basis of V. Geometrically, these bases are arranged as in Figure 1. We will show in Corollary 17 that, with this decomposition of V, the matrix corresponding to $\lambda_{0,0}$ is block upper triangular. The block sizes are $(h-1)(d-1), d-2, 1, h-1, 1, d-1, 1$.

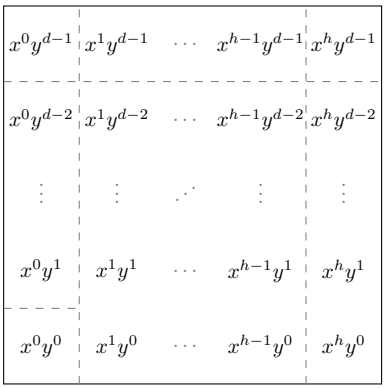


FIGURE 1. Partition of the basis of V into seven sets, which generate the subspaces $\langle x^0y^{d-1}\rangle$, V_{t}°, $\langle x^hy^{d-1}\rangle$, V_{ℓ}°, V°, V_{r}°, and $\langle x^0y^0\rangle$.

Example 15. As in Example 14, let $h = 2$, $d = 4$, and

$$P = (x^2 + x + 2)y^4 + xy^3 + (2x + 1)y^2 + (x^2 + 1)y + 2x^2 + x \in \mathbb{F}_3[x, y].$$

The basis of V, ordered according to (8), is

$$\left(x^1y^0, x^1y^1, x^1y^2, \quad x^0y^1, x^0y^2, \quad x^0y^0, \quad x^1y^3, \quad x^0y^3, \quad x^2y^0, x^2y^1, x^2y^2, \quad x^2y^3\right).$$

With this basis, the operators $\lambda_{0,0}$, $\lambda_{1,0}$, and $\lambda_{2,0}$ are represented by the 12×12 matrices

$$L_{0,0} = \begin{bmatrix}
1 & 1 & 1 & 2 & 1 & 2 & 0 & 0 & 2 & 2 & 0 & 0 \\
1 & 2 & 1 & 2 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 \\
2 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 1 \\
 & & & 1 & 2 & 1 & & 1 & & & & \\
 & & & 0 & 1 & 1 & & 1 & & & & \\
 & & & & & 1 & & & & & & \\
 & & & & & & 2 & 2 & & & & 1 \\
 & & & & & & & 1 & & & & \\
 & & & & & & & & 1 & 1 & 1 & 0 \\
 & & & & & & & & 2 & 1 & 0 & 1 \\
 & & & & & & & & 1 & 0 & 0 & 2 \\
 & & & & & & & & & & & 1
\end{bmatrix}$$

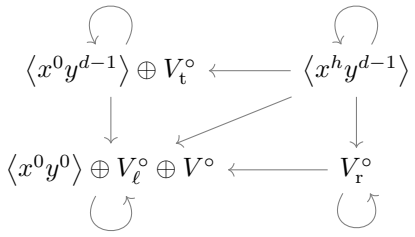
$$L_{1,0} = \begin{bmatrix}
2 & 2 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
2 & 2 & 2 & 1 & 0 & 2 & 2 & 1 & 1 & 2 & 1 & 1 \\
2 & 2 & 1 & 0 & 0 & 1 & 2 & 2 & 2 & 2 & 1 & 1 \\
1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 2 & 1 & 0 & 0 & 0 & 2 \\
 & & & & & & 1 & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 0
\end{bmatrix}$$

$$L_{2,0} = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 0 \\
2 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 2 & 2 & 2 \\
1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 2 & 1 & 2 \\
1 & 1 & 1 & 2 & 1 & 1 & 2 & 1 & 1 & 1 & 2 & 1 \\
1 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 0 & 1 & 1 \\
2 & 2 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & & & 1 & 0 & 0 & 0 & 0 & 2 \\
 & & & & & & 1 & 2 & 0 & 0 & 0 & 1 \\
 & & & & & & & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 0
\end{bmatrix}.$$

The first three columns of $L_{0,0}$ have 0s in rows 4–12, since Proposition 13 tells us that the V° component of S has no impact on the V_{ℓ}, V_{r}, and V_{t} components of

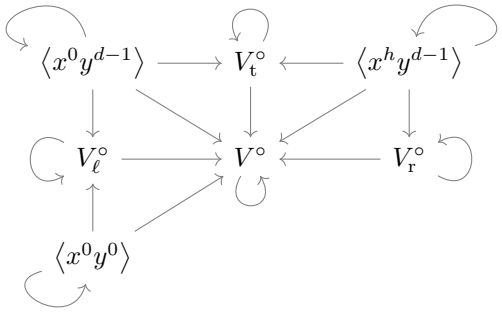
$\lambda_{0,0}(S)$. Conversely, several nonzero entries in the last column of $L_{0,0}$ indicate that the monomial $x^2y^3 \in V_{\mathrm{r}}$ in S has an effect on monomials of $\lambda_{0,0}(S)$ outside of V_{r}. In general, entries guaranteed to be 0 by Theorem 16 below have been omitted from $L_{0,0}$, and entries guaranteed to be 0 by Corollary 17 have been omitted from $L_{1,0}$ and $L_{2,0}$. In addition, the second statement in Proposition 9 implies that the last $d = 4$ rows of $L_{1,0}$ and $L_{2,0}$ are zero rows.

Proposition 13 shows that, under applications of $\lambda_{0,0}$, information flows from V_ℓ to its complement subspace but not in the other direction, and similarly for V_{r} and V_{t}. That is, information flows between four subspaces of V according to the following diagram.



We can refine this further.

Theorem 16. *Under applications of $\lambda_{0,0}$ on V, information flows according to the following diagram. Namely, if $S \in V$ and U is one of the seven distinguished subspaces of V, then the projection of $\lambda_{0,0}(S)$ to U is determined by the projections of S onto the subspaces with arrows pointing to U.*



Proof. We will rule out all arrows that do not appear in the diagram.

Part 1 of Proposition 13 implies that the left subspaces $\langle x^0y^0\rangle$, V_ℓ°, and $\langle x^0y^{d-1}\rangle$ have no incoming arrows from the other four subspaces. Similarly, Part 2 implies that the right subspaces V_{r}° and $\langle x^hy^{d-1}\rangle$ have no incoming arrows from the other five subspaces, and Part 3 implies that the top subspaces $\langle x^0y^{d-1}\rangle$, V_{t}°, and $\langle x^hy^{d-1}\rangle$ have no incoming arrows from the other four subspaces. It follows that the top corner subspaces $\langle x^0y^{d-1}\rangle$ and $\langle x^hy^{d-1}\rangle$ have no incoming arrows other than their loops.

To see that $\langle x^0y^0\rangle$ has no incoming arrows other than its loop, let $j \in \{1, \dots, d-1\}$. We have $\lambda_{0,0}(x^0y^j) = \Lambda_{0,0}(x^0y^jQ^{q-1})$. The coefficient of x^0y^0 in $\lambda_{0,0}(x^0y^j)$ is equal to the coefficient of x^0y^0 in $x^0y^jQ^{q-1}$. However, since $P(0,0) = 0$ and $Q = P/y$, the only monomials x^Iy^J with $J \le -1$ that appear in Q with a nonzero

coefficient satisfy $I \geq 1$. Therefore the coefficient of x^0y^0 in $x^0y^jQ^{q-1}$ is 0, and $\langle x^0y^0 \rangle$ has only one incoming arrow. □

Theorem 16 and Proposition 9 imply the following, where the seven blocks correspond to the seven subspaces in Theorem 16.

Corollary 17. *If the basis of V is ordered according to* (8)*, then the matrix corresponding to $\lambda_{0,0}$ is block upper triangular with seven blocks. Moreover, for all $r \in \{1, 2, \ldots, q-1\}$, the matrix corresponding to $\lambda_{r,0}$ is block upper triangular with four blocks (whose sizes are $h(d-1)$, h, $d-1$, and 1).*

5. Orbit size of a univariate polynomial under λ_0

Proposition 13 (and, more explicitly, Theorem 16) shows that the orbit size of a bivariate polynomial $S \in V$ under $\lambda_{0,0}$ depends in part on the orbit sizes of the univariate polynomials $\pi_\ell(S), \pi_\mathrm{r}(S), \pi_\mathrm{t}(S)$ under λ_0 for the respective values $R = A_0/y, R = A_h/y, R = B_d$. (Recall from Equation (6) that the definition of λ_0 depends on R.) The main result of this section is Theorem 30, which establishes an upper bound on orbit sizes under λ_0 for a general element $R \in z^{-1}\mathbb{F}_q[z]$; this includes the case $R = A_h/y$ (where $z = y$), which is not necessarily a polynomial (for instance, as in Example 14).

We will use the following lemma several times.

Lemma 18. *Let $q \geq 2$.*

- *If $k \in \mathbb{Z}$ and $f(x) = \left\lfloor \frac{x+k(q-1)}{q} \right\rfloor$, then, for every $x \geq k$ and $n \geq \lfloor \log_q(x-k) \rfloor + 1$, we have $f^n(x) = k$.*
- *If $k \geq 1$ and $f(x) = \left\lceil \frac{x+k(q-1)}{q} \right\rceil$, then, for every $x \geq 0$ and $n \geq \lfloor \log_q k \rfloor + 1$, we have $f^n(x) \geq k$.*

Proof. The function $f(x) = \left\lfloor \frac{x+k(q-1)}{q} \right\rfloor = k + \left\lfloor \frac{x-k}{q} \right\rfloor$ has an attracting fixed point k for $x \geq k$. Since $\left\lfloor \frac{\lfloor (x-k)/q^n \rfloor}{q} \right\rfloor = \left\lfloor \frac{x-k}{q^{n+1}} \right\rfloor$, a straightforward induction shows that $f^n(x) = k + \left\lfloor \frac{x-k}{q^n} \right\rfloor$ for all $n \geq 0$. The first statement follows.

For the second statement, we have $f^n(x) = k + \left\lceil \frac{x-k}{q^n} \right\rceil$ for all $n \geq 0$. If $n \geq \lfloor \log_q k \rfloor + 1$, then $\left\lceil -\frac{k}{q^n} \right\rceil = 0$. Therefore $f^n(x) = k + \left\lceil \frac{x-k}{q^n} \right\rceil \geq k + \left\lceil -\frac{k}{q^n} \right\rceil = k$. □

The following proposition shows that if $\deg S > \deg R$ then the orbit of S under λ_0 eventually consists of polynomials with degree at most $\deg R$. Looking ahead, this will let us restrict attention to polynomials S with $\deg S \leq \deg R$ in later results (namely, Theorem 23, Corollary 25, and Theorem 30). We will use it directly in the proof of Lemma 31.

Proposition 19. *Let $R \in z^{-1}\mathbb{F}_q[z]$ be a Laurent polynomial, let $r = \deg R$, and define λ_0 on $\mathbb{F}_q[z]$ by $\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big)$. Let $S \in \mathbb{F}_q[z]$, let $s = \deg S$, and suppose that $s > r$. If $n \geq \lfloor \log_q(s-r) \rfloor + 1$, then $\deg \lambda_0^n(S) \leq r$.*

In particular, if $r = -1$ then $\lambda_0^n(S) = 0$ for sufficiently large n since λ_0 maps polynomials to polynomials.

Proof. We have $\deg \lambda_0(S) \le \frac{s+r(q-1)}{q}$. We track the behavior of $\deg \lambda_0^n(S)$ by iterating the function $f(x) = \left\lfloor \frac{x+r(q-1)}{q} \right\rfloor$. Applying Lemma 18 with $k = r$, the result follows. □

Example 20. Let $q = 3$ and $R = (z^2+1)(z^3+z^2+2) \in \mathbb{F}_3[z]$. By computing $\mathrm{orb}_{\lambda_0}(S)$ from each $S \in \mathbb{F}_3[z]$ with $\deg S \le \deg R = 5$, one finds that each orbit is periodic with period length 1, 2, 3, or 6. For example, the orbit of z^4+z^2 is constant, the orbit of z^2+2z+1 has period length 2, the orbit of $z+2$ has period length 3, and the orbit of 1 has period length 6.

The orbit of $S \in \mathbb{F}_q[z]$ under λ_0 is eventually periodic, since an argument similar to Proposition 19 shows that the elements in the orbit have bounded degree. As we vary q and r and consider all polynomials $R \in \mathbb{F}_q[z]$ with fixed degree $\deg R = r$, one finds that the maximal size of the orbit is independent of q and depends only on r. We prove this in Theorem 23, which is an important step in proving Theorem 30. The proof uses the periodicity of the series expansion of $\frac{1}{R}$ to establish the periodicity of the orbit under λ_0, as in the following example.

Example 21. Let $q = 2$ and $R = z^2+z+1 \in \mathbb{F}_2[z]$. In light of Proposition 19, we consider polynomials $S \in \mathbb{F}_2[z]$ such that $\deg S \le \deg R = 2$. Let $j \in \{0,1,2\}$, so that each monomial in S is of the form $c\,z^j$. Proposition 4 implies $\lambda_0(z^j) = \Lambda_0(z^j R^{q-1}) = \Lambda_0(\frac{z^j}{R})R$. Iterating λ_0 gives $\lambda_0^n(z^j) = \Lambda_0^n(\frac{z^j}{R})R$ for all $n \ge 0$. We show that $\Lambda_0^2(\frac{z^j}{R}) = \frac{z^j}{R}$; this implies $\lambda_0^2(z^j) = z^j$, which, by linearity, implies $\lambda_0^2(S) = S$ for all $S \in \mathbb{F}_2[z]$ with $\deg S \le 2$. We will only use two facts about the series expansion $\sum_{n\ge 0} a(n)z^n := \frac{1}{R} = 1 + 1z + 0z^2 + 1z^3 + 1z^4 + 0z^5 + \cdots$: it is periodic with period length 3, and $a(2) = 0$. We start by rewriting

$$\frac{z^j}{R} = \sum_{n\ge 0} a(n) z^{n+j} = \sum_{n \ge j} a(n-j) z^n.$$

Since $\Lambda_0^2(z^n) = 0$ if $n \not\equiv 0 \mod 4$, this implies

$$\Lambda_0^2\left(\frac{z^j}{R}\right) = \Lambda_0^2\left(\sum_{n \ge \lceil j/4 \rceil} a(4n-j) z^{4n}\right) = \sum_{n \ge \lceil j/4 \rceil} a(4n-j) z^n.$$

If $j = 0$ or $j = 1$, then $\lceil j/4 \rceil = j$, so this series is $\sum_{n\ge j} a(4n-j)z^n$. If $j = 2$, then the coefficient for $n = \lceil j/4 \rceil = 1$ is $a(4\cdot 1 - j) = a(2) = 0$, so again the series is $\sum_{n\ge 2} a(4n-j)z^n = \sum_{n\ge j} a(4n-j)z^n$. Since $a(n)_{n\ge 0}$ is periodic with period length 3, we have $a(4n-j) = a((4n-j) \bmod 3) = a(n-j)$ for all $n \ge j \ge 0$. Therefore

$$\Lambda_0^2\left(\frac{z^j}{R}\right) = \sum_{n\ge j} a(4n-j) z^n = \sum_{n \ge j} a(n-j) z^n = \sum_{n\ge 0} a(n) z^{n+j} = \frac{z^j}{R},$$

as desired.

In general, periodicity of the series expansion of $\frac{1}{R}$ is guaranteed by the following standard argument.

Lemma 22. *Let $R \in \mathbb{F}_q[z]$ be a polynomial with $\deg R \ge 1$. If the coefficient of z^0 is nonzero, then $\frac{1}{R}$ has a power series expansion, and the sequence of coefficients of $\frac{1}{R}$ is periodic.*

Proof. The fact that $\frac{1}{R}$ has a power series expansion follows from $R(0) \neq 0$ and the geometric series formula. Write $\frac{1}{R} = \sum_{n\geq 0} a(n)z^n$. The relation $R\sum_{n\geq 0} a(n)z^n = 1$ gives a recurrence for the coefficient sequence $a(n)_{n\geq 0}$. Since there are only finitely many $(\deg R)$-tuples of elements from $\mathbb{F}_q$, the sequence $a(n)_{n\geq 0}$ is eventually periodic. Since the coefficient of $z^{\deg R}$ is invertible, we can run the recurrence backward as well as forward, so $a(n)_{n\geq 0}$ is periodic. □

The main result of this section is that the size of the orbit under λ_0 is related to the factorization of R. The *factorization into irreducibles* of an element $R \in z^{-1}\mathbb{F}_q[z]$ is $R = cz^{e_0}R_1^{e_1}\cdots R_k^{e_k}$, where $z, R_1, \ldots, R_k \in \mathbb{F}_q[z]$ are distinct, monic, irreducible polynomials, $c \in \mathbb{F}_q$, $e_0 \geq -1$, and $e_i \geq 1$ for all $i \in \{1, \ldots, k\}$. We say that R is *square-free* if $e_i = 1$ for all $i \in \{1, \ldots, k\}$. If $R \in z^{-1}\mathbb{F}_q[z]$ and $R \neq 0$, define $\deg R$ to be the largest exponent of z with a nonzero coefficient in the expansion of R in the monomial basis.

First we establish a bound on the orbit size for certain square-free Laurent polynomials R with positive degree. We use the convention that $\operatorname{lcm}() = 1$.

Theorem 23. *Let $R \in z^{-1}\mathbb{F}_q[z]$ be a nonzero square-free Laurent polynomial such that $\deg R \geq 1$, whose factorization into irreducibles is of the form $cz^{e_0}R_1\cdots R_k$, where $e_0 \in \{-1, 0\}$. Let $\ell = \operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$. Define λ_0 on $\mathbb{F}_q[z]$ by $\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big)$. Then $\lambda_0^\ell(S) = S$ for all $S \in \mathbb{F}_q[z]$ with $\deg S \leq \deg R$.*

To prove Theorem 23, we use the following classical result to bound the period length of the series expansion of $\frac{1}{R}$ and to conclude that certain coefficients are 0.

Proposition 24. *Let $\ell \geq 1$. The product of all monic irreducible polynomials in $\mathbb{F}_q[z]$ with degree dividing ℓ is $z^{q^\ell} - z$.*

Proposition 24 follows from the fact that $\mathbb{F}_{q^\ell}$ is the splitting field of $z^{q^\ell} - z$ over $\mathbb{F}_q$; since each element in $\mathbb{F}_{q^\ell}$ has a minimal polynomial over $\mathbb{F}_q$, the product of all those minimal polynomials is $z^{q^\ell} - z$.

Now we prove Theorem 23.

Proof of Theorem 23. Let $r := \deg R$. Since $e_0 \in \{-1, 0\}$, we have a power series expansion $\frac{1}{R} = \sum_{n\geq 0} a(n)z^n \in \mathbb{F}_q[\![z]\!]$. Let $j \in \{0, 1, \ldots, r\}$. By Proposition 4, $\lambda_0(z^j) = \Lambda_0(z^j R^{q-1}) = \Lambda_0(\frac{z^j}{R})R$. Therefore, by iterating, $\lambda_0^\ell(z^j) = \Lambda_0^\ell(\frac{z^j}{R})R$. We show $\Lambda_0^\ell(\frac{z^j}{R}) = \frac{z^j}{R}$; this implies $\lambda_0^\ell(z^j) = z^j$, and the statement will follow from the linearity of λ_0. Since $\Lambda_0^\ell(z^n) = 0$ if $n \not\equiv 0 \mod q^\ell$, we have

$$\begin{aligned}\Lambda_0^\ell\left(\frac{z^j}{R}\right) &= \Lambda_0^\ell\left(\sum_{n\geq j} a(n-j)z^n\right) = \Lambda_0^\ell\left(\sum_{n\geq \lceil j/q^\ell\rceil} a(q^\ell n - j)z^{q^\ell n}\right)\\ &= \sum_{n\geq \lceil j/q^\ell\rceil} a(q^\ell n - j)z^n.\end{aligned}$$

We will use the fact that the series expansion of $\frac{1}{R}$ is periodic to rewrite $a(q^\ell n - j)$. Since each $\deg R_k$ divides ℓ, Proposition 24 implies that the polynomial $z^{-e_0}R$ divides $z^{q^\ell - 1} - 1$. Write $1 - z^{q^\ell - 1} = RT$ where $T \in z^{-e_0}\mathbb{F}_q[z]$; then

$$\frac{1}{R} = \frac{T}{1 - z^{q^\ell - 1}}. \tag{9}$$

Since $r \geq 1$, $a(n)_{n\geq 0}$ is periodic by Lemma 22. Moreover, $\deg T < q^\ell - 1$, so its period length divides $q^\ell - 1$. Therefore $a(q^\ell n - j) = a((q^\ell n - j) \bmod (q^\ell - 1)) = a(n-j)$ for all $n \geq j$, so

$$\sum_{n\geq j} a(q^\ell n - j) z^n = \sum_{n\geq j} a(n-j) z^n = \frac{z^j}{R},$$

and it follows that

$$\Lambda_0^\ell\left(\frac{z^j}{R}\right) = \sum_{n=\lceil j/q^\ell \rceil}^{j-1} a(q^\ell n - j) z^n + \frac{z^j}{R}.$$

It remains to show that $a(q^\ell n - j) = 0$ for all $n \in \{\lceil j/q^\ell \rceil, \ldots, j-2, j-1\}$. If $j = 0$ or $j = 1$, this is vacuously true, so assume $j \in \{2, 3, \ldots, r\}$. We identify certain 0 coefficients in the series $\frac{1}{R}$. From Equation (9), we obtain

$$\sum_{n\geq 0} a(n) z^n = \frac{1}{R} = \frac{T}{1 - z^{q^\ell - 1}} = T + Tz^{q^\ell - 1} + Tz^{2(q^\ell - 1)} + \cdots.$$

Since $\deg T = q^\ell - 1 - \deg R = q^\ell - 1 - r$, this implies $0 = a(q^\ell - r) = a(q^\ell - r + 1) = \cdots = a(q^\ell - 2)$; that is, $a(q^\ell - i) = 0$ for all $i \in \{2, 3, \ldots, r\}$. For all $n \in \{\lceil j/q^\ell \rceil, \ldots, j-2, j-1\}$, we have $j - n + 1 \in \{2, 3, \ldots, j - \lceil j/q^\ell \rceil + 1\} \subseteq \{2, 3, \ldots, r\}$. Therefore, since the period length of $a(n)_{n\geq 0}$ divides $q^\ell - 1$, we have $a(q^\ell n - j) = a(q^\ell - (j - n + 1)) = 0$ for $n \in \{\lceil j/q^\ell \rceil, \ldots, j-2, j-1\}$, as desired. □

In Theorem 23 we assumed that $\deg R \geq 1$. However in general $\deg R \geq -1$; the next result extends Theorem 23.

Corollary 25. *Let $R \in z^{-1}\mathbb{F}_q[z]$ be a nonzero square-free Laurent polynomial whose factorization into irreducibles is of the form $cz^{e_0} R_1 \cdots R_k$, where $e_0 \in \{-1, 0\}$. Let $\ell = \operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$. Define λ_0 on $\mathbb{F}_q[z]$ by $\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big)$. Then $\lambda_0^\ell(S) = S$ for all $S \in \mathbb{F}_q[z]$ with $\deg S \leq \deg R$.*

Proof. Let $r := \deg R$. Theorem 23 covers the case $r \geq 1$. If $r = -1$, then $S = 0$ and the conclusion holds.

Suppose $r = 0$, so that $R = bz^{-1} + c$ for some $b, c \in \mathbb{F}_q$ with $c \neq 0$. Here $\ell = 1$. Let $S \in \mathbb{F}_q$. By Proposition 4, $\lambda_0(S) = \Lambda_0(SR^{q-1}) = S\Lambda_0(\frac{1}{R})R$. We show that $\Lambda_0(\frac{1}{R}) = \frac{1}{R}$, which will imply $\lambda_0(S) = S$. If $b = 0$, then

$$\Lambda_0\left(\frac{1}{R}\right) = \Lambda_0\left(\frac{1}{c}\right) = \frac{1}{c} = \frac{1}{R}.$$

If $b \neq 0$, then

$$\begin{aligned}\Lambda_0\left(\frac{1}{R}\right) &= \Lambda_0\left(\frac{z}{b(1 - (-c/b)z)}\right) = \Lambda_0\left(\frac{1}{b}\sum_{n\geq 0}(-c/b)^n z^{n+1}\right) \\ &= \frac{1}{b}\sum_{n\geq 1}(-c/b)^{nq-1} z^n = \frac{1}{b}\sum_{n\geq 0}(-c/b)^n z^{n+1} = \frac{1}{R}\end{aligned}$$

since $nq - 1 \equiv n - 1 \mod q - 1$. □

Proposition 13 tells us that the orbit of $y\frac{\partial P}{\partial y}$ under $\lambda_{0,0}$, when restricted to the left, right, and top borders of V, is dictated by the orbits of its projection onto these borders under λ_0, where the latter is defined using A_0/y, A_h/y, and B_d. These orbits can be studied using Corollary 25 as follows.

Example 26. We continue to use the polynomial P from Examples 7 and 12. The initial state is $S_0 = y\frac{\partial P}{\partial y}$, and the second state in the orbit of S_0 under $\lambda_{0,0}$ is

$$S_1 := \lambda_{0,0}(S_0) = xy^3 + (x^2 + x + 1)y^2 + (2x^2 + 2)y + x^2 + x.$$

In Example 14, we reduced to the univariate operator λ_0 with $R = A_0/y$, $R = A_h/y$, and $R = B_d$. We verify that the conditions of Corollary 25 hold. The factorizations into irreducibles of the three Laurent polynomials R are

$$\begin{aligned} A_0/y &= 2y^3 + y + 1 = 2(y^3 + 2y + 2) \\ A_h/y &= y^3 + 1 + 2y^{-1} = y^{-1}(y^4 + y + 2) \\ B_d &= x^2 + x + 2. \end{aligned}$$

Moreover, $\deg \pi_\ell(S_1) \le \deg(A_0/y)$, $\deg \pi_{\mathrm{r}}(S_1) \le \deg(A_h/y)$, and $\deg \pi_{\mathrm{t}}(S_1) \le \deg B_d$. Therefore, by Corollary 25, the three relevant orbits under the three operators λ_0 are periodic and have respective period lengths dividing 3, 4, and 2. For $R = A_0/y$, the orbit of $\pi_\ell(S_1)$ under λ_0 is

$$y^2 + 2y,\ y^2 + y,\ y^2,\ y^2 + 2y,\ \ldots.$$

For $R = A_h/y$, the orbit of $\pi_{\mathrm{r}}(S_1)$ under λ_0 is

$$y^2 + 2y + 1,\ y^2 + y + 1,\ y^2,\ 1,\ y^2 + 2y + 1,\ \ldots.$$

Lastly, for $R = B_d$, the orbit of $\pi_{\mathrm{t}}(S_1)$ under λ_0 is

$$x,\ 2x,\ x,\ \ldots.$$

In particular, the upper bounds on the period lengths are attained.

It remains to remove the restriction that R is square-free. Unlike the square-free case, the orbit of S under λ_0 may have a transient (in other words, may be eventually periodic but not periodic). First we give two propositions showing that elements sufficiently far out in the orbit are necessarily divisible by a certain polynomial; if S is not divisible by this polynomial then the orbit has a transient.

Proposition 27. *Let $R \in z^{-1}\mathbb{F}_q[z]$ be a nonzero Laurent polynomial such that $R = F^e G$ for some $F \in \mathbb{F}_q[z]$, $G \in z^{-1}\mathbb{F}_q[z]$, and $e \ge 1$. For all $S \in \mathbb{F}_q[z]$ and all $n \ge \lceil \log_q e \rceil$, the polynomial $\lambda_0^n(S)$ is divisible by F^{e-1}.*

Note that there are potentially multiple ways to decompose R in Proposition 27. For example, if $R = z^4$ then we could write $F = z, e = 4, G = 1$ or $F = z, e = 5, G = z^{-1}$. The latter choice leads to a stronger conclusion regarding divisibility.

Proof of Proposition 27. Let $S \in \mathbb{F}_q[z]$, and write $S = F^s T$ for some $s \ge 0$. (We do not require s to be maximal.) We have

$$\begin{aligned} \lambda_0(S) = \Lambda_0(SR^{q-1}) &= \Lambda_0\Big(F^s T F^{e(q-1)} G^{q-1}\Big) \\ &= \Lambda_0\Big(F^{(s+e(q-1)) \bmod q} T G^{q-1}\Big)\, F^{\lfloor (s+e(q-1))/q \rfloor} \end{aligned}$$

by Proposition 4. Therefore $\lambda_0(S)$ is divisible by $F^{\lfloor (s+e(q-1))/q \rfloor} = F^{e+\lfloor (s-e)/q \rfloor}$, so we iterate the function $f(x) = e + \left\lfloor \frac{x-e}{q} \right\rfloor$. Let $n \geq \lceil \log_q e \rceil$, so that $\left\lfloor -\frac{e}{q^n} \right\rfloor = -1$. As in the proof of Lemma 18, we have $f^n(s) = e + \left\lfloor \frac{s-e}{q^n} \right\rfloor \geq e + \left\lfloor -\frac{e}{q^n} \right\rfloor = e-1$. Therefore $\lambda_0^n(S)$ is divisible by F^{e-1}. □

Example 28. Let $q = 3$ and $R = z^{-1}(z+1)^3(z+2)$. The orbit of $S = 1$ under λ_0 is $1, (z+1)^2, (z+1)^2, \ldots$. It has transient length 1 (and period length 1).

If $F = z$ and if G is a polynomial, then we can slightly increase the exponent to which F eventually divides elements in the orbit under λ_0.

Proposition 29. *Let $R \in \mathbb{F}_q[z]$ be a nonzero polynomial such that $R = z^e G$ for some $G \in \mathbb{F}_q[z]$ where $e \geq 1$ and G is not divisible by z. For all $S \in \mathbb{F}_q[z]$ and all $n \geq \lfloor \log_q e \rfloor + 1$, the polynomial $\lambda_0^n(S)$ is divisible by z^e.*

Proof. Let $S \in \mathbb{F}_q[z]$, and write $S = z^s T$ where $s \geq 0$ and T is not divisible by z. We have

$$\lambda_0(S) = \Lambda_0(SR^{q-1}) = \Lambda_0\Big(z^{s+e(q-1)} T G^{q-1}\Big).$$

Since $z^{s+e(q-1)} T G^{q-1}$ is divisible by $z^{s+e(q-1)}$, it follows that $\lambda_0(S)$ is divisible by $z^{f(s)}$, where $f(x) = e + \left\lceil \frac{x-e}{q} \right\rceil$. Applying Lemma 18, if $n \geq \lfloor \log_q e \rfloor + 1$ then $\lambda_0^n(S)$ is divisible by z^e. □

In the following theorem, we show that the situation for a general (not necessarily square-free) element $R \in z^{-1}\mathbb{F}_q[z]$ reduces to Corollary 25 by Propositions 27 and 29. The idea of the proof is that if R is divisible by F^e, then every application of λ_0 pushes the image into a smaller vector space until we are emulating the map λ_0 for a square-free polynomial R'. We define $\log_q 0 = -\infty$, $\lfloor -\infty \rfloor = -\infty$, $\lceil -\infty \rceil = -\infty$, and $\max() = 0$. (When $\deg R = -1$, the only polynomial S satisfying $\deg S \leq \deg R$ is $S = 0$, so the theorem does not say much in this case.)

Theorem 30. *Let $R \in z^{-1}\mathbb{F}_q[z]$ be a nonzero Laurent polynomial. Let $R = cz^{e_0} R_1^{e_1} \cdots R_k^{e_k}$ be its factorization into irreducibles. Let*

$$t = \max\Big(\big\lfloor \log_q \max(e_0, 0) \big\rfloor + 1, \big\lceil \log_q \max(e_1, \ldots, e_k) \big\rceil, 0\Big) \tag{10}$$

and $\ell = \operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$. Define λ_0 on $\mathbb{F}_q[z]$ by $\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big)$. For all $S \in \mathbb{F}_q[z]$ with $\deg S \leq \deg R$, the orbit size of S under λ_0 is at most $t + \ell$.

Proof. Define the *radical* of R by $\operatorname{rad} R = cz^{\min(e_0,0)} R_1 \cdots R_k$. Let

$$U = z^{\max(e_0,0)} R_1^{e_1-1} \cdots R_k^{e_k-1}$$

so that $U \operatorname{rad} R = R$. Let $S \in \mathbb{F}_q[z]$ with $\deg S \leq \deg R$. We show that the orbit of S under λ_0 is eventually periodic with transient length at most t and period length dividing ℓ.

We claim that $\lambda_0^t(S) = TU$ for some $T \in \mathbb{F}_q[z]$ satisfying $\deg T \leq \deg \operatorname{rad} R$. To see that $R_i^{e_i-1}$ divides $\lambda_0^t(S)$ for each $i \in \{1, 2, \ldots, k\}$, we apply Proposition 27 with $F = R_i$. If $e_0 \in \{-1, 0\}$, then $z^{\max(e_0,0)} = 1$. If $e_0 \geq 1$, then Proposition 29

implies that $z^{\max(e_0,0)} = z^{e_0}$ divides $\lambda_0^t(S)$. To see that $\deg T \le \deg \operatorname{rad} R$, we have

$$\begin{aligned}\deg T &= \deg \lambda_0^t(S) - \deg U \\ &\le \deg R - \deg U \\ &= \left(e_0 + \sum_{i=1}^{k} e_i \deg R_i\right) - \left(\max(e_0,0) + \sum_{i=1}^{k} (e_i - 1)\deg R_i\right) \\ &= \min(e_0,0) + \sum_{i=1}^{k} \deg R_i = \deg \operatorname{rad} R.\end{aligned}$$

This completes the proof of the claim.

Next we use the identity $e_i - 1 + e_i(q-1) = e_i q - 1 = q - 1 + (e_i - 1)q$. For all $T \in \mathbb{F}_q[z]$ (and in particular for the T satisfying $\lambda_0^t(S) = TU$),

$$\begin{aligned}\lambda_0(TU) = \Lambda_0\big(TUR^{q-1}\big) &= \Lambda_0\Big(Tc^{q-1} z^{\max(e_0,0)+e_0(q-1)} R_1^{e_1 q-1} \cdots R_k^{e_k q-1}\Big) \\ &= \Lambda_0\Big(Tc^{q-1} z^{\min(e_0,0)(q-1)} R_1^{q-1} \cdots R_k^{q-1} U^q\Big) \\ &= \Lambda_0\big(T\,(\operatorname{rad} R)^{q-1}\big)\,U\end{aligned}$$

by Proposition 4. Accordingly, define $\kappa_0 \colon \mathbb{F}_q[z] \to \mathbb{F}_q[z]$ by $\kappa_0(T) = \Lambda_0(T\,(\operatorname{rad} R)^{q-1})$, so that $\lambda_0(TU) = \kappa_0(T)U$. Iterating, we have $\lambda_0^\ell(TU) = \kappa_0^\ell(T)U$. Applying Corollary 25 to κ_0, we have $\kappa_0^\ell(T) = T$ since $\operatorname{rad} R$ is square-free and $\deg T \le \deg \operatorname{rad} R$. Therefore

$$\lambda_0^{t+\ell}(S) = \lambda_0^\ell\big(\lambda_0^t(S)\big) = \lambda_0^\ell(TU) = \kappa_0^\ell(T)U = TU = \lambda_0^t(S),$$

so the orbit of S under λ_0 contains at most $t + \ell$ elements. □

6. Orbit size under $\lambda_{0,0}$

In this section, we prove Theorem 1. Our aim is to bound the size of the q-kernel for $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!]$, which satisfies $P(x,F) = 0$. From Corollary 11, it remains to bound $|\operatorname{orb}_{\Lambda_0}(F)|$, equivalently, $|\operatorname{orb}_{\lambda_{0,0}}(S_0)|$ where $S_0 = y\frac{\partial P}{\partial y}$.

To do this, we will use Theorem 30 to obtain bounds on orbit sizes under $\lambda_{0,0}$. We need the following lemma, which bounds the degree of the three border polynomials of $\lambda_{0,0}\Big(y\frac{\partial P}{\partial y}\Big)$. Recall the definitions of A_i and B_j from Equation (7), that $\deg(A_0/y) \ge 0$, and that A_h and B_d are nonzero.

Lemma 31. *Let $S_0 = y\frac{\partial P}{\partial y}$. Then*

1. $\deg \pi_\ell(\lambda_{0,0}(S_0)) \le \deg(A_0/y)$,
2. $\deg \pi_{\mathrm{r}}(\lambda_{0,0}(S_0)) \le \deg(A_h/y)$, *and*
3. $\deg \pi_{\mathrm{t}}(\lambda_{0,0}^n(S_0)) \le \deg B_d$ *for all* $n \ge \lfloor \log_q h \rfloor + 2$.

Proof. For the first two statements, we will use

$$\pi_\ell(S_0) = y\tfrac{dA_0}{dy} \quad \text{and} \quad \pi_{\mathrm{r}}(S_0) = y\tfrac{dA_h}{dy}.$$

For the first statement, let $R = A_0/y$. Part 1 of Proposition 13 gives $\pi_\ell(\lambda_{0,0}(S_0)) = \lambda_0(\pi_\ell(S_0)) = \Lambda_0\Big(y\frac{dA_0}{dy}\cdot(A_0/y)^{q-1}\Big)$. The degree of this polynomial is at most $\deg(A_0/y)$.

The second statement follows in the same way by applying Part 2 of Proposition 13 since $\deg_x S_0 \le h$.

For the first two statements, we applied Proposition 13 to $S_0 = y\frac{\partial P}{\partial y}$. For the third statement, we apply Part 3 of Proposition 13 to $R = B_d$ and $\lambda_{0,0}(S_0)$ since $\deg_y \lambda_{0,0}(S_0) \le d-1$ by Proposition 9. Let $S = \pi_{\mathrm{t}}(\lambda_{0,0}(S_0)) = \pi_{\mathrm{t}}\big(\Lambda_{0,0}\big(S_0 Q^{q-1}\big)\big)$. All nonzero monomials in $\pi_{\mathrm{t}}\big(\Lambda_{0,0}\big(S_0 Q^{q-1}\big)\big)$ come from applying Λ_0 to terms in $(d-1)B_{d-1}\cdot B_d^{q-1}$ or $dB_d\cdot(q-1)B_{d-1}B_d^{q-2}$, so $S = \Lambda_0\Big((qd-1)B_{d-1}B_d^{q-1}\Big)$. We have $r := \deg R = \deg B_d \ge 0$ and $s := \deg S \le h$. By Proposition 19, if $n-1 \ge \big\lfloor \log_q h \big\rfloor + 1 \ge \big\lfloor \log_q \max(s-r,1) \big\rfloor + 1$ then $\deg \pi_{\mathrm{t}}(\lambda_{0,0}^n(S_0)) \le \deg B_d$. $\square$

Example 32. For the polynomial P in Examples 7 and 26, it suffices to take $n = 1$ to achieve $\deg \pi_{\mathrm{t}}(\lambda_{0,0}^n(S_0)) \le \deg B_d$, since

$$\deg \pi_{\mathrm{t}}(\lambda_{0,0}(S_0)) = \deg \pi_{\mathrm{t}}(S_1) = \deg x \le \deg(x^2+x+2) = \deg B_d.$$

The eventual period lengths in Theorem 30 are bounded by $\operatorname{lcm}(\deg R_1, \dots, \deg R_k)$. We will use the function $\mathcal{L}(l,m,n)$ defined in Section 1 to obtain a bound that is independent of the factorizations of A_0/y, A_h/y, and B_d. We rephrase Theorem 1 in terms of $\ker_q(a(n)_{n\ge 0})$, since it has the same size as the minimal automaton for $a(n)_{n\ge 0}$.

Theorem 1. *Let $F = \sum_{n\ge 0} a(n)x^n \in \mathbb{F}_q[\![x]\!] \setminus \{0\}$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x,y]$ of height h and degree d. Then*

$$|\ker_q(a(n)_{n\ge 0})| \le q^{hd} + q^{(h-1)(d-1)}\mathcal{L}(h,d,d) + \big\lfloor \log_q h \big\rfloor + \big\lceil \log_q \max(h,d-1) \big\rceil + 3.$$

Proof. By Corollary 11, $|\ker_q(a(n)_{n\ge 0})| \le q^{hd} + |\mathrm{orb}_{\Lambda_0}(F)|$, so we now bound $|\mathrm{orb}_{\Lambda_0}(F)| \le |\mathrm{orb}_{\lambda_{0,0}}(S_0)|$. We do this by emulating $\lambda_{0,0}$ with the appropriate univariate operators λ_0 on the left, right, and top borders of V and using a crude upper bound for the rest. Lemma 31 and Proposition 13 will allow us to do this.

We use the following fact. Let V be a finite vector space with basis $\mathcal{B}$. Let $(\mathcal{B}_1, \mathcal{B}_2)$ be a partition of $\mathcal{B}$, and let U_1 and U_2 be the subspaces generated by $\mathcal{B}_1$ and $\mathcal{B}_2$. Let π_U denote projection onto U. If $f\colon V \to V$ and $\tilde{f}\colon U_1 \to U_1$ are linear transformations satisfying $\pi_{U_1} \circ f = \tilde{f} \circ \pi_{U_1}$, then

$$f(x) = \pi_{U_1}(f(x)) + \pi_{U_2}(f(x)) = \tilde{f}(\pi_{U_1}(x)) + \pi_{U_2}(f(x)),$$

so that $|\mathrm{orb}_f(x)| \le |U_2| \cdot |\mathrm{orb}_{\tilde{f}}(\pi_{U_1}(x))|$ for all $x \in V$.

We apply this fact to $U_2 = V^\circ$, where V° is defined in Equation (5), and $f = \lambda_{0,0}$. To define $\tilde{f}$, note that Proposition 13 gives us an operator $\tilde{\lambda}_0\colon U_1 \to U_1$ that satisfies $\pi_{U_1} \circ \lambda_{0,0} = \tilde{\lambda}_0 \circ \pi_{U_1}$. (The operator $\tilde{\lambda}_0$ acts as the appropriate λ_0 on the three respective borders.) Set $\tilde{f} = \tilde{\lambda}_0$. The fact in the previous paragraph now implies $|\mathrm{orb}_{\lambda_{0,0}}(x)| \le |V^\circ| \cdot |\mathrm{orb}_{\tilde{\lambda}_0}(\pi_{U_1}(x))|$ for all $x \in V$. Let

$$t := \big\lfloor \log_q h \big\rfloor + 2 + \big\lceil \log_q \max(h, d-1) \big\rceil + 1;$$

we will justify the definition of t below. Set $S_t := \lambda_{0,0}^t(y\frac{\partial P}{\partial y})$. We have $S_t \in V$ by Proposition 9 since $t \ge 1$. Write $S_t = \pi_{V^\circ}(S_t) + T$ where $T \in U_1$. Therefore, using $|V^\circ| = q^{(h-1)(d-1)}$, we have

$$|\mathrm{orb}_{\lambda_{0,0}}(S_t)| \le q^{(h-1)(d-1)} \cdot |\mathrm{orb}_{\tilde{f}}(T)|. \tag{11}$$

It remains to bound $|\mathrm{orb}_{\tilde{f}}(T)|$. We will do this by bounding the orbit sizes of the projections $\pi_\ell(T)$, $\pi_{\mathrm{r}}(T)$, and $\pi_{\mathrm{t}}(T)$ under the respective operators λ_0, defined by the Laurent polynomials $R = A_0/y$ on the left border, $R = A_h/y$ on the right

border, and $R = B_d$ on the top border. Equation (10) in Theorem 30 will give transient lengths t_ℓ, t_r, and t_t in terms of the factorization of R. These transient lengths are at most

$$\max(t_\ell, t_\mathrm{r}, t_\mathrm{t}) \leq \left\lceil \log_q \max(h, d-1) \right\rceil + 1,$$

where the upper bound here is achieved in the extreme case $e_1 = \cdots = e_k = 0$ and $e_0 = h$ or $e_0 = d-1$. Since $t \geq \left\lfloor \log_q h \right\rfloor + 2$, Lemma 31 and Proposition 13 tell us that, on U_1, we can emulate the action of $\lambda_{0,0}$ on S_t with the three operators λ_0. Since $\pi_\ell(T) = \pi_\ell(S_t)$, $\pi_\mathrm{r}(T) = \pi_\mathrm{r}(S_t)$, and $\pi_\mathrm{t}(T) = \pi_\mathrm{t}(S_t)$, we consider the sizes of the orbits

$$\begin{aligned}
\mathrm{orb}_\ell(S_t) &= \{\lambda_0^n(\pi_\ell(S_t)) : n \geq 0\} \\
\mathrm{orb}_\mathrm{r}(S_t) &= \{\lambda_0^n(\pi_\mathrm{r}(S_t)) : n \geq 0\} \\
\mathrm{orb}_\mathrm{t}(S_t) &= \{\lambda_0^n(\pi_\mathrm{t}(S_t)) : n \geq 0\}.
\end{aligned}$$

By Theorem 30 and our choice of t, these three orbits are periodic, i.e. have no transient. Lemma 31 implies $\deg S_t \leq \deg B_d$, so we can apply Theorem 30 with $R = B_d$ to $\pi_\mathrm{t}(S_t)$. It tells us that $|\mathrm{orb}_\mathrm{t}(S_t)| = \mathrm{lcm}(\sigma)$ for some integer partition $\sigma \in \mathrm{parts}(\deg B_d)$. Similarly, for $|\mathrm{orb}_\ell(S_t)|$ and $|\mathrm{orb}_\mathrm{r}(S_t)|$ we obtain integer partitions in $\mathrm{parts}(1 + \deg A_0/y) = \mathrm{parts}(\deg A_0)$ and $\mathrm{parts}(1 + \deg A_h/y) = \mathrm{parts}(\deg A_h)$.

We now use

$$|\mathrm{orb}_{\tilde{f}}(T)| \leq \mathrm{lcm}\big(|\mathrm{orb}_\ell(S_t)|, |\mathrm{orb}_\mathrm{r}(S_t)|, |\mathrm{orb}_\mathrm{t}(S_t)|\big) \tag{12}$$

and maximize over the orbit sizes that arise. By Equation (12) and the definition of $\mathcal{L}$, we have $|\mathrm{orb}_{\tilde{f}}(T)| \leq \mathcal{L}(h, d, d)$ since $\deg A_0 \leq d$, $\deg A_h \leq d$, and $\deg B_d \leq h$. Equation (11) gives

$$|\mathrm{orb}_{\lambda_{0,0}}(S_t)| \leq q^{(h-1)(d-1)}\mathcal{L}(h, d, d).$$

It follows that $|\mathrm{orb}_{\Lambda_0}(F)| \leq |\mathrm{orb}_{\lambda_{0,0}}(S_t)| + t \leq q^{(h-1)(d-1)}\mathcal{L}(h, d, d) + t$ as desired. □

Example 33. We continue Examples 7 and 32, where $h = 2$ and $d = 4$. We have $\mathcal{L}(h, d, d) = 12 = \mathrm{lcm}(3, 4, 2)$. Computing the orbit of S_0 under $\lambda_{0,0}$, one finds that it has size 157, consisting of 1 transient state followed by a period with length $156 = 13 \cdot 12$. This period length is less than the theoretical maximum $q^{(h-1)(d-1)}\mathcal{L}(h, d, d) = 27 \cdot 12$. The number of states in the constructed automaton is $5989 \approx 3^{7.917}$, which is on the order of the upper bound

$$\begin{aligned}
&q^{hd} + q^{(h-1)(d-1)}\mathcal{L}(h, d, d) + \left\lfloor \log_q h \right\rfloor + \left\lceil \log_q \max(h, d-1) \right\rceil + 3 \\
&\quad = 3^8 + 3^3 \cdot 12 + 4 \\
&\quad = 6889 \approx 3^{8.044}.
\end{aligned}$$

Minimizing the automaton reduces the number of states by 1 to 5988.

Asymptotically, we have the following.

Theorem 2. *Let $F = \sum_{n \geq 0} a(n)x^n \in \mathbb{F}_q[\![x]\!]$ be the Furstenberg series associated with a polynomial $P \in \mathbb{F}_q[x, y]$ of height h and degree d. Then $|\ker_q(a(n)_{n \geq 0})|$ is in $(1 + o(1))q^{hd}$ as any of q, h, or d tends to infinity and the others remain constant.*

Proof. Recall that the conditions on a Furstenberg series guarantee that $d \geq 1$. If $h = 0$, then the power series F is the 0 series, so $|\ker_q(a(n)_{n\geq 0})| = 1$. Therefore we assume $h \geq 1$.

As before, let $g(n)$ be the Landau function. The set of triples of integer partitions of h, d, d gives rise to a subset of integer partitions of $h + 2d$. Thus $\mathcal{L}(h, d, d) \leq g(h + 2d)$. By Theorem 1,

$$|\ker_q(a(n)_{n\geq 0})| \leq q^{hd} + q^{(h-1)(d-1)}g(h + 2d) + \lfloor \log_q h \rfloor + \lceil \log_q \max(h, d - 1) \rceil + 3.$$

The expression $\lfloor \log_q h \rfloor + \lceil \log_q \max(h, d - 1) \rceil + 3$ is clearly in $o(1)q^{hd}$. It remains to show that $q^{(h-1)(d-1)}g(h + 2d)$ is also in $o(1)q^{hd}$. Landau [17] proved that $\log g(n) \sim \sqrt{n \log n}$, that is, $g(n) = e^{(1+\epsilon(n))\sqrt{n \log n}}$, where $\epsilon(n) \to 0$ as $n \to \infty$. Therefore

$$\frac{q^{(h-1)(d-1)}g(h + 2d)}{q^{hd}} = \frac{g(h + 2d)}{q^{h+d-1}} = \frac{e^{(1+\epsilon(h+2d))\sqrt{(h+2d)\log(h+2d)}}}{q^{h+d-1}},$$

and this tends to 0 as any of q, h, or d tends to infinity and the others remain constant. □

Bridy used a similar argument, also bounding the orbit size by $g(h + 2d)$ [7, Proof of Theorem 1.2].

Example 34. The factor $1+o(1)$ cannot be removed from the bound in Theorem 2. Let $q = 2$, and consider

$$P = (x^3 + x^2 + 1)y^3 + (x^3 + 1)y^2 + (x^3 + x^2 + x + 1)y + x^3 + x^2 \in \mathbb{F}_2[x, y]$$

with height $h = 3$ and degree $d = 3$. The coefficient sequence $a(n)_{n\geq 0}$ of the series $F \in \mathbb{F}_2[\![x]\!]$ satisfying $P(x, F) = 0$ is $0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, \ldots$. The constructed automaton has 532 states. Minimizing reduces the number of states by only 1 to 531, which is larger than $q^{hd} = 512$.

With the same techniques as in the proof of Theorem 1, one obtains the following result, which concerns diagonals of rational functions that are not necessarily of the form in Theorem 5. To state it, we extend the function $\mathcal{L}(n_1, n_2, n_3)$ from Section 1 to $\mathcal{L}(n_1, n_2, n_3, n_4)$, defined analogously as the maximum value of $\operatorname{lcm}(\operatorname{lcm}(\sigma_1), \operatorname{lcm}(\sigma_2), \operatorname{lcm}(\sigma_3), \operatorname{lcm}(\sigma_4))$ over integer partitions σ_i of integers in $\{1, 2, \ldots, n_i\}$. The reason for this is that Theorem 35 is symmetric in x and y, unlike Theorem 5. This symmetry leads to the appearance of $\mathcal{L}(h, h, d, d)$ in Theorem 35 instead of $\mathcal{L}(h, d, d)$ as in Theorem 1.

Theorem 35. *Let $P(x, y)$ and $Q(x, y)$ be polynomials in $\mathbb{F}_q[x, y]$ such that $Q(0, 0) \neq 0$. Let*

$$F = \mathcal{D}\left(\frac{P(x, y)}{Q(x, y)}\right),$$

and write $F(x) = \sum_{n\geq 0} a(n)x^n$. Let

$$h = \max(\deg_x P, \deg_x Q)$$
$$d = \max(\deg_y P, \deg_y Q),$$

and assume $h \geq 1$ and $d \geq 1$. Then the size of $\ker_q((a(n) \bmod p^\alpha)_{n\geq 0})$ is at most

$$q^{hd} + q^{(h-1)(d-1)}\mathcal{L}(h, h, d, d) + \lfloor \log_q \max(h, d) \rfloor + \lceil \log_q \max(h, d) \rceil + 2.$$

Consequently, $|\ker_q((a(n))_{n\geq 0})|$ *is in* $(1+o(1))q^{hd}$ *as any of* q*,* h*, or* d *tends to infinity and the others remain constant.*

When comparing Theorem 35 to the rest of the paper, note that a Furstenberg series is the diagonal of a rational function, whose denominator is also called Q, but in Theorem 35 the denominator Q has degree d and not $d-1$ as before.

The structure of the proof of Theorem 35 is similar to that of Theorem 1. One difference is that the diagonal in Theorem 5 contains expressions of the form $P(xy,y)$, which led us to shear and to consider the maps $\lambda_{r,0}$ on Laurent polynomials. For general diagonals, the symmetry in x and y means that no shearing is required, the relevant maps are $\lambda_{r,r}$, and no Laurent polynomials enter the picture. We define the main objects and state the modifications of relevant results used in the proof. With h and d defined as in Theorem 35, let

$$V := \left\langle x^i y^j : 0 \leq i \leq h \text{ and } 0 \leq j \leq d \right\rangle$$

and

$$V^\circ = \left\langle x^i y^j : 1 \leq i \leq h-1 \text{ and } 1 \leq j \leq d-1 \right\rangle.$$

The initial state of the automaton is $S_0 = P$. Define π_ℓ and π_{r} as in Section 4. For a polynomial $S = \sum_{i,j} c_{i,j} x^i y^j$, define $\pi_{\mathrm{b}}(S) = \sum_i c_{i,0} x^i$ and $\pi_{\mathrm{t}}(S) = \sum_i c_{i,d} x^i$. The following results are analogues of Proposition 13 and Lemma 31; their proofs are similar.

Proposition 36. *We have the following.*

(1) *Let* $R = \pi_\ell(Q)$*. For all* $S \in \mathbb{F}_q[x,y]$*,*
$$\pi_\ell(\lambda_{0,0}(S)) = \lambda_0(\pi_\ell(S)).$$
(2) *Let* $R = \pi_{\mathrm{r}}(Q)$*. For all* $S \in \mathbb{F}_q[x,y]$ *with height at most* h*,*
$$\pi_{\mathrm{r}}(\lambda_{0,0}(S)) = \lambda_0(\pi_{\mathrm{r}}(S)).$$
(3) *Let* $R = \pi_{\mathrm{b}}(Q)$*. For all* $S \in \mathbb{F}_q[x,y]$*,*
$$\pi_{\mathrm{b}}(\lambda_{0,0}(S)) = \lambda_0(\pi_{\mathrm{b}}(S)).$$
(4) *Let* $R = \pi_{\mathrm{t}}(Q)$*. For all* $S \in \mathbb{F}_q[x,y]$ *with degree at most* d*,*
$$\pi_{\mathrm{t}}(\lambda_{0,0}(S)) = \lambda_0(\pi_{\mathrm{t}}(S)).$$

Lemma 37. *Let*

$$\begin{aligned} u_\ell &= \left\lfloor \log_q \max((d - \deg \pi_\ell(Q)), 1) \right\rfloor + 1 \\ u_{\mathrm{r}} &= \left\lfloor \log_q \max((d - \deg \pi_{\mathrm{r}}(Q)), 1) \right\rfloor + 1 \\ u_{\mathrm{b}} &= \left\lfloor \log_q \max((h - \deg \pi_{\mathrm{b}}(Q)), 1) \right\rfloor + 1 \\ u_{\mathrm{t}} &= \left\lfloor \log_q \max((h - \deg \pi_{\mathrm{t}}(Q)), 1) \right\rfloor + 1. \end{aligned}$$

For all $S \in V$*, we have*

(1) $\deg \pi_\ell(\lambda_{0,0}^n(S)) \leq \deg \pi_\ell(Q)$ *for all* $n \geq u_\ell$*,*
(2) $\deg \pi_{\mathrm{r}}(\lambda_{0,0}^n(S)) \leq \deg \pi_{\mathrm{r}}(Q)$ *for all* $n \geq u_{\mathrm{r}}$*,*
(3) $\deg \pi_{\mathrm{b}}(\lambda_{0,0}^n(S)) \leq \deg \pi_{\mathrm{b}}(Q)$ *for all* $n \geq u_{\mathrm{b}}$*, and*
(4) $\deg \pi_{\mathrm{t}}(\lambda_{0,0}^n(S)) \leq \deg \pi_{\mathrm{t}}(Q)$ *for all* $n \geq u_{\mathrm{t}}$*.*

With Proposition 36 and Lemma 37, one follows the proof of Theorem 1 to prove Theorem 35. The term $\left\lfloor \log_q \max(h,d) \right\rfloor + 1$ in Theorem 35 comes from Lemma 37 by bounding u_ℓ, u_{r} by $\left\lfloor \log_q d \right\rfloor + 1$ and $u_{\mathrm{b}}, u_{\mathrm{t}}$ by $\left\lfloor \log_q h \right\rfloor + 1$.

7. Subspaces of univariate polynomials

In this section, we give two conjectures that were discovered in earlier attempts to prove results in Section 5 bounding the period length of an orbit under the linear operator λ_0. They were not needed in the end, but they are interesting in their own right since they identify additional structure in λ_0. For a polynomial $R \in \mathbb{F}_q[z]$, define $\lambda_0(S) = \Lambda_0\big(SR^{q-1}\big)$ as in Equation (6).

The first conjecture implies the conclusion of Theorem 23, given in Proposition 39 below. For an integer $m \geq 1$, consider the set of polynomials $S \in \mathbb{F}_q[z]$ such that $\deg S \leq \deg R$ and $\lambda_0^m(S) = S$. This set forms a vector space.

Conjecture 38. *Let $R \in \mathbb{F}_q[z]$ such that $\deg R \geq 1$ and R is not divisible by z. Let $R = cR_1^{e_1} \cdots R_k^{e_k}$ be its factorization into irreducibles. For every divisor m of $\operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$, the vector space*

$$\{S \in \mathbb{F}_q[z] : \deg S \leq \deg R \text{ and } \lambda_0^m(S) = S\} \tag{13}$$

has dimension $1 + \sum_{i=1}^{k} \gcd(m, \deg R_i)$.

In particular, the exponents e_i do not affect the dimension.

Proposition 39. *Let $R \in \mathbb{F}_q[z]$ be a nonzero square-free polynomial such that $\deg R \geq 1$ and R is not divisible by z. Let $R = cR_1 \cdots R_k$ be its factorization into irreducibles, and let $\ell = \operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$. Conjecture 38 implies that $\lambda_0^\ell(S) = S$ for all $S \in \mathbb{F}_q[z]$ with $\deg S \leq \deg R$.*

Proof. For $m = \ell$, Conjecture 38 states that the vector space (13) has dimension

$$1 + \sum_{i=1}^{k} \gcd(\ell, \deg R_i) = 1 + \sum_{i=1}^{k} \deg R_i = 1 + \deg R,$$

so in fact it is the entire space $\{S \in \mathbb{F}_q[z] : \deg S \leq \deg R\}$. □

A natural question is whether we can write down an explicit basis of the vector space (13). For $m = 1$, Conjecture 38 implies that the subspace of fixed points has dimension $k + 1$. The next conjecture provides a basis of this subspace, for certain polynomials R. One basis element is R itself, since $\lambda_0(R) = \Lambda_0(R^q) = R$ by Proposition 4. We get k additional basis elements from the following operation. For a polynomial $S = \sum_{j=0}^{s} c_j z^j \in \mathbb{F}_q[z]$ where $c_s \neq 0$, define $\Delta(S) = \sum_{j=0}^{s} (s-j) c_j z^j$. Equivalently, $\Delta(S) = z^{s-1} \frac{\mathrm{d}}{\mathrm{d}w}(w^s S)$ where $w = \frac{1}{z}$. From this it follows that Δ is a derivation. That is, $\Delta(ST) = \Delta(S)T + S\Delta(T)$ for all $S, T \in \mathbb{F}_q[z]$.

Conjecture 40. *Let $R \in \mathbb{F}_q[z]$ such that $\deg R \geq 1$. Let $R = cR_1^{e_1} \cdots R_k^{e_k}$ be its factorization into irreducibles. For each $i \in \{1, 2, \ldots, k\}$, the polynomial $R_1^{e_1} \cdots R_{i-1}^{e_{i-1}} \Delta(R_i^{e_i}) R_{i+1}^{e_{i+1}} \cdots R_k^{e_k}$ is a fixed point of λ_0. Moreover, if R is not divisible by z and $e_i \not\equiv 0 \mod p$ for all i, where p is the characteristic of $\mathbb{F}_q$, then these k fixed points, along with R, are linearly independent.*

If R is divisible by z, then we don't get a basis element because $\Delta(z) = 0$. Similarly, if $e_i \equiv 0 \mod p$ then $\Delta(R_i^{e_i}) = e_i R_i^{e_i - 1} \Delta(R_i) = 0$.

Conjecture 40 implies that $\Delta(R)$ is a fixed point of λ_0, since we can use the fact that Δ is a derivation to write $\Delta(R)$ as a sum of fixed points.

For $m \geq 2$, it would be interesting to know how to extend the basis of fixed points in Conjecture 40 to a basis of the vector space (13).

Acknowledgment

The third author thanks Boris Adamczewski for several helpful discussions.

Appendix

The next few pages contain the tables and figures mentioned in Section 2.

$q = 2$:

h	d	P	aut. size	q^{hd}	bound
1	1	$y + x$	3	2	6
2	1	$(x^2 + x + 1)y + x^2$	6	4	11
3	1	$(x^3 + x + 1)y + x^3$	11	8	17
4	1	$(x^4 + x + 1)y + x^4$	20	16	27
5	1	$(x^5 + x^3 + 1)y + x^5$	37	32	46
6	1	$(x^6 + x + 1)y + x^6$	70	64	78
7	1	$(x^7 + x + 1)y + x^7$	135	128	148
8	1	$(x^8 + x^7 + x^2 + x + 1)y + x^8$	264	256	280
9	1	$(x^9 + x^5 + 1)y + x^9$	521	512	542
10	1	$(x^{10} + x^3 + 1)y + x^{10}$	1034	1024	1064
1	2	$xy^2 + (x + 1)y + x$	7	4	9
2	2	$x^2y^2 + (x^2 + x + 1)y + x^2$	14	16	25
3	2	$(x^3 + x^2 + 1)y^2 + (x^3 + 1)y + x$	68	64	94
4	2	$(x^4 + x + 1)y^2 + (x^4 + x^2 + x + 1)y + x$	252	256	311
5	2	$(x^5 + x^3 + 1)y^2 + (x^5 + x + 1)y + x$	1052	1024	1192
6	2	$(x^6 + x^5 + 1)y^2 + (x^6 + x^2 + x + 1)y + x$	4062	4096	4424
7	2	$(x^7 + x + 1)y^2 + (x^7 + x^4 + x^3 + x + 1)y + x$	16424	16384	17288
1	3	$xy^3 + y^2 + (x + 1)y + x$	11	8	18
2	3	$(x^2 + x + 1)y^3 + y^2 + (x^2 + 1)y + x^2 + x$	61	64	93
3	3	$(x^3 + x + 1)y^3 + y^2 + (x^3 + x^2 + x + 1)y + x^3 + x^2$	533	512	614
4	3	$(x^4 + x + 1)y^3 + y^2 + (x^4 + 1)y + x^4 + x^3 + x$	4213	4096	4871
1	4	$(x + 1)y^4 + y^2 + (x + 1)y + x$	20	16	33
2	4	$(x^2 + x + 1)y^4 + y^3 + (x^2 + x + 1)y + x^2 + x$	216	256	358
3	4	$(x^3 + x + 1)y^4 + y^3 + (x^3 + 1)y + x^2 + x$	3956	4096	4870
1	5	$(x + 1)y^5 + (x + 1)y^2 + y + x$	37	32	67
2	5	$(x^2 + x + 1)y^5 + y^4 + y^3 + x^2y^2 + y + x^2 + x$	889	1024	1510
3	5	$(x^3 + x^2 + 1)y^5 + y^4 + x^3y^2 + (x + 1)y + x^3 + x^2 + x$	43913	32768	48134

$q = 3$:

h	d	P	aut. size	q^{hd}	bound
1	1	$(x + 1)y + x$	4	3	7
2	1	$(2x^2 + x + 1)y + x^2$	11	9	15
3	1	$(x^3 + 2x + 1)y + x^3$	30	27	35
4	1	$(2x^4 + x + 1)y + x^4$	85	81	91
5	1	$(x^5 + 2x + 1)y + x^5$	248	243	255
6	1	$(2x^6 + x + 1)y + x^6$	735	729	741
1	2	$(x + 1)y^2 + y + x$	9	9	14
2	2	$(x^2 + x + 2)y^2 + y + x^2$	79	81	91
3	2	$(x^3 + x^2 + 2x + 1)y^2 + y + x^3 + x$	727	729	788
4	2	$(x^4 + x^3 + 2)y^2 + y + x^4 + x$	6533	6561	6729

TABLE 1. Polynomials in $\mathbb{F}_q[x, y]$ achieving the maximum unminimized automaton size for given values of q, h, and d, for comparison with the bound in Theorem 1.

$q = 2$:

h	d	P	orbit size	bound
1	1	$y + x$	2	4
2	1	$y + x^2$	3	7
3	1	$(x^3 + x + 1)y + x$	4	9
4	1	$(x^4 + x^3 + 1)y + x$	5	11
5	1	$(x^5 + x + 1)y + x$	7	14
6	1	$(x^6 + x^3 + 1)y + x$	7	14
7	1	$(x^7 + x^6 + x^2 + x + 1)y + x$	13	20
8	1	$(x^8 + x^3 + 1)y + x$	16	24
9	1	$(x^9 + x^2 + 1)y + x$	21	30
10	1	$(x^{10} + x^6 + x^3 + x^2 + 1)y + x$	31	40
1	2	$xy^2 + (x + 1)y + x$	3	5
2	2	$x^2y^2 + (x^2 + x + 1)y + x^2$	6	9
3	2	$(x^3 + x^2 + 1)y^2 + (x^3 + 1)y + x$	12	30
4	2	$(x^4 + x^2 + x)y^2 + (x^4 + x + 1)y + x^4$	25	55
5	2	$(x^5 + x^3 + 1)y^2 + (x^5 + x + 1)y + x$	60	168
6	2	$(x^6 + x^4 + x)y^2 + (x^5 + x + 1)y + x$	61	328
7	2	$(x^7 + x + 1)y^2 + (x^7 + x^4 + x^3 + x + 1)y + x$	168	904
8	2	$(x^8 + x^3 + 1)y^2 + (x^8 + x^7 + x^2 + x + 1)y + x^8$	240	3849
1	3	$(x + 1)y^3 + y^2 + (x + 1)y + x$	7	10
2	3	$(x^2 + 1)y^3 + y^2 + (x^2 + x + 1)y + x^2$	14	29
3	3	$(x^3 + x + 1)y^3 + x^2y^2 + y + x^3$	85	102
4	3	$(x^4 + x^2 + x + 1)y^3 + y^2 + y + x^3 + x^2 + x$	373	775
5	3	$(x^5 + x^2 + 1)y^3 + y^2 + (x^5 + x^3 + 1)y + x^5 + x^4 + x$	7621	7688
1	4	$(x + 1)y^4 + y^2 + (x + 1)y + x$	12	17
2	4	$x^2y^4 + (x^2 + x + 1)y^3 + (x + 1)y + x^2 + x$	26	102
3	4	$(x^3 + x^2 + x + 1)y^4 + (x^3 + x + 1)y^3 + (x^2 + 1)y + x^3$	375	774
4	4	$(x^4 + 1)y^4 + (x^4 + x^3 + 1)y^3 + (x + 1)y + x^4 + x^3$	5209	6151
1	5	$(x + 1)y^5 + (x + 1)y^2 + y + x$	21	35
2	5	$(x^2 + 1)y^5 + y^4 + xy^3 + x^2y^2 + (x + 1)y + x^2 + x$	122	486
3	5	$(x^3 + x^2 + 1)y^5 + y^4 + x^3y^2 + (x + 1)y + x^3 + x^2 + x$	15241	15366

$q = 3$:

h	d	P	orbit size	bound
1	1	$y + x$	2	4
2	1	$(x^2 + 1)y + x$	3	6
3	1	$(x^3 + 2x^2 + 1)y + x$	4	8
4	1	$(2x^4 + x + 1)y + x$	5	10
5	1	$(x^5 + 2x^2 + 1)y + x$	7	12
6	1	$(x^6 + x + 1)y + x^2$	7	12
7	1	$(x^7 + 2x^3 + 1)y + x$	13	18
1	2	$xy^2 + y + x$	3	5
2	2	$(x^2 + 1)y^2 + y + x^2 + x$	7	10
3	2	$(x^3 + 2x + 2)y^2 + y + x^2 + x$	25	59
4	2	$(x^3 + 2x + 2)y^2 + y + x^4$	79	168

TABLE 2. Polynomials in $\mathbb{F}_q[x, y]$ for which the initial state achieves the maximum orbit size under $\lambda_{0,0}$ for given values of q, h, and d. The final column contains the value of $q^{(h-1)(d-1)}\mathcal{L}(h, d, d) + \lfloor \log_q h \rfloor + \lceil \log_q \max(h, d-1) \rceil + 3$ from Theorem 1.

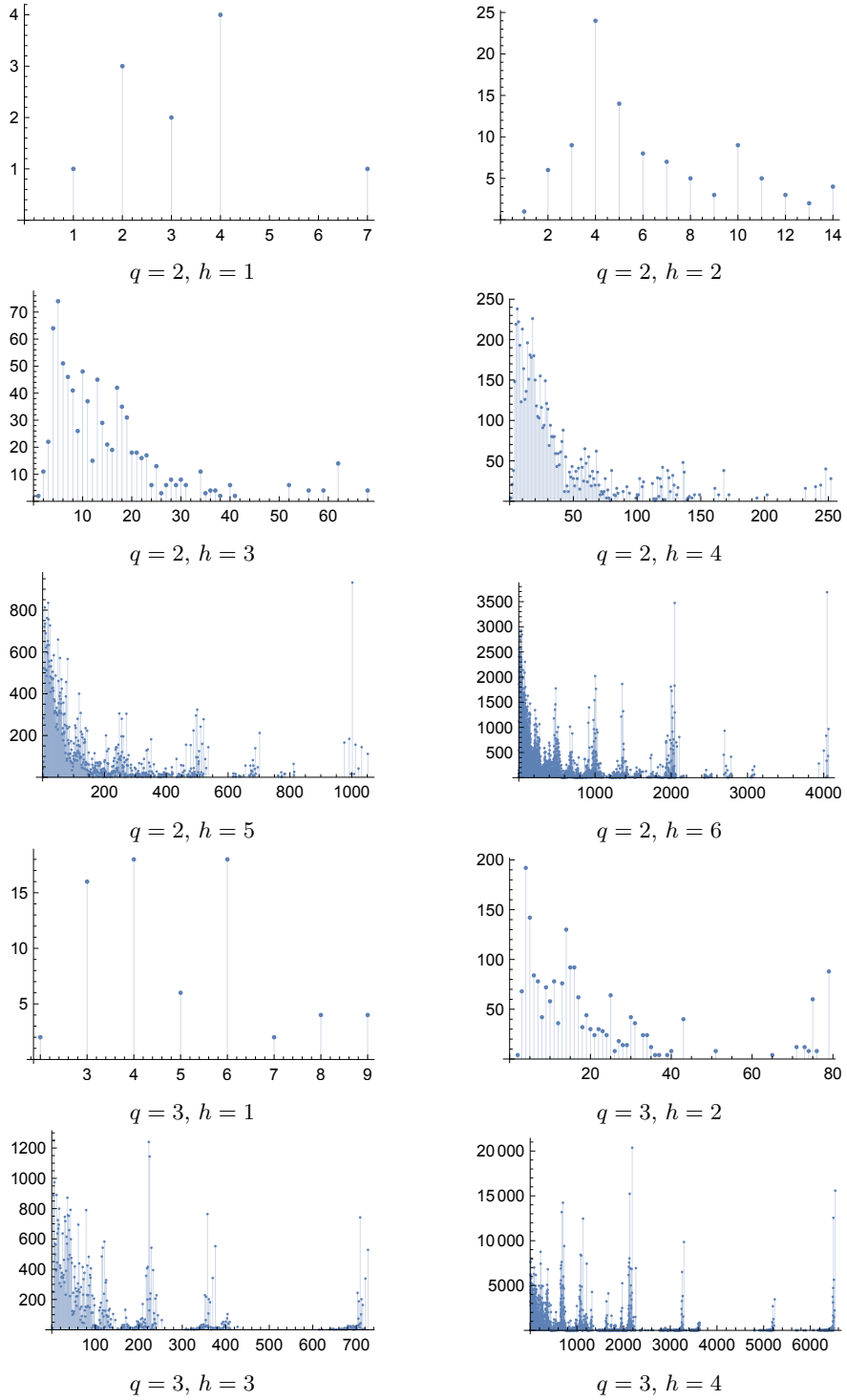


FIGURE 2. Number of polynomials (vertical axis) with degree $d = 2$ that produce unminimized automata with a given size (horizontal axis). The top six plots are for $q = 2$ and vary $h \in \{1, 2, \ldots, 6\}$. In the bottom four, $q = 3$ and $h \in \{1, 2, 3, 4\}$.

References

[1] Boris Adamczewski and Jason Bell, Diagonalization and rationalization of algebraic Laurent series, *Annales Scientifiques de l'École Normale Supérieure* **46** (2013) 963–1004.

[2] Boris Adamczewski, Alin Bostan, and Xavier Caruso, A sharper multivariate Christol's theorem with applications to diagonals and Hadamard products, `https://arxiv.org/abs/2306.02640`.

[3] Boris Adamczewski and Reem Yassawi, A note on Christol's theorem, `https://arxiv.org/abs/1906.08703`.

[4] Jean-Paul Allouche and Jeffrey Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press (2003).

[5] Peter Beelen, A generalization of Baker's theorem, *Finite Fields and Their Applications* **15** (2009) 558–568.

[6] Alin Bostan, Xavier Caruso, Gilles Christol, and Philippe Dumas, Fast coefficient computation for algebraic power series in positive characteristic, *The Open Book Series* **2** (Proceedings of the Thirteenth Algorithmic Number Theory Symposium, 2019) 119–135.

[7] Andrew Bridy, Automatic sequences and curves over finite fields, *Algebra & Number Theory* **11** (2017) 685–712.

[8] Rob Burns, Structure and asymptotics for Catalan numbers modulo primes using automata, `https://arxiv.org/abs/1701.02975`.

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

[10] Gilles Christol, Fonctions et éléments algébriques, *Pacific Journal of Mathematics* **125** (1986) 1–37.

[11] Gilles Christol, Teturo Kamae, Michel Mendès France, and Gérard Rauzy, Suites algébriques, automates et substitutions, *Bulletin de la Société Mathématique de France* **108** (1980) 401–419.

[12] Pierre Deligne, Intégration sur un cycle évanescent, *Inventiones Mathematicae* **76** (1984) 129–143.

[13] Jan Denef and Leonard Lipshitz, Algebraic power series and diagonals, *Journal of Number Theory* **26** (1987) 46–67.

[14] Harry Furstenberg, Algebraic functions over finite fields, *Journal of Algebra* **7** (1967) 271–277.

[15] Takashi Harase, Algebraic elements in formal power series rings II, *Israel Journal of Mathematics* **67** (1989) 62–66.

[16] Manuel Kauers and Peter Paule, *The Concrete Tetrahedron*, SpringerWienNewYork, Vienna (2011).

[17] Edmund Landau, Über die Maximalordung der Permutation gegbenen Grades, *Archiv der Mathematik und Physik* Series 3, **5** (1903) 92–103.

[18] Eric Rowland, IntegerSequences, `https://github.com/ericrowland/IntegerSequences`.

[19] Eric Rowland, IntegerSequences: a package for computing with k-regular sequences, International Congress on Mathematical Software, *Lecture Notes in Computer Science* **10931** (2018) 414–421.

[20] Eric Rowland, What is an automatic sequence?, *Notices of the American Mathematical Society* **62** (2015) 274–276.

[21] Eric Rowland and Reem Yassawi, Automatic congruences for diagonals of rational functions, *Journal de Théorie des Nombres de Bordeaux* **27** (2015) 245–288.

[22] Eric Rowland and Reem Yassawi, Algebraic power series and their automatic complexity modulo prime powers, `https://arxiv.org/abs/2408.00750`.

[23] Neil Sloane et al., The On-Line Encyclopedia of Integer Sequences, `https://oeis.org`.

Department of Mathematics, Hofstra University, Hempstead, NY, USA

Department of Mathematics, University of Liège, Allée de la Découverte 12, 4000 Liège, Belgium

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK