NONCOMMUTATIVE RATIONAL PÓLYA SERIES

JASON BELL AND DANIEL SMERTNIG

ABSTRACT. A (noncommutative) Pólya series over a field K is a formal power series whose nonzero coefficients are contained in a finitely generated subgroup of $K^\times$. We show that rational Pólya series are unambiguous rational series, proving a 40 year old conjecture of Reutenauer. The proof combines methods from noncommutative algebra, automata theory, and number theory (specifically, unit equations). As a corollary, a rational series is a Pólya series if and only if it is Hadamard sub-invertible. Phrased differently, we show that every weighted finite automaton taking values in a finitely generated subgroup of a field (and zero) is equivalent to an unambiguous weighted finite automaton.

1. INTRODUCTION AND MAIN RESULTS

Let K be a field. A univariate formal power series

$$S = \sum_{n\geq 0} s(n)x^n \in K[\![x]\!]$$

is a rational series if it is the power series expansion of a rational function at 0. Necessarily, this rational function does not have a pole at 0. Equivalently, the coefficients of a rational series satisfy a linear recurrence relation, that is, there exist $\alpha_1, \ldots, \alpha_m \in K$ such that

$$s(n+m) = \alpha_1 s(n+m-1) + \cdots + \alpha_m s(n) \quad \text{for all } n \geq 0.$$

Pólya [Pó21] considered arithmetical properties of rational series over $K = \mathbb{Q}$, and characterized the univariate rational series whose coefficients are supported at finitely many prime numbers. This was later extended to number fields by Benzaghou [Ben70, Chapitre 5], and to arbitrary fields, in particular fields of positive characteristic, by Bézivin [Bé87]. Ultimately, they proved the following theorem.

We call a rational series $S \in K[\![x]\!]$ a Pólya series if there exists a finitely generated subgroup $G \leq K^\times$, such that all coefficients of S are contained in $G \cup \{0\}$.

2010 *Mathematics Subject Classification.* Primary 68Q45, 68Q70; Secondary 11B37.

Key words and phrases. noncommutative rational series, weighted finite automata, Pólya series, Hadamard sub-invertibility, unambiguous rational series.

Theorem 1.1 (Pólya; Benzaghou; Bézivin)**.** *Let K be a field, let $S = P/Q \in K(x)$ be a rational function with $Q(0) \neq 0$, and let*

$$S = \sum_{n=0}^{\infty} s(n)x^n \in K[\![x]\!]$$

be the power series expansion of S at 0. Suppose that S is a Pólya series. Then there exist a polynomial $T \in K[x]$, $d \in \mathbb{Z}_{\geq 0}$, and for each $r \in [0, d-1]$ elements $\alpha_r \in K$, $\beta_r \in K^\times$ such that

$$S = T + \sum_{r=0}^{d-1} \frac{\alpha_r x^r}{1 - \beta_r x^d}.$$

Equivalently, there exists a finite set $F \subseteq \mathbb{Z}_{\geq 0}$ such that

$$s(kd + r) = \alpha_r \beta_r^k \qquad \text{for all } k \geq 0 \text{ and } r \in [0, d-1] \text{ with } kd + r \notin F.$$

The converse of the previous theorem, that every series with such coefficients is a Pólya series, is of course trivial.

Let R be a (commutative) domain. A noncommutative formal power series $S = \sum_{w \in X^*} S(w)w \in R\langle\langle X \rangle\rangle$ is rational if it can be obtained from noncommutative polynomials in $R\langle X \rangle$ by successive applications of addition, multiplication, and the star operation $S^* = (1 - S)^{-1} = \sum_{n \geq 0} S^n$ (if S has zero constant coefficient). See Section 2 below for formal definitions, and the book by Berstel–Reutenauer [BR11] for more background. Extending the correspondence between univariate rational series and linear recurrence relations to the noncommutative setting, a theorem of Schützenberger shows that S is rational if and only if it has a linear representation, or, equivalently, is recognized by a weighted finite automaton.

The definition of Pólya series extends to this noncommutative setting: A rational series $S \in R\langle\langle X \rangle\rangle$ is a Pólya series if its nonzero coefficients are contained in a finitely generated subgroup $G \leq K^\times$ of the quotient field K of R. Noncommutative rational Pólya series were first studied by Reutenauer in 1979 [Reu79]. Reutenauer introduced the notion of an unambiguous rational series and conjectured that these should be precisely the rational Pólya series [Reu79]; see also [Reu80], [Reu96, §6] and [BR11, p.233, Open Problem 4]. He proved many equivalent characterizations of unambiguous rational series, for instance, showing that they are precisely the ones being recognized by an unambiguous weighted finite automaton. The conjecture that rational Pólya series are unambiguous however so far remained open.

The goal of the present paper is to prove this conjecture. We also recover a new proof of Theorem 1.1 as a special case of our more general theorem, and give a characterization of those rational series recognized by a deterministic weighted finite automaton. Moreover, this resolves all three conjectures in [Reu80, Chapitre 6].

A rational series is unambiguous if it can be obtained from noncommutative polynomials and the operations of addition, multiplication, and the star operation

$S^* = (1-S)^{-1} = \sum_{n\geq 0} S^n$ in such a way that, in these operations, one never forms a sum of two nonzero coefficients. (This is defined more formally in Definition 2.4 below; in Section 2 we also recall the definitions of rational series and (unambiguous) weighted automata.) A formal series $S \in \mathbb{Z}\langle\langle X\rangle\rangle$ is linearly bounded if there exists $C \geq 0$ such that $|S(w)| \leq C|w|$ for all nonempty words $w \in X^*$.

Let R be a (commutative) domain and K its quotient field. An element $a \in K$ is almost integral over R if there exists $0 \neq c \in R$ such that $ca^n \in R$ for all $n \geq 0$. A domain R is completely integrally closed if it contains all such almost integral elements. We are mostly interested in the cases where $R = K$ is a field or $R = \mathbb{Z}$. In general, one cannot relax the completely integrally closed condition in the following theorem to integrally closed—see Remark 9.2 and the example following it.

Theorem 1.2. *Let R be a completely integrally closed domain with quotient field K. Let X be a finite non-empty set, and let $S \in R\langle\langle X\rangle\rangle$ be a rational series. Then the following statements are equivalent.*

(a) *S is a Pólya series.*
(b) *S is recognized by an unambiguous weighted finite automaton with weights in R.*
(c) *S is unambiguous (over R).*
(d) *There exist $\lambda_1, \ldots, \lambda_k \in R \smallsetminus \{0\}$, linearly bounded rational series $a_1, \ldots, a_k \in \mathbb{Z}\langle\langle X\rangle\rangle$, and a rational language $\mathcal{L} \subseteq X^*$ such that $\operatorname{supp}(a_i) \subseteq \mathcal{L}$ for all $i \in [1,k]$ and*
$$S(w) = \begin{cases} \lambda_1^{a_1(w)} \cdots \lambda_k^{a_k(w)} & \text{if } w \in \mathcal{L}, \\ 0 & \text{if } w \notin \mathcal{L}. \end{cases}$$
(e) *S is Hadamard sub-invertible, that is, the series*
$$\sum_{w \in \operatorname{supp}(S)} S(w)^{-1} w \in K\langle\langle X\rangle\rangle$$
is a rational series (over K).

The 'hard' part of this theorem is showing (a) ⇒ (b) in the case where $R = K$ is a field. It involves the use of finiteness results on unit equations in characteristic 0 and a recent extension of Derksen–Masser to positive characteristic [DM12]. We also make use of a new invariant associated to a linear representation, its linear hull. The linear hull also allows a characterization of determinizable weighted automatons, see Theorem 1.3 below.

The other implications of Theorem 1.2 are comparatively straightforward and are largely known. The equivalence (b) ⇔ (c) was first noted by Reutenauer [Reu80, Chapitre VI, Théorème 1]. The implications (c) ⇒ (e) and (e) ⇒ (a) are also known [BR11, Exercise 3.1 of Chapter 6], and so once (a) ⇒ (b) is shown, the equivalence of (a), (b), (c), and (e) is clear. Finally, (a) ⇔ (d) appears in the proof of [Reu79,

Proposition 4(ii)]. Despite this, we opt to give a self-contained proof of all of Theorem 1.2 in the present paper.

Denote by $\mathbb{1}_{\mathcal{L}}$ the characteristic series of a set $\mathcal{L} \subseteq X^*$. By a theorem of Schützenberger [BR11, Corollary 9.2.6], any linearly bounded rational series $a \in \mathbb{Z}\langle\langle X \rangle\rangle$ can be expressed as a $\mathbb{Z}$-linear combination of series of the form $\mathbb{1}_{\mathcal{L}}$ and $\mathbb{1}_{\mathcal{L}}\mathbb{1}_{\mathcal{K}}$ for rational languages $\mathcal{L}$, $\mathcal{K}$. This gives a more explicit description of the series appearing as exponents in (d) of Theorem 1.2.

In computer science the question whether a given weighted automaton is equivalent to a deterministic (sequential) one has received considerable attention; we mention the survey [LS06]. The question is of theoretical importance but also of practical relevance in natural language processing [Moh97, BGW00, MR17]. In this context, in contrast to our setting, however K is usually a tropical semiring. Nevertheless one can also pose these questions for K a field, as is done for instance in [LS06, §5].

When K is a field, the linear hull (see Definition 3.6) allows a characterization of rational series recognized by a deterministic weighted automaton. See Section 10 for another characterization, using bounded variation, in the spirit of Mohri [Moh97].

Theorem 1.3. *Let K be a field and X a finite non-empty set. For a rational series $S \in K\langle\langle X \rangle\rangle$, the following statements are equivalent.*

(a) *S is recognized by a deterministic weighted automaton.*
(b) *If (u, μ, v) is a minimal linear representation of S, then its linear hull has dimension at most 1.*

Rephrasing Theorem 1.3, a weighted automaton (with weights in a field K) is equivalent to a deterministic one if and only if (b) holds. We also obtain the following (already known) corollary.

Corollary 1.4. *If $S \in K\langle\langle X \rangle\rangle$ is a rational series whose coefficients take only* finitely *many values, then S is recognized by a deterministic weighted automaton. In particular, if K is a finite field, then every rational series over K is recognized by a deterministic weighted automaton.*

Notation. Throughout the paper, let R be a (commutative) domain with quotient field K. Often we will be concerned only with the case where $R = K$ is a field. Let X be a finite non-empty set. Let $G \leq K^\times$ be a finitely generated subgroup, and set $G_0 = G \cup \{0\}$. When considering Pólya series S, we will assume that G is such that G_0 contains all coefficients of S.

Outline. The paper is organized as follows. In Section 2 we recall necessary background on rational series. In Section 3 we introduce a useful topology and the notion of a linear hull; we also make a first crucial reduction in Lemma 3.13 and obtain Corollary 1.4. In Section 4 we use unit equations to prove a key lemma, with the

majority of the work dedicated to dealing with positive characteristic. In Sections 3 and 4 we restrict to the case where $R = K$ is a field. Now we can prove Theorem 1.2 over fields: in Section 5 we prove the hard direction (a) $\Rightarrow$ (b). In Section 6 we show (b) $\Leftrightarrow$ (c), in Section 7 we show (c) $\Rightarrow$ (d). The implications (c) $\Rightarrow$ (e) and (e) $\Rightarrow$ (a) are shown in Section 8. In Section 9 we put all these pieces together and extend the main implication (a) $\Rightarrow$ (b) from fields to completely integrally closed domains, then prove Theorems 1.1 and 1.2. In Section 10 we conclude the proof of Theorem 1.3.

Acknowledgments. We thank Christophe Reutenauer for many helpful comments on an earlier version of this manuscript, and Daniela Petrisan, Amaury Pouly, as well as Jacques Sakarovitch for fruitful discussions on topics related to the paper. We are grateful to the anonymous referee for their careful reading and helpful comments; in particular their request for the addition of illustrative examples will surely be appreciated by readers.

Bell was supported by NSERC grant RGPIN-2016-03632. Smertnig was supported by the Austrian Science Fund (FWF) project J4079-N32.

2. Preliminaries: rational series, linear representations, and weighted automata

We briefly recall the definitions of (noncommutative) rational series, linear representations, and weighted automata and how they relate to each other. We largely follow the notation and terminology from [BR11].

Let X^* denote the free monoid on the alphabet X. For a (noncommutative) formal power series $S \in R\langle\langle X \rangle\rangle$ and a word $w \in X^*$, we write $S(w)$ for the coefficient of w, that is

$$S = \sum_{w \in X^*} S(w)w.$$

The support of S is $\operatorname{supp}(S) = \{\, w \in X^* : S(w) \neq 0 \,\}$.

The ring of rational series in X is the smallest subring of the power series ring $R\langle\langle X \rangle\rangle$ that contains the noncommutative polynomials $R\langle X \rangle$ and is closed under addition, multiplication, and the partial operation

$$S \mapsto S^* := (1 - S)^{-1} = \sum_{n \geq 0} S^n$$

whenever S has zero constant coefficient.

If $X = \{x\}$ is a singleton, then $R\langle\langle X \rangle\rangle = R[\![x]\!]$. One can easily check that $S \in R[\![x]\!]$ is a rational series if and only if there exist polynomials P, $Q \in R[x]$ with $Q(0) = 1$ such that $S = P/Q$ [BR11, Proposition 6.1.1]. In other words, S is rational if and only if it is the power series expansion, at the point 0, of a rational function not having a pole at 0. If $R = K$ is a field, it is also well-known (and not hard to check) that this is the case if and only if the coefficients of S satisfy a linear recurrence relation.

Equivalently, there exist vectors $u \in K^{1\times n}$, $v \in K^{n\times 1}$, and a matrix $A \in K^{n\times n}$ such that $S(x^i) = uA^iv$ for every $i \geq 0$.

A fundamental theorem of Schützenberger extends this description to multivariate noncommutative rational series. A linear representation of rank (dimension) n is a triple (u, μ, v) where $u \in R^{1\times n}$ and $v \in R^{n\times 1}$ are vectors, and $\mu\colon X^* \to R^{n\times n}$ is a monoid homomorphism from the free monoid X^* to multiplicative monoid of $n \times n$-matrices. Schützenberger showed that $S \in R\langle\langle X\rangle\rangle$ is rational if and only if there exists a linear representation (u, μ, v) such that $S(w) = u\mu(w)v$ for every $w \in X^*$; see [BR11, Theorem 1.7.1].

Suppose $R = K$ is a field. A linear representation is minimal if the dimension n is minimal among all possible linear representations of S. This is the case if and only if the span of $u\mu(X^*) = \{\, u\mu(w) : w \in X^* \,\}$ is $K^{1\times n}$ and the span of $\mu(X^*)v$ is $K^{n\times 1}$.

There is another, graph-theoretical, way to view linear representations over the domain R that will come in handy. A weighted (finite) automaton $\mathcal{A} = (Q, I, E, T)$ over the alphabet X with weights in R consists of a finite set of states Q and three maps

$$I\colon Q \to R, \quad E\colon Q \times X \times Q \to R, \quad T\colon Q \to R.$$

A triple $(p, x, q) \in Q \times X \times Q$ is an edge if $E(p, x, q) \neq 0$. More specifically, we say that there is an edge from p to q labeled by x and with weight $E(p, x, q)$. A state $p \in Q$ is initial if $I(p) \neq 0$ and terminal if $T(p) \neq 0$.

A path is a sequence of edges

$$P = (p_0, x_1, p_1)(p_1, x_2, p_2) \cdots (p_{l-1}, x_l, p_l).$$

Its weight is $E(P) = \prod_{i=1}^{l} E(p_{i-1}, x_i, p_i)$ and its label is the word $x_1 \cdots x_l \in X^*$. The path is accepting if p_0 is an initial state and p_l is a terminal state. The automaton is trim if every state lies on an accepting path.

The series $S \in R\langle\langle X\rangle\rangle$ is recognized by $\mathcal{A}$ if

$$S(w) = \sum_{\substack{p_0,p_1,\ldots,p_l \in Q \\ w = x_1\cdots x_l,\, x_i \in X}} I(p_0)E(p_0, x_1, p_1) \cdots E(p_{l-1}, x_l, p_l)T(p_l). \tag{1}$$

Thus, the coefficient $S(w)$ is obtained by summing the weights of all accepting paths labeled by the word w, weighing each path by initial/terminal weights. Two automata are equivalent if they recognize the same series. Obviously every weighted automaton is equivalent to a trim one.

There is an easy correspondence between linear representations and weighted automata. Explicitly, the weighted automaton associated to a linear representation (u, μ, v) is given by $Q = [1, n]$, with $I(k) = u_k$, with $T(k) = v_k$, and $E(k, x, l) = \mu(x)_{k,l}$; here the subscripts denote the corresponding coordinates of the vectors u and v, respectively the entries of the matrix $\mu(x)$. Conversely, for a weighted automaton, one may without loss of generality assume $Q = [1, n]$, and then the correspondence

above yields a linear representation (a different labeling of the states gives a conjugate linear representation, corresponding to a permutation of the basis vectors). A series is recognized by the weighted finite automaton if and only if it is recognized by the associated linear representation. Hence series recognized by automata and series with linear representations are the same, and by Schützenberger's Theorem coincide with rational series.

Definition 2.1. Let $\mathcal{A}$ be a weighted automaton. Then $\mathcal{A}$ is unambiguous if each $w \in X^*$ labels at most one accepting path. It is deterministic (or sequential) if

- there exists at most one initial state; and
- for each $(p, x) \in Q \times X$, there exists at most one $q \in Q$ with $E(p, x, q) \neq 0$.

Note that for an unambiguous automaton, in the expression (1) for $S(w)$, at most one summand is nonzero. Deterministic weighted automata are clearly unambiguous.

Remark 2.2. For automata without weights (equivalently, weights in the Boolean semiring $\mathcal{B} = \{0, 1\}$ with $1 + 1 = 1$), it is well known that every automaton is equivalent to a deterministic one. This is no longer true for weighted automata; there exist unambiguous weighted automata that are not equivalent to deterministic ones, and there exist weighted automata that are not equivalent to unambiguous ones.

Definition 2.3. A rational series $S \in R\langle\langle X\rangle\rangle$ is a Pólya series if there exists a finitely generated subgroup $G \leq K^\times$ of the quotient field K of R such that $S(w) \in G_0 = G \cup \{0\}$ for all $w \in X^*$.

Let $S, T \in R\langle\langle X\rangle\rangle$ be two series with $\mathcal{K} = \operatorname{supp}(S)$ and $\mathcal{L} = \operatorname{supp}(T)$. The addition $S + T$ is unambiguous if $\operatorname{supp}(S) \cap \operatorname{supp}(T) = \emptyset$; the product ST is unambiguous if every $w \in \mathcal{K}\mathcal{L}$ has a *unique* expression $w = w_1 w_2$ with $w_1 \in \mathcal{K}$ and $w_2 \in \mathcal{L}$; and the star operation S^* is unambiguous if $\mathcal{K}$ is a code, that is, every $w \in \mathcal{K}^*$ has a *unique* expression $w = w_1 \cdots w_k$ with $w_i \in \mathcal{K}$.

Definition 2.4. The set of unambiguous rational series is the smallest subset of $R\langle\langle X\rangle\rangle$ that contains $R\langle X\rangle$ and is closed under *unambiguous* addition, multiplication, and star operation.

Note that unambiguous operations are defined in such a way that every coefficient of the resulting series is a product of coefficients of the initial series. That is, one never forms a sum of two nonzero coefficients. We thus we have the following.

Lemma 2.5. *Every unambiguous rational series is a Pólya series.*

3. The linear hull of a linear representation

In this section we consider only the case where $R = K$ is a field. We introduce a topology and a related invariant of a linear representation that will be essential in the proof of the implication (a) $\Rightarrow$ (b) of Theorem 1.2.

Definition 3.1. For a finite-dimensional vector space V, let $\mathcal{F}(V)$ be the collection of all subsets $Y \subseteq V$ of the form $Y = V_1 \cup \cdots \cup V_l$ with $l \in \mathbb{Z}_{\geq 0}$ and $V_i \subseteq V$ vector subspaces.

Lemma 3.2. *Every finite-dimensional vector space V has a noetherian topology for which $\mathcal{F}(V)$ is the collection of closed sets.*

Proof. Set $\mathcal{F} = \mathcal{F}(V)$. Clearly $V \in \mathcal{F}$ and $\emptyset \in \mathcal{F}$ (with $\emptyset$ represented by the empty union). By definition $\mathcal{F}$ is closed under finite unions. To show that $\mathcal{F}$ is the collection of closed sets of a topology, it remains to verify that $\mathcal{F}$ is closed under intersections. Every $Y \in \mathcal{F}$ is closed in the Zariski topology (identifying V with $\mathbb{A}^n_K$ for $n = \dim_K V$), which is noetherian. Hence, any intersection is equal to a finite subintersection. The claim follows since intersections distribute over unions, and intersections of vector subspaces are again vector subspaces. Since every $Y \in \mathcal{F}$ is Zariski-closed, the topology is noetherian. □

Definition 3.3. Let V be a finite-dimensional vector space. The linear Zariski topology on V is the topology whose collection of closed sets is $\mathcal{F}(V)$.

If $W \subseteq V$ is a vector subspace, then the subspace topology induced on W is the linear Zariski topology on W. All topological notions occurring in the remainder of the paper will refer to the linear Zariski topology. We mention [Bou72, §II.4.1 and §II.4.2] and [Sta19, Sections 004U and 0050] as references on irreducible and noetherian topological spaces.

A topological space X is irreducible if it is non-empty and $X = Z_1 \cup Z_2$ with Z_1, Z_2 closed implies $X = Z_1$ or $X = Z_2$. A subset $Z \subseteq X$ is an irreducible component if it is a maximal irreducible subspace. By $\mathcal{Z}(X)$ we denote the set of all irreducible components of X. Then $X = \bigcup_{Z \in \mathcal{Z}(X)} Z$.

Lemma 3.4. *The closed irreducible subsets of a vector space V are exactly the*
(1) *vector subspaces of V if K is infinite,*
(2) *vector subspaces of V of dimension ≤ 1 if K is finite.*

Proof. If K is infinite, then a vector space cannot be expressed as a finite union of proper vector subspaces. It follows that the closed irreducible subsets of V are exactly the vector subspaces. If K is finite, we can write every nonzero vector subspace of V as a finite union of one-dimensional vector subspaces. □

The dimension of a closed set is the maximal dimension of its irreducible components, with $\dim \emptyset = -\infty$.

We recall the following basic properties, of which we will make use throughout.

Lemma 3.5. *Let Y be a topological space.*
(1) *If Y is irreducible, $Y' \subsetneq Y$ is closed, and $\Omega \subseteq Y$ is dense, then $\overline{\Omega \setminus Y'} = Y$.*

(2) *If $Z \subseteq Y$ is irreducible, and $f\colon Y \to Y'$ is continuous, then $f(Z)$ is irreducible.*
(3) *If $Z \in \mathcal{Z}(Y)$ and $f\colon Y \to Y'$ is continuous, then $f(Z) \subseteq Z'$ for some $Z' \in \mathcal{Z}(Y')$.*
(4) *If Y is noetherian, then $\mathcal{Z}(Y)$ is finite.*
(5) *If $\mathcal{Z}(Y)$ is finite, $\Omega \subseteq Y$ is dense, and $Z \in \mathcal{Z}(Y)$, then $\overline{\Omega \cap Z} = Z$.*

Proof. (1) The set $U = Y \setminus Y'$ is non-empty and open and therefore dense in Y by irreducibility [Bou72, Proposition 1 of §II.4.1]. By basic topology, the intersection of a dense subset with an open subset is dense in the open subset. Thus $\Omega \setminus Y' = \Omega \cap U$ is dense in U, and by transitivity, in Y.

(2) [Bou72, Proposition 4 of §II.4.1] or [Sta19, Lemma 0379].

(3) By (2), the subset $f(Z)$ of Y' is irreducible. Every irreducible subset of Y' is a subset of an irreducible component by [Bou72, Proposition 5 of §II.4.2] or [Sta19, Lemma 004W].

(4) [Bou72, Proposition 10 of §II.4.2] or [Sta19, Lemma 0052].

(5) Let $Z_1, \ldots, Z_m$ be the irreducible components of Y, with $Z = Z_1$. Then $Y = \bigcup_{i=1}^m Z_i$. Let $U \subseteq Z_1$ be relatively open in Y and assume $U \cap \Omega = \emptyset$. We have to show $U = \emptyset$. Since irreducible components are closed ([Bou72, Proposition 2 of §II.4.1] or [Sta19, Lemma 004W]), the set $Z_1 \setminus U$ is closed in Y. Therefore $\Omega \subseteq (Z_1 \setminus U) \cup Z_2 \cup \cdots \cup Z_m$ implies $Z_1 \subseteq \overline{\Omega} \subseteq (Z_1 \setminus U) \cup Z_2 \cup \cdots \cup Z_m$. The irreducibility of Z_1, together with the incomparability of irreducible components, implies $Z_1 \subseteq Z_1 \setminus U$, that is $U = \emptyset$. □

We can now define a key invariant associated to a linear representation.

Definition 3.6. Let (u, μ, v) be a linear representation over the field K, and let

$$\Omega \coloneqq u\mu(X^*) = \{\, u\mu(w) : w \in X^* \,\}$$

be the (left) reachability set. The closed set $\overline{\Omega}$ is the (left) linear hull of (u, μ, v).

Before continuing, we illustrate the linear hull on two examples.

Example 3.7. Let $K = \mathbb{Q}$, let $X = \{a, b, c\}$, and define a linear representation (u, μ, v) by $u = (1, 1, 1)$, by $v = (1, 1, 0)^T$, and by

$$\mu(a) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \qquad \mu(b) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad \mu(c) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

The corresponding automaton is depicted in Figure 1. It is easy to see that (u, μ, v) is minimal. We claim $\overline{\Omega} = \langle e_1 + e_2, e_3 \rangle \cup \langle e_1 - e_2, e_3 \rangle$. The inclusion $\subseteq$ follows from $u \in \overline{\Omega}$ and the fact that $\langle e_1 + e_2, e_3 \rangle \cup \langle e_1 - e_2, e_3 \rangle$ is closed under right action by $\mu(x)$ for all $x \in \{a, b, c\}$. For the inclusion $\supseteq$, observe $u\mu(a)^n = (2^n, (-1)^n 2^n, 3^n)$. Therefore $\{\, u\mu(a)^{2n} : n \geq 0 \,\} \subseteq \Omega$ is dense in $\langle e_1 + e_2, e_3 \rangle$ while $\{\, u\mu(a)^{2n+1} : n \geq 0 \,\} \subseteq \Omega$ is dense in $\langle e_1 - e_1, e_3 \rangle$.

One could check that all nonzero coefficients of the series $S = 2 + 2b + 8a^2 + 10cb + 6ab + \cdots$ recognized by (u, μ, v) are of the form $2^e 3^f 5^g$ with e, f, $g \geq 0$, and that S is therefore a Pólya series. However, since we will construct an unambiguous linear representation recognizing the same series, we will see that S is Pólya a posteriori.

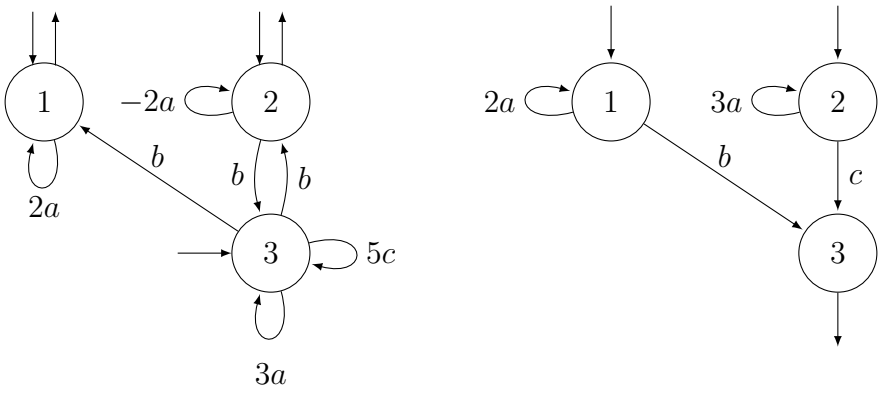

Figure 1. *Left: Example 3.7.* A weighted automaton recognizing a Pólya series. The automaton is minimal but ambiguous. A non-minimal unambiguous automaton on four states recognizes the same series (see Example 5.4 and Figure 2). *Right: Example 3.8.* A weighted automaton that is not determinizable and has a two-dimensional *left* linear hull. The same series can be recognized by a co-deterministic weighted automaton (that is, deterministic when reading words from right to left). Correspondingly the *right* linear hull is one-dimensional.

Example 3.8. Let $K = \mathbb{Q}$, let $X = \{a, b, c\}$, and define a minimal linear representation (u, μ, v) by $u = (1, 1, 0)$, $v = (0, 0, 1)^T$, and

$$\mu(a) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mu(b) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mu(c) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

The automaton is depicted in Figure 1. Here the (left) linear hull is $\overline{\Omega} = \overline{u\mu(X^*)} = \langle e_1, e_2 \rangle \cup \langle e_3 \rangle$. The dually defined right linear hull is $\overline{\mu(X^*)v} = \langle e_1 \rangle \cup \langle e_2 \rangle \cup \langle e_3 \rangle$. Observe that neither the dimension nor the number of components of the left and right linear hull coincide.

Let S be the series recognized by (u, μ, v). Then $S(a^n b) = 2^n$ and $S(a^n c) = 3^n$ for every $n \geq 0$ and $S(w) = 0$ for every other word w. By our convention, automata read words left to right. In view of Theorem 1.3, and its natural dual for reading words right to left, the asymmetry in the linear hull reflects that S can be recognized by

a deterministic automaton when reading words right to left, but not when reading words left to right.

Our next goal is to show that we can change the linear representation in such a way that the irreducible components of the linear hull, which are vector subspaces, form a direct sum (Lemma 3.13). We do this by forming the *external* direct sum of the irreducible components of the linear hull, thereby passing to a linear representation of possibly larger dimension. For instance, in Example 3.7, the linear hull is a union of two planes in affine 3-space, necessarily intersecting in a line, while the new linear representation will be defined in affine 4-space and have a linear hull consisting of two 2-dimensional planes (intersecting only in the origin). We will moreover do this in such a way that, if S is a Pólya series, then all coefficients appearing in vectors of Ω have their coordinates in G_0.

Let $Y \subseteq V$ be a closed subset of a vector space V, and $\varphi \in \mathrm{End}_K(V)$ with $\varphi(Y) \subseteq Y$. Then φ is continuous in our topology. Thus, if Z is an irreducible component of Y, there exists an irreducible component Z' of Y such that $\varphi(Z) \subseteq Z'$ (see (3) of Lemma 3.5). In particular, there exists a map $f\colon \mathcal{Z}(Y) \to \mathcal{Z}(Y)$ such that $\varphi(Z) \subseteq f(Z)$ for all $Z \in \mathcal{Z}(Y)$. In general, there are several possible choices for this map if $\varphi(Z)$ lies in an intersection of multiple irreducible components.

Definition 3.9. Let V be a finite-dimensional vector space and $Y \subseteq V$ a closed set. We define a vector space $\widehat{Y}$ together with a homomorphism $\sigma\colon \widehat{Y} \to V$ by

$$\widehat{Y} = \bigoplus_{Z \in \mathcal{Z}(Y)} Z$$

and $\sigma = \sum_{Z \in \mathcal{Z}(Y)} \pi_Z$, where $\pi_Z\colon \widehat{Y} \to Z \subseteq V$ denotes the canonical projection.

For $Z \in \mathcal{Z}(Y)$, let $\varepsilon_Z\colon Z \to \widehat{Y}$ denote the canonical embedding. For $\varphi \in \mathrm{End}_K(V)$ with $\varphi(Y) \subseteq Y$ and a map $f\colon \mathcal{Z}(Y) \to \mathcal{Z}(Y)$ with $\varphi(Z) \subseteq f(Z)$ for all $Z \in \mathcal{Z}(Y)$, let

$$\widehat{\varphi} = {}_f\widehat{\varphi}\colon \widehat{Y} \to \widehat{Y} \quad \text{be defined by} \quad \widehat{\varphi} \circ \varepsilon_Z = \varepsilon_{f(Z)} \circ \varphi|_Z.$$

The definition of $\widehat{\varphi}$ strongly depends on the choice of f. However, as we will see in a moment, $\sigma \circ \widehat{\varphi}$ does not depend on f. Since we will ultimately be interested in this composition, the particular choice of f will not matter, and we suppress f in the notation when this does not cause confusion.

Lemma 3.10. *Let all notation be as in Definition 3.9.*

(1) *For $Z \in \mathcal{Z}(Y)$ and $z \in Z$, we have $\varphi(z) = \sigma \circ \widehat{\varphi} \circ \varepsilon_Z(z)$.*
(2) *If $\varphi, \psi \in \mathrm{End}_K(V)$ with $\varphi(Y) \subseteq Y$ and $\psi(Y) \subseteq Y$, then*

$$\sigma \circ {}_h\widehat{(\psi \circ \varphi)} = \sigma \circ {}_g\widehat{\psi} \circ {}_f\widehat{\phi}.$$

for any choice of f, g, $h\colon \mathcal{Z}(Y) \to \mathcal{Z}(Y)$ with $\varphi(Z) \subseteq f(Z)$, $\psi(Z) \subseteq g(Z)$, and $\psi \circ \varphi(Z) \subseteq h(Z)$ for all $Z \in \mathcal{Z}(Y)$.

Proof. (1) We have

$$\sigma \circ \widehat{\varphi} \circ \varepsilon_Z(z) = \sigma \circ \varepsilon_{f(Z)} \circ \varphi(z) = \sum_{Z' \in \mathcal{Z}(Y)} \pi_{Z'} \circ \varepsilon_{f(Z)} \circ \varphi(z)$$
$$= \pi_{f(Z)} \circ \varepsilon_{f(Z)} \circ \varphi(z) = \varphi(z).$$

(2) Let $Z \in \mathcal{Z}(Y)$ and $z \in Z$. Then, by applying (1), we have

$$\sigma \circ {}_h\widehat{(\psi \circ \varphi)} \circ \varepsilon_Z(z) = \psi \circ \varphi(z).$$

On the other hand,

$$\sigma \circ {}_g\widehat{\psi} \circ {}_f\widehat{\varphi} \circ \varepsilon_Z(z) = \sigma \circ {}_g\widehat{\psi} \circ \varepsilon_{f(Z)} \circ \varphi(z) = \psi(\varphi(z)),$$

where have again applied (1) in the second equality. □

Lemma 3.11. *Let $S \in K\langle\langle X \rangle\rangle$ be a rational series and $\Gamma = \{ S(w) : w \in X^* \} \subseteq K$. Then S has a minimal linear representation (u, μ, v) with $u\mu(w) \in \Gamma^{1\times n}$ for all $w \in X^*$. If $n \geq 1$, we can take $v = e_1$.*

Proof. Let (u', μ', v') be a minimal linear representation of S. By minimality we have $\langle \mu'(w)v' : w \in X^* \rangle_K = K^{n\times 1}$. Let $w_1, \ldots, w_n \in X^*$ be such that the vectors $b_i = \mu'(w_i)v'$ for $i \in [1, n]$ form a basis of $K^{n\times 1}$. If $n \geq 1$, then $v' \neq 0$ by minimality, and we may take $w_1 = 1$ (the empty word), ensuring $b_1 = v'$. Let $B \in K^{n\times n}$ be the matrix whose i-th column is b_i, and let $\mu\colon X^* \to K^{n\times n}$ be defined by $\mu(w) = B^{-1}\mu'(w)B$. Let $u = u'B$ and $v = B^{-1}v'$. Then (u, μ, v) is a minimal linear representation of S.

Let $w \in X^*$ and $i \in [1, n]$. Then $u\mu(w) = u'\mu'(w)B$. The i-th coordinate of this vector is

$$u'\mu'(w)b_i = u'\mu'(ww_i)v' = S(ww_i) \in \Gamma.$$

By construction, we also have $v = B^{-1}v' = e_1$ if $n \geq 1$. □

Lemma 3.12. *Let V be a vector space with basis $e_1, \ldots, e_n$. Let $\Gamma \subseteq K$. For every vector subspace $W \subseteq V$, there exists a basis $f_1, \ldots, f_m$ of W such that*

$$(\Gamma e_1 + \cdots + \Gamma e_n) \cap W \subseteq \Gamma f_1 + \cdots + \Gamma f_m.$$

Proof. Using standard reductions and possibly renumbering the basis elements of V, we can find a basis $f_1, \ldots, f_m$ of W with $e_i^*(f_j) = \delta_{i,j}$ for $i, j \in [1, m]$ (essentially Gram-Schmidt).

Let

$$w = \alpha_1 e_1 + \cdots + \alpha_n e_n = \beta_1 f_1 + \cdots + \beta_m f_m$$

with $\alpha_1, \ldots, \alpha_n \in \Gamma$ and $\beta_1, \ldots, \beta_m \in K$. Applying e_i^* to the equation shows $\beta_i = \alpha_i \in \Gamma$ for $i \in [1, m]$. □

Lemma 3.13. *Let (u, μ, v) be a linear representation representing a rational series $S \in K\langle\langle X\rangle\rangle$ and $\Omega = u\mu(X^*)$. Let $\Gamma \subseteq K$ be such that $\Omega \subseteq \Gamma^{1\times n}$ and $\Omega v \subseteq \Gamma$. Let $\mathcal{Z}(\overline{\Omega}) = \{W_1, \dots, W_k\}$, let $m_i = \dim W_i$, and let $m = m_1 + \cdots + m_k$.*

For $i \in [1, k]$, let $W_i' \subseteq K^{1\times m}$ be the m_i-dimensional subspace spanned by the standard basis vectors $e_j \in K^{1\times m}$ with $j \in [m_1 + \cdots + m_{i-1} + 1, m_1 + \cdots + m_i]$, so that in particular

$$K^{1\times m} = W_1' \oplus \cdots \oplus W_k'.$$

Then there exists an m-dimensional linear representation (u', μ', v') of S with

$$\mathcal{Z}(\overline{\Omega}') = \{W_1', \dots, W_k'\}$$

where $\Omega' = u'\mu'(X^*)$. Moreover $\Omega' \subseteq \Gamma^{1\times m}$ and $\Omega' v' \subseteq \Gamma$.

Proof. If $n = 0$, then $(u', \mu', v') = (u, \mu, v)$ trivially has the desired properties. To avoid this degenerate case, from now on assume $n \geq 1$.

For each W_i choose a basis $f_{(i,1)}, \dots, f_{(i,m_i)}$ as in Lemma 3.12. Denote by $e_1, \dots, e_n$ the standard basis of $K^{1\times n}$. For $x \in X$ let $\varphi_x \colon K^{1\times n} \to K^{1\times n}$ denote the homomorphism that is represented by $\mu(x)$, and let $\psi \colon K^{1\times n} \to K$ denote the homomorphism represented by v. That is, $u\mu(x_1 \cdots x_l)v = \psi \circ \varphi_{x_l} \circ \cdots \circ \varphi_{x_1}(u)$ for $x_1, \dots, x_l \in X$.

For each $x \in X$, the homomorphism φ_x is continuous and closed and $\varphi_x(\Omega) \subseteq \Omega$. Hence also $\varphi_x(\overline{\Omega}) \subseteq \overline{\Omega}$. Without restriction $u \in W_1$. Let $Y = \overline{\Omega}$. We denote by $\varepsilon_i \colon W_i \to \widehat{Y}$ the canonical embedding. By Lemma 3.10,

$$\psi \circ \varphi_{x_l} \circ \cdots \circ \varphi_{x_1}(u) = \psi \circ \sigma \circ \widehat{\varphi}_{x_l} \circ \cdots \circ \widehat{\varphi}_{x_1} \circ \varepsilon_1(u). \tag{2}$$

Set $Q = \{\, (i,j) : i \in [1,k],\, j \in [1, m_i] \,\}$. Then the family $(\varepsilon_i(f_{(i,j)}))_{(i,j)\in Q}$ is a basis of $\widehat{Y}$. With respect to this choice of basis, let $u' \in K^{1\times Q}$ represent $\varepsilon_1(u)$, let $v' \in K^{Q\times 1}$ represent $\psi \circ \sigma \colon K^{1\times Q} \to K$, and let A_x' represent $\widehat{\varphi}_x \colon K^{Q\times Q} \to K^{Q\times Q}$. Setting $\mu'(x) = A_x'$, the tuple (u', μ', v') is a linear representation of S by Eq. (2).

Clearly $\Omega' \subseteq \varepsilon_1(W_1) \cup \cdots \cup \varepsilon_k(W_k)$, and we show that the right side is the decomposition of $\overline{\Omega'}$ into irreducible components. The sets $\varepsilon_i(W_i)$ are irreducible closed subsets of $\widehat{Y}$, and $\varepsilon_i(W_i) \cap \varepsilon_j(W_j) = 0$ if $i \neq j$. Let $i \in [1, k]$. The set

$$\Omega_i = (\Omega \cap W_i) \setminus \bigcup \{\, W_j : j \in [1,k],\, i \neq j \,\}$$

is dense in W_i. If $w \in X^*$ with $u\mu(w) \in \Omega_i$, then necessarily $u'\mu'(w) \in \varepsilon_i(W_i) \cap \Omega'$. Hence $\varepsilon_i(W_i) \cap \Omega'$ is dense in $\varepsilon_i(W_i)$.

Finally, by choice of the $f_{(i,j)}$, we have

$$\Omega \cap W_i \subseteq (\Gamma e_1 + \cdots + \Gamma e_n) \cap W_i \subseteq \Gamma f_{(i,1)} + \cdots + \Gamma f_{(i,m_i)}.$$

Thus $u'\mu'(w) \in \Gamma^{1\times Q}$ for all $w \in X^*$. If $w \in X^*$ with $u'\mu'(w) \in \Omega'$, then $u\mu(w) \in \Omega$ and $u'\mu'(w)v' = u\mu(w)v \in \Gamma$. □

We are now in a position to deal with the (relatively easy) case in which the linear hull has dimension at most 1. As an immediate corollary we obtain the direction (a) ⇒ (b) of Theorem 1.2 in the special case where S has only *finitely many* distinct coefficients.

Proposition 3.14. *Let $S \in K\langle\langle X\rangle\rangle$ be a rational series with a linear representation (u, μ, v) whose linear hull has dimension ≤ 1. Then S is recognized by a deterministic weighted automaton.*

Proof. Replacing the representation by one as in Lemma 3.13, we can assume that the spaces in $\mathcal{Z}(\overline{\Omega})$ form a direct sum. Thus, if $W \in \mathcal{Z}(\overline{\Omega})$, $a \in W$, and $x \in X$ with $0 \neq a\mu(x)$, then there exists a unique $W' \in \mathcal{Z}(\overline{\Omega})$ with $a\mu(x) \in W'$. This means that each column of $\mu(x)$ contains at most one non-zero entry. Hence the weighted automaton associated to this linear representation is deterministic. □

Proof of Corollary 1.4. Choose a representation as in Lemma 3.11 with $\Gamma = \{ S(w) : w \in X^* \}$ being finite. Then $\Omega = u\mu(X^*) \subseteq \Gamma^{1\times n}$ is a finite set, and hence $\overline{\Omega}$ is a finite union of vector spaces of dimension ≤ 1. Thus $\dim \overline{\Omega} \leq 1$, and Proposition 3.14 implies the claim. □

Remark 3.15. Corollary 1.4 was known before. If G is finite, then S is a finite linear combination of characteristic series of rational languages [BR11, Corollary 3.2.6]. Since each of these characteristic series can be recognized by a deterministic automaton, the result follows [Reu96, §6]. However, the proof we give here is more in line with the one for our general result.

4. An important lemma

In this section we again consider only the case where $R = K$ is a field. We are now ready to prove a key lemma in characteristic 0. Its proof depends on unit equations. As a consequence, a variant for positive characteristic is more complicated and will follow at the end of the section.

We recall the fundamental finiteness result on unit equations in characteristic 0 that we will be using. For number fields it was proved independently by Evertse [Eve84] and van der Poorten–Schlickewei [vdPS82]; the extension to arbitrary fields appears in [vdPS91]. We refer to [EG15, Chapter 6] or [BG06, Theorem 7.4.1] for more details.

Proposition 4.1 (Evertse; van der Poorten–Schlickewei)**.** *Suppose* $\operatorname{char} K = 0$. *Let $m \geq 2$, and $a_1, \ldots, a_m \in K^\times$. Then there exist only finitely many projective points $(x_1 : \cdots : x_m)$ with coordinates $x_1, \ldots, x_m \in G$ such that*

$$a_1x_1 + \cdots + a_mx_m = 0 \tag{3}$$

and $\sum_{i\in I} a_ix_i \neq 0$ for any non-empty, proper subset I of $[1, m]$.

A solution $(x_1, \dots, x_m)$ of (3) with $\sum_{i\in I} a_i x_i \neq 0$ for every $\emptyset \neq I \subsetneq [1,m]$ is called non-degenerate. So, counted as projective points, there are only finitely many non-degenerate solutions with coordinates in G. It is easily seen that there can be infinitely many degenerate solutions (even when considered as projective points), but by definition, the affine coordinates of the degenerate solutions lie in a finite union of proper vector subspaces.

Lemma 4.2. *Suppose* $\operatorname{char} K = 0$. *Let* V *be a vector space with basis* $e_1, \dots, e_n$. *Suppose that* $\Omega \subseteq G_0 e_1 + \cdots + G_0 e_n$ *is a dense subset of* V. *Then, for all* $\varphi \in \operatorname{Hom}_K(V, K)$ *with* $\varphi(\Omega) \subseteq G_0$, *there exists at most one* $i \in [1,n]$ *with* $\varphi(e_i) \neq 0$.

Proof. The claim is trivial for $n = 1$. Suppose $n \geq 2$. With respect to the basis $(e_1, \dots, e_n)$, the homomorphism φ is represented by $(\alpha_1, \dots, \alpha_n) \in K^{1\times n}$ with $\alpha_i = \varphi(e_i)$. Enlarging G if necessary, we may assume $\alpha_1, \dots, \alpha_n \in G_0$.

Let $I = \{\, i \in [1,n] : \alpha_i \neq 0 \,\}$. We must show $|I| \leq 1$. Suppose to the contrary that $|I| \geq 2$. After renumbering the basis vectors if necessary, we may assume $I = [1,m]$ for some $m \geq 2$.

For $\emptyset \neq J \subseteq I$ let

$$V_J = \Big\{ \lambda_1 e_1 + \cdots + \lambda_n e_n \in V : \sum_{j\in J} \alpha_j \lambda_j = 0 \Big\}.$$

By choice of I, each V_J is a proper subspace of V. Set $Y = \bigcup_{\emptyset\neq J\subseteq I} V_J$. Since K is an infinite field, a vector space cannot be covered by a finite union of proper subspaces. Thus $Y \subsetneq V$. Hence $\Omega' = \Omega \setminus Y$ is dense in V by Lemma 3.5.

If $v = \lambda_1 e_1 + \cdots + \lambda_n e_n \in \Omega'$, then

$$\sum_{i\in I} \alpha_i \lambda_i = g, \tag{4}$$

for some g in G_0 by assumption on φ. If $\emptyset \neq J \subseteq I$, then

$$\sum_{j\in J} \alpha_j \lambda_j \neq 0,$$

since $v \notin V_J$. Thus $(\alpha_1\lambda_1, \dots, \alpha_m\lambda_m, -g)$ is a non-degenerate solution of the unit equation $X_1 + \cdots + X_{m+1} = 0$.

Hence there exists a finite subset $M \subseteq \mathbb{P}^{m-1}(K)$ with $(\alpha_1\lambda_1 \colon \cdots \colon \alpha_m\lambda_m) \in M$ for all $v = \lambda_1 e_1 + \cdots + \lambda_n e_n \in \Omega'$. In particular, since $m \geq 2$, we see that λ_1/λ_2 can take only finitely many values. Thus Ω' can be covered by finitely many proper vector subspaces of V, in contradiction to $\overline{\Omega'} = V$. □

4.1. **Positive characteristic.** For this subsection, we now make the additional assumptions that $\operatorname{char} K = p > 0$, and that K is finitely generated over its prime field $\mathbb{F}_p$.

In extending Lemma 4.2 to positive characteristic, we face the problem that unit equations may have infinitely many non-degenerate solutions. However, a result of Derksen and Masser [DM12] is useful in bounding the number of solutions of bounded height. In this way, we will be able to recover a version of Lemma 4.2 with the original density hypothesis replaced by a quantitative one.

As in [DM12, Section 2], we can define a set of discrete valuations and associated absolute values on K in such a way that the absolute values satisfy the product formula. These absolute values depend on a choice of transcendence basis; we will always work with a fixed such set.

Associated to this set of absolute values, we define a (logarithmic, projective) height of an element $a = (\alpha_1 : \cdots : \alpha_n) \in \mathbb{P}^{n-1}(K)$ by

$$h(a) = \log \prod_v \max\{|\alpha_1|_v, \ldots, |\alpha_n|_v\}.$$

For $0 \neq (\alpha_1, \ldots, \alpha_n) \in K^{1\times n}$ we set $h(\alpha_1, \ldots, \alpha_n) = h(\alpha_1 : \cdots : \alpha_n)$. This height satisfies the Northcott property, that is, for every $N \geq 0$, the set $\{\, a \in \mathbb{P}^{n-1}(K) : h(a) \leq N \,\}$ is finite (at this point we are using that the prime field is finite). Moreover, for every $A \in K^{n\times n}$ there exists a constant C_A, such that for any $a \in K^{1\times n}$ with $a \notin \ker A$,

$$h(aA) \leq h(a) + C_A.$$

If A is invertible, then even

$$h(a) - C_A \leq h(aA) \leq h(a) + C_A.$$

(See [HS00, Theorem B.2.5].)

If V is a finite-dimensional vector space, then any choice of basis gives an isomorphism $V \to K^{1\times n}$, and therefore induces a corresponding height h_V on V and on the projective space $\mathbb{P}(V)$. If h_V, h'_V are two such heights, induced by different bases, then $h'_V(a) = h_V(a) + O(1)$. The exact choice of height will not matter.

We shall also need the following property.

Lemma 4.3. *The group*

$$\sqrt{G} \coloneqq \{\, a \in K : a^n \in G \text{ for some } n \geq 1 \,\} \leq K^\times$$

is finitely generated.

Proof. By assumption K is finitely generated over its prime field $\mathbb{F}_p$ and $G \leq K^\times$ is a finitely generated subgroup. Let $\mathbb{F}_q$ be the algebraic closure of $\mathbb{F}_p$ in K, so that K is a regular extension of $\mathbb{F}_q$. Let R be the finitely generated $\mathbb{F}_q$-subalgebra of K generated by G. Then the integral closure $\overline{R}$ is a finitely generated R-module by [Lan83, Proposition 2.4.1] or [Eis95, Corollary 13.13]. Hence $\overline{R}$ also is a finitely generated $\mathbb{F}_q$-algebra. Now [Lan83, Corollary 2.7.3] implies that $\overline{R}^\times$ is a finitely generated group; since $\sqrt{G} \subseteq \overline{R}^\times$, the claim follows. □

Let

$$\mathbb{P}^{n-1}(G) = \{\, (\alpha_1 : \cdots : \alpha_n) : \alpha_1, \ldots, \alpha_n \in G \,\}.$$

As a consequence of a theorem of Derksen–Masser [DM12, Theorem 3], we have an upper bound on the number of solutions of bounded height of a unit equation in positive characteristic.

Lemma 4.4. *Let $n \geq 2$ and let $S \subseteq \mathbb{P}^{n-1}(G)$ be the set of non-degenerate solutions in G to*

$$a_1 x_1 + \cdots + a_n x_n = 0, \qquad a_1, \ldots, a_n \in K^\times, \tag{5}$$

and for all $N \geq 0$ let $c(N) \coloneqq |\{\, a \in S : h(a) \leq N \,\}|$. Then there exists $D \in \mathbb{Z}_{\geq 0}$ such that

$$c(N) = O(\log(N)^D). \tag{6}$$

Proof. By the Lemma 4.3, the group $\sqrt{G}$ is finitely generated. A $\sqrt{G}$-automorphism is a map

$$\psi \colon \mathbb{P}^{n-1}(K) \to \mathbb{P}^{n-1}(K), \ (\alpha_1 : \cdots : \alpha_n) \mapsto (g_1\alpha_1 : \ldots : g_n\alpha_n)$$

with $g_1, \ldots, g_n \in \sqrt{G}$. For q a power of p, let

$$\varphi_q(\alpha_1 : \cdots : \alpha_n) = (\alpha_1^q : \cdots : \alpha_n^q).$$

Finally, for $\sqrt{G}$-automorphisms $\psi_1, \ldots, \psi_k$ and $a \in \mathbb{P}^{n-1}(K)$ let

$$[\psi_1, \ldots, \psi_k]_q(a) \coloneqq \{\, (\psi_1^{-1}\varphi_q^{e_1}\psi_1)(\psi_2^{-1}\varphi_q^{e_2}\psi_2)\cdots(\psi_k^{-1}\varphi_q^{e_k}\psi_k)(a) : e_1, \ldots, e_k \in \mathbb{Z}_{\geq 0} \,\}.$$

(We suppress the composition operator $\circ$ for brevity.)

By [DM12, Theorem 3], the set of solutions S of (5) is contained in a finite union of sets of the form $[\psi_1, \ldots, \psi_k]_q(a)$. It therefore suffices to show that (6) holds for S such a set. Thus, suppose $S = [\psi_1, \ldots, \psi_k]_q(a)$ for some $a \in \mathbb{P}^{n-1}(K)$ and $\sqrt{G}$-automorphisms $\psi_1, \ldots, \psi_k$.

By induction on k, we show that S contains $O((\log N)^k)$ elements of height at most N. The case $k = 0$ is clear. Suppose $k \geq 1$ and that the claim holds for $k-1$. There exists $C' \geq 0$ such that $h(\psi_1^{-1}(b)) \geq h(b) - C'$ and $h(\psi_1(b)) \geq h(b) - C'$ for all $b \in \mathbb{P}^{n-1}(K)$. Let

$$T_0 = \{\, b \in [\psi_2, \ldots, \psi_k]_q(a) : h(b) \leq C' + 1 \,\}, \text{ and}$$
$$T_1 = \{\, b \in [\psi_2, \ldots, \psi_k]_q(a) : h(b) > C' + 1 \,\}.$$

The set T_0 is finite. For each $b \in T_0$ and $b' = \psi_1^{-1}\varphi_q^{e_1}\psi_1(b)$ with $e_1 \geq 0$, we have $h(b') \geq q^{e_1}h(\psi_1(b)) - C'$. Hence, the number of such elements b' with $h(b') \leq N$ is $O_b(\log N)$. Exploiting the finiteness of T_0, altogether there are $O(\log N)$ elements $b' \in \psi_1^{-1}\varphi_q^{e_1}\psi_1(T_0)$, with $e_1 \geq 0$, such that $h(b') \leq N$.

For $b \in T_1$ and $b' = \psi_1^{-1}\varphi_q^{e_1}\psi_1(b)$ with $e_1 \geq 0$, we have

$$h(b') \geq q^{e_1}(h(b) - C') - C' > q^{e_1} - C'.$$

Thus, if $h(b') \leq N$ then $e_1 \leq \log(N + C')/\log(q)$ and $h(b) \leq N + 2C'$. By the induction hypothesis, there are $O(\log(N + 2C')^{k-1}) = O(\log(N)^{k-1})$ elements $b \in T_1$ with $h(b) \leq N + 2C'$. Thus there are $O(\log(N)^k)$ elements $b' \in \psi_1^{-1}\varphi_q^{e_1}\psi_1(T_1)$, with $e_1 \geq 0$, for which $h(b') \leq N$.

Altogether,

$$S = \bigcup_{e_1 \geq 0} \psi_1^{-1}\varphi_q^{e_1}\psi_1(T_0) \cup \psi_1^{-1}\varphi_q^{e_1}\psi_1(T_1)$$

contains $O(\log(N)^k)$ elements of height at most N. □

We can now obtain a variant of Lemma 4.2, using a slightly stronger hypothesis, that also holds in positive characteristic. To ensure that a vector subspace of $K^{1\times n}$ is irreducible in our topology, we do still need to assume that K is infinite.

For a subset S of a vector space V, let $\mathbb{P}(S)$ be the image of $S \smallsetminus \{0\}$ in the projective space $\mathbb{P}(V)$.

Lemma 4.5. *Suppose K is an* infinite *field of positive characteristic, finitely generated over its prime field. Let V be a vector space with basis $e_1, \ldots, e_n$, and $\Omega \subseteq G_0e_1 + \cdots + G_0e_n$ a dense subset of V. Assume that, for every standard projection $\pi\colon V \to W$ with $W = \langle e_{i_1}, \ldots, e_{i_m}\rangle_K$ and $m \geq 2$, every closed subset $Y \subsetneq W$, and every $C \in \mathbb{R}_{\geq 0}$, $D \in \mathbb{Z}_{\geq 0}$, there exist arbitrarily large N such that*

$$|\{\, a \in \mathbb{P}(\pi(\Omega) \smallsetminus Y) : h_W(a) \leq N \,\}| > C\log(N)^D.$$

Then, for all $\varphi \in \mathrm{Hom}_K(V, K)$ with $\varphi(\Omega) \subseteq G_0$, there exists at most one $i \in [1, n]$ with $\varphi(e_i) \neq 0$.

Proof. The beginning of the proof and the overall strategy are analogous to Lemma 4.2. Let $\varphi(e_i) = \alpha_i \in G_0$ and $I = \{\, i \in [1, m] : \alpha_i \neq 0 \,\}$. Without restriction $I = [1, m]$ and we have to show $m = 1$. Assume $m \geq 2$. Let $\pi\colon V \to W$ with $W = \langle e_1, \ldots, e_m\rangle_K$ denote the standard projection.

For $\emptyset \neq J \subseteq I$, let $W_J = \{\, \lambda_1e_1 + \cdots + \lambda_me_m \in W : \sum_{j\in J} \alpha_j\lambda_j = 0 \,\}$ and $Y = \bigcup_{\emptyset\neq J\subseteq I} W_J$. Note that $Y \subsetneq W$. Thus, applying our assumption to points of bounded height in $\mathbb{P}(\pi(\Omega) \smallsetminus Y)$, we find that, for any $D \in \mathbb{Z}_{\geq 0}$, the set $\mathbb{P}(\pi(\Omega) \smallsetminus Y)$ contains more than $O(\log(N)^D)$ points of height at most N.

Suppose $\lambda_1e_1 + \cdots + \lambda_me_m$ represents a point in $\mathbb{P}(\pi(\Omega) \smallsetminus Y)$. Then there exists a $g \in G$ with

$$\sum_{j=1}^{m} \alpha_i\lambda_i = g.$$

Hence $(\alpha_1\lambda_1 : \cdots : \alpha_m\lambda_m : -g)$ is a non-degenerate solution of the unit equation $X_1 + \cdots + X_{m+1} = 0$. By Lemma 4.4, we conclude that there exist $O(\log(N)^D)$ such points $(\alpha_1\lambda_1 : \cdots : \alpha_m\lambda_m)$ of height at most N, a contradiction to the size of $\mathbb{P}(\pi(\Omega) \smallsetminus Y)$. □

The following example shows that the conclusion of the previous lemma is trivially false for finite fields.

Example 4.6. Let K be a finite field and $n \geq 2$. Let $A_1, \ldots, A_k \in \mathrm{GL}(n, K)$ be such that the residue classes generate $\mathrm{PGL}(n, K)$ as a semigroup, and let $X = \{a_1, \ldots, a_k\}$. Let $0 \neq u \in K^{1\times n}$, let $0 \neq v \in K^{n\times 1}$, and let $\mu(a_i) = A_i$ for $i \in [1, k]$. Since $\mathrm{PGL}(n, K) = \mu(X^*)$ acts transitively on $\mathbb{P}^{n-1}(K)$, the linear representation (u, μ, v) is minimal and its linear hull is $K^{1\times n}$, that is, the set $\Omega = u\mu(X^*)$ is dense in all of $K^{1\times n}$. With $G = K^\times$ and $G_0 = K$, any choice of $\lambda_1, \ldots, \lambda_n \in K$, yields a linear map $\varphi\colon K^{1\times n} \to K, (\beta_1, \ldots, \beta_n) \mapsto \lambda_1\beta_1 + \cdots + \lambda_n\beta_n$ with $\varphi(\Omega) \subseteq G_0$. In particular, we may take all λ_i to be nonzero, in contrast to the conclusion of the previous lemma.

We will later apply Lemma 4.2, respectively Lemma 4.5, to an irreducible component V of the linear hull $\overline{\Omega} = \overline{u\mu(X^*)}$ of a linear representation (u, μ, v). In characteristic 0 we may use Lemma 4.2, and in this case it is clear that $\Omega \cap V$ is dense in V. However, in positive characteristic, where we need to apply Lemma 4.5, it is necessary to verify that the stronger conditions of this lemma are indeed satisfied. This is the subject of the next two lemmas.

Let (u, μ, v) be a linear representation and $\Omega = u\mu(X^*)$. For $N \in \mathbb{Z}_{\geq 0}$, let $\Omega_{\leq N} = \{\, u\mu(w) : w \in X^*, |w| \leq N \,\}$.

Lemma 4.7. *Let $Y \subseteq K^{1\times n}$ be a closed set of dimension m with k irreducible components, and let $N = k^{|X|(m-1)} + 1$. If $\Omega_{\leq N} \subseteq Y$ then $\Omega \subseteq Y$.*

Proof. Starting from Y we iteratively construct a sequence of subsequently smaller closed subsets; it is easiest to keep track of the necessary data using disjoint unions of k rooted trees.

We thus construct a sequence of graphs $\mathcal{T}_1, \ldots, \mathcal{T}_l$ whose vertices are labeled by vector spaces contained in Y, and with each $\mathcal{T}_i$ having the following properties:

(1) If W' labels a child of a vertex labeled by W, then $W' \subsetneq W$.
(2) $\mathcal{T}_i$ is s-regular (except for the leaves) with $s = k^{|X|}$.
(3) If W labels a leaf and $x \in X$, then there exists a vertex labeled by W' such that $W\mu(x) \subseteq W'$.
(4) If W labels an internal vertex and $W_1, \ldots, W_s$ label its children, then

$$\Omega_{\leq N-i} \cap W \;\subseteq\; W_1 \cup \cdots \cup W_s.$$

The graph $\mathcal{T}_1$ has k roots labeled by the elements of $\mathcal{Z}(Y)$. For $W \in \mathcal{Z}(Y)$ and $F\colon X \to \mathcal{Z}(Y)$, let

$$W_F = \{\, a \in W : a\mu(x) \in F(x) \text{ for all } x \in X \,\}.$$

If, for a fixed W, each of the vector spaces W_F is a proper subspace of W, then we attach s children to the root labeled by W. These children are labeled by W_F for $F\colon X \to \mathcal{Z}(Y)$. On the other hand, if $W = W_F$ for some F, we do not attach any children to the vertex labeled by W. In this case, observe that $W\mu(x) \subseteq F(x)$ for every $x \in X$.

It is clear that $\mathcal{T}_1$ satisfies (1) and (2). Property (3) holds by the choice of the spaces. For (4) let $a \in \Omega_{\leq N-1} \cap W$. Taking any $x \in X$, we have $a\mu(x) \in W'$ for some $W' \in \mathcal{Z}(Y)$. Then $a \in W_F$ for any $F\colon X \to \mathcal{Z}(Y)$ with $F(x) = W'$.

We now iteratively construct $\mathcal{T}_i$ from $\mathcal{T}_{i-1}$ for $i \geq 2$. If, for every leaf of $\mathcal{T}_{i-1}$, say, labeled by W, and every $x \in X$, there exists a leaf labeled by W' such that $W\mu(x) \subseteq W'$, then we stop and set $l = i - 1$.

Otherwise, fix a leaf α labeled by W and an $x \in X$ such that $W\mu(x)$ is not contained in any label of a leaf of $\mathcal{T}_{i-1}$. By (3) there exists an internal vertex β labeled by W' such that $W\mu(x) \subseteq W'$ but $W\mu(x) \not\subseteq W''$ for any W'' labeling a child of β. By our construction W' has s children labeled by $W'_1, \ldots, W'_s$. Set $W_j = \{\, a \in W : a\mu(x) \in W'_j \,\}$, and attach s new children to α, labeled by $W_1, \ldots, W_s$.

It is clear that (1)–(3) are preserved for $\mathcal{T}_i$. Property (4) also carries over for vertices other than α. To verify (4) for α, let $a \in \Omega_{\leq N-i} \cap W$. Then $a\mu(x) \in W' \cap \Omega_{\leq N-(i-1)}$ and hence $a\mu(x) \in W'_j$ for some j by (4). Thus $a \in W_j$.

To see that this process terminates, note that each $\mathcal{T}_i$ is s-regular (except for the leaves) of height at most m, and hence has at most s^m vertices. Since each $\mathcal{T}_i$ has s vertices more than $\mathcal{T}_{i-1}$, the process terminates after at most s^{m-1} steps, that is $l \leq s^{m-1} + 1$.

Instead of (3), the final graph $\mathcal{T}_l$ has the stronger property that if W labels a leaf and $x \in X$, then there exists W' labeling a leaf of $\mathcal{T}_l$ such that $W\mu(x) \subseteq W'$. Defining Y' to be the union of all labels of leaves of T_l, this implies $Y'\mu(x) \subseteq Y'$ for each $x \in X$ and hence $Y'\mu(X^*) \subseteq Y'$. As Y' contains $u \in \Omega_{\leq N - s^{m-1} - 1}$, this implies $\Omega \subseteq Y' \subseteq Y$. □

We give an example in dimension $n = 3$ with $\mathcal{T}_2 \neq \mathcal{T}_1$, illustrating the iterative construction in the previous proof.

Example 4.8. Let e_1, e_2, $e_3 \in K^{1\times 3}$ be the standard unit vectors and let $X = \{a\}$. Consider (u, μ, v) with $u = e_1$, with $v = e_1^T$, and with $\mu(a) \in K^{3\times 3}$ the permutation matrix defined by $e_1\mu(a) = e_2$, $e_2\mu(a) = e_3$, and $e_3\mu(a) = e_1$. Clearly $\Omega = \{e_1, e_2, e_3\}$ and $\overline{\Omega} = \langle e_1\rangle \cup \langle e_2\rangle \cup \langle e_3\rangle$. Let $Y = \langle e_1, e_2\rangle \cup \langle e_2, e_3\rangle$ and note $\Omega \subseteq Y$.

Now $\mathcal{T}_1$ has two roots, labeled by $W_{1,2} \coloneqq \langle e_1, e_2\rangle$ and $W_{2,3} \coloneqq \langle e_2, e_3\rangle$. Since $W_{1,2}\mu(a) = W_{2,3}$, no children are attached to $W_{1,2}$ in $\mathcal{T}_1$. However $W_{2,3}\mu(a) = \langle e_1, e_3\rangle \not\subseteq Y$. Thus, in $\mathcal{T}_1$, two children are attached to the root labeled by $W_{2,3}$. These children are labeled by $\langle e_2\rangle = \{\, w \in W_{2,3} : w\mu(a) \subseteq W_{2,3} \,\}$ and $\langle e_3\rangle = \{\, w \in W_{2,3} : w\mu(a) \in W_{1,2} \,\}$.

In the next step, we observe that $W_{1,2}$ is not mapped into labels of leaves of $\mathcal{T}_1$ by $\mu(a)$. Thus, to the vertex labeled by $W_{1,2}$, we attach two new children labeled by $\langle e_1\rangle = \{\, w \in W_{1,2} : w\mu(a) \in \langle e_2\rangle \,\}$ and $\langle e_2\rangle = \{\, w \in W_{1,2} : w\mu(a) \in \langle e_3\rangle \,\}$. (There are two different vertices with label $\langle e_2\rangle$, one attached to each of the two roots.) At this point the process stops, because every label of a leaf is mapped into a label of a leaf by $\mu(a)$.

Lemma 4.9. *Let V be an irreducible component of $\overline{\Omega}$. Let $\pi\colon V \to W$ be an epimorphism with $\dim W \geq 2$. If $Y \subsetneq W$ is a closed subset, then there exists $C \in \mathbb{R}_{>0}$ such that*

$$|\{\, a \in \mathbb{P}(\pi(\Omega) \smallsetminus Y) : h_W(a) \leq N \,\}| \;\geq\; CN^{\frac{1}{e|X|}} \qquad \text{for } N \in \mathbb{Z}_{\geq 0},$$

where $e = \max\{\dim Y, 1\} + n - m$.

Proof. Extend π to $\pi\colon K^{1\times n} \to W$. Let $C > 0$ be such that $h(a\mu(x)) \leq h(a) + C$ for all $x \in X$ and $a \in K^{1\times n} \smallsetminus \ker(\mu(x))$. We may moreover assume $h_W(\pi(a)) \leq h(a) + C$ for all $a \in K^{1\times n} \smallsetminus \ker(\pi)$. If $a \in \Omega_{\leq M} \smallsetminus \ker(\pi)$, then $h_W(\pi(a)) \leq h(u) + (M+1)C$. Choosing $M = (N - h(u))/C - 1$ we find that all $a \in \Omega_{\leq M} \smallsetminus \ker(\pi)$ have $h_W(\pi(a)) \leq N$. Thus it suffices to show that $\mathbb{P}(\pi(\Omega_{\leq M}) \smallsetminus Y)$ contains at least $C'M^{\frac{1}{e|X|}}$ points for some $C' > 0$.

If $\mathbb{P}(\pi(\Omega_{\leq M}) \smallsetminus Y)$ contains l points, then $\Omega_{\leq M} \subseteq \pi^{-1}(Y) \cup \pi^{-1}(P_1) \cup \cdots \cup \pi^{-1}(P_l)$ for some 1-dimensional vector spaces $P_1, \ldots, P_l$. Since $\dim Y < m$ and $\overline{\pi(\Omega)} = W$, we have $\Omega \not\subseteq \pi^{-1}(Y) \cup \pi^{-1}(P_1) \cup \cdots \cup \pi^{-1}(P_l)$.

By Lemma 4.7, for $l \geq \max\{1, k\}$ with $k = |\mathcal{Z}(Y)|$,

$$M < (l+k)^{(e-1)|X|} + 1 \leq (2l)^{e|X|}. \qquad \square$$

5. Proof of (a) ⇒ (b)

Having made the necessary preparations, in this section we prove (a) ⇒ (b) of Theorem 1.2. Let $\mathcal{P}(Q)$ denote the power set of a set Q. The following lemmas very closely parallel the corresponding results on semi-monomial matrices used in the proof of a decomposition theorem for rational functions between free monoids; see [Sak09, Chapter V.2] and Remark 5.5.

Lemma 5.1. *Let $\mathcal{A} = (Q, I, E, T)$ be a weighted automaton on the alphabet X with coefficients in R. Suppose that there exists $\mathcal{S} \subseteq \mathcal{P}(Q)$ such that $Q = \bigcup_{M\in\mathcal{S}} M$, and all of the following conditions are satisfied.*

(1) *There exists an $M \in \mathcal{S}$ containing all initial states of $\mathcal{A}$.*
(2) *For $M \in \mathcal{S}$ and $x \in X$, there exists an $N \in \mathcal{S}$ such that, whenever there is an edge from $p \in M$ to $q \in Q$ labeled by x, then $q \in N$.*
(3) *For every state $q \in Q$, $x \in X$, and $M \in \mathcal{S}$, there exists at most one state $p \in M$ that has an edge from p to q labeled by x.*
(4) *Every $M \in \mathcal{S}$ contains at most one terminal state.*

Then $\mathcal{A}$ is unambiguous.

Proof. We need to show that for a word $w = a_1 \cdots a_l \in X^*$ with $a_1, \ldots, a_d \in X$, there exists at most one accepting path in $\mathcal{A}$ that is labeled by w. Suppose that there are two accepting paths

$$(p_0, a_1, p_1)(p_1, a_2, p_2) \cdots (p_{l-1}, a_l, p_l) \quad \text{and} \quad (q_0, a_1, q_1)(q_1, a_2, q_2) \cdots (q_{l-1}, a_l, q_l).$$

We first show that for every $j \in [0, l]$, there exists a set $M_j \in \mathcal{S}$ with p_j, $q_j \in M_j$. For $j = 0$, note that p_0 and q_0 are initial states, hence by (1), there exists $M_0 \in \mathcal{S}$ with p_0, $q_0 \in M_0$. Now, if $j \in [1, l]$ and p_{j-1}, $q_{j-1} \in M_{j-1}$, then (2) implies that there exists $M_j \in \mathcal{S}$ with p_j, $q_j \in M_j$.

Since the paths are accepting, p_l and q_l are terminal states. Since p_l, $q_l \in M_l$, condition (4) implies $p_l = q_l$. If $p_j = q_j$ for some $j \in [1, l]$, then, since we already know p_{j-1}, $q_{j-1} \in M_{j-1}$, condition (3) implies $p_{j-1} = q_{j-1}$. Thus, altogether we have $p_j = q_j$ for all $j \in [1, l]$ and hence the two paths are the same. Thus we have shown that $\mathcal{A}$ is unambiguous. □

For the statement of the next lemma we fix the following notation: Let $e_1, \ldots, e_n$ denote the standard basis vectors of $R^{1\times n}$. For $M \subseteq [1, n]$ we set $V(M) = \langle e_\nu : \nu \in M \rangle_K$. The subscripts v_i and $\mu(x)_{\nu,j}$ below refer to the respective coordinates.

Lemma 5.2. *Let (u, μ, v) be a linear representation of rank n with coefficients in R. Suppose that there exists $\mathcal{S} \subseteq \mathcal{P}([1, n])$ such that $[1, n] = \bigcup_{M\in\mathcal{S}} M$, and all of the following conditions are satisfied.*

(1) *There exists an $M \in \mathcal{S}$ with $u \in V(M)$.*
(2) *For every $M \in \mathcal{S}$ and $x \in X$, there exists $N \in \mathcal{S}$ with $V(M)\mu(x) \subseteq V(N)$.*
(3) *For every $M \in \mathcal{S}$, $x \in X$, and $j \in [1, n]$, there exists at most one $\nu \in M$ with $\mu(x)_{\nu,j} \neq 0$.*
(4) *For every $M \in \mathcal{S}$ there exists at most one $\nu \in M$ with $v_\nu \neq 0$.*

Then the weighted automaton $\mathcal{A}$ associated to the linear representation (u, μ, v) is unambiguous.

Proof. Following the construction of the associated automaton $\mathcal{A}$, the conditions above translate directly into the ones of Lemma 5.1. □

We are now ready to prove the main implication over a field.

Proposition 5.3. *Every rational Pólya series over K is recognized by an unambiguous weighted automaton (with weights in K).*

Proof. Let S be a rational Pólya series, and let (u, μ, v) be a minimal linear representation of S, chosen as in Lemma 3.11. Hence $\Omega \coloneqq u\mu(X^*) \subseteq G_0^{1\times n}$ and $\Omega v \subseteq G_0$. If $n = 0$, then $S = 0$, and S is recognized by the trivial automaton with empty set of states. To avoid this corner case, from now on assume $n \geq 1$. Then $v = e_1 \in K^{n\times 1}$ is the first standard basis vector by Lemma 3.11.

We may replace K by the field generated by all coefficients in u, v, and $\mu(x)$ for $x \in X$. Thus, we may without restriction assume that K is finitely generated over its prime field. We may also assume that K is infinite; otherwise Proposition 3.14 implies the even stronger claim that S is recognized by a deterministic automaton.

Applying Lemma 3.13, we can assume $K^{1\times n} = W_1 \oplus \cdots \oplus W_k$ with $W_i = \langle e_{m_1+\cdots+m_{i-1}+1}, \ldots, e_{m_1+\cdots+m_i}\rangle_K$ and $\mathcal{Z}(\overline{\Omega}) = \{W_1, \ldots, W_k\}$. Without restriction $u \in W_1$. Note that, taking $\Gamma = G_0$ in Lemma 3.13, also the properties $\Omega \subseteq G_0^{1\times n}$ and $\Omega v \subseteq G_0$ are preserved by this change of linear representation.

For $i \in [1, k]$ let

$$M_i = [m_1 + \cdots + m_{i-1} + 1, m_1 + \cdots + m_i],$$

so that $\{ e_\nu : \nu \in M_i \}$ is a basis for W_i. We show that the linear representation (u, μ, v), with

$$\mathcal{S} = \{ M_i : i \in [1, k] \}$$

satisfies the conditions of Lemma 5.2, from which the claim will follow.

Indeed, (1) holds for M_1 since $u \in W_1$. Since $\mu(x)$ is continuous, statement (3) of Lemma 3.5 implies that for every $i \in [1, k]$ and $x \in X$, there exists $j \in [1, k]$ such that $W_i\mu(x) \subseteq W_j$. This implies (2).

Let $i \in [1, k]$, let $x \in X$, and let $j \in [1, n]$. Let $\varphi\colon W_i \to K$ be defined by $\varphi(a) = a\mu(x)e_j^T$. By (5) of Lemma 3.5, the set $\Omega \cap W_i$ is dense in W_i. Moreover $\Omega \cap W_i \subseteq \sum_{\nu\in M_i} G_0 e_\nu$ since $\Omega \subseteq G_0^{1\times n}$. Finally, if $a \in \Omega$, then $a\mu(x) \in \Omega \subseteq G_0^{1\times n}$, so that $\varphi(\Omega \cap W_i) \subseteq G_0$. If K has characteristic 0, we can thus apply Lemma 4.2 to the vector space W_i, its dense subset $\Omega \cap W_i$, and the homomorphism $\varphi\colon W_i \to K$. We conclude that there exists at most one $\nu \in M_i$ with $0 \neq \varphi(e_\nu) = e_\nu\mu(x)e_j^T$. Since $e_\nu\mu(x)e_j^T$ is the (ν, j)-entry of the matrix $\mu(x)$, there is at most one $\nu \in M_i$ with $\mu(x)_{\nu,j} \neq 0$. Thus (3) holds in characteristic 0.

If K has positive characteristic, we apply Lemma 4.5 instead of Lemma 4.2; the additional condition in this lemma is satisfied by Lemma 4.9. This shows (3) in positive characteristic.

Similarly, applying Lemma 4.2 in characteristic 0 (respectively Lemma 4.5 together with Lemma 4.9 in positive characteristic) to the map $W_i \to K$ given by $a \mapsto av$, we find that there exists at most one $\nu \in M_i$ with $v_\nu \neq 0$, implying (4). □

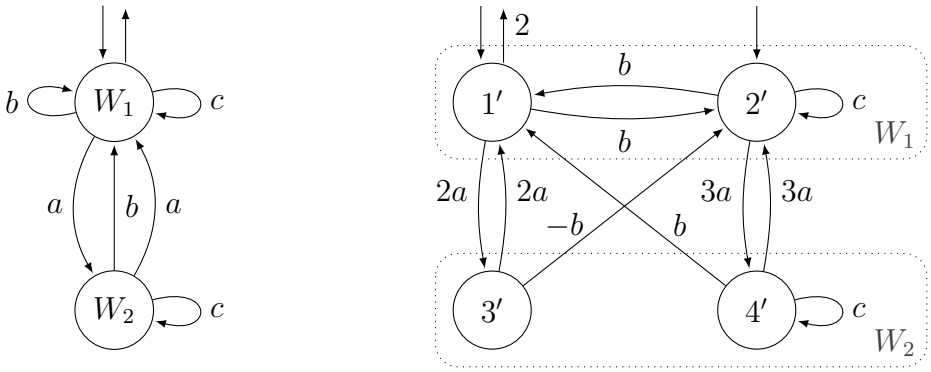

Figure 2. (Example 5.4) *Left:* Our choice of transitions between irreducible components can be depicted as a deterministic automaton. *Right:* An unambiguous weighted automaton recognizing the same series as in Example 3.7 and Figure 1.

Example 5.4. (Continuation of Example 3.7.) We illustrate the construction of the previous proof using the linear representation from Example 3.7. The linear hull decomposes as $\overline{\Omega} = W_1 \cup W_2$ with $W_1 = \langle e_1 + e_2, e_3 \rangle$ and $W_2 = \langle e_1 - e_2, e_3 \rangle$. Now $W_1 \oplus W_2 \cong K^{1\times 4} = \langle e_1', \dots, e_4' \rangle$, where we fix the embeddings $W_1 \hookrightarrow K^{1\times 4}$, $e_1 + e_2 \mapsto e_1'$ and $e_3 \mapsto e_2'$, as well as $W_2 \hookrightarrow K^{1\times 4}$, $e_1 - e_2 \mapsto e_3'$ and $e_3 \mapsto e_4'$. For every $x \in \{a, b, c\}$ we need to choose $f_x \colon \{W_1, W_2\} \to \{W_1, W_2\}$ such that $W_i \mu(x) \subseteq f_x(W_i)$ for $i \in \{1, 2\}$. For a, b there is a unique such choice. Since $W_i \mu(c) \subseteq W_1 \cap W_2$, there are several choices for c and we pick $f_c(W_i) = W_i$ for $i \in \{1, 2\}$. This choice fixes a deterministic automaton describing the transitions between irreducible components (left side of Figure 2).

The newly constructed linear representation on $K^{1\times 4}$ is given by (u', μ', v') with $u' = (1, 1, 0, 0)$, with $v' = (2, 0, 0, 0)^T$, and with

$$\mu'(a) = \left(\begin{array}{cc:cc} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ \hdashline 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{array}\right), \quad \mu'(b) = \left(\begin{array}{cc:cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hdashline 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right), \quad \mu'(c) = \left(\begin{array}{cc:cc} 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{array}\right).$$

Here the block structure is determined by the choice of transitions between irreducible components, e.g., a different choice of f_c would yield a different matrix $\mu'(c)$. The resulting automaton is depicted in the right side of Figure 2.

Remark 5.5. Let X, Y be finite sets. A rational function is a function $f \colon X^* \to Y^*$ whose graph is a rational subset of $X^* \times Y^*$. By the *Decomposition Theorem* of Elgot–Mezei every such rational function is a composition of a (pure) sequential function

with a (pure) co-sequential function [Sak09, Chapter V.2] (varying terminology is used, see for instance [AL79]; we follow Sakarovitch).

One of the proofs of this theorem produces, through the use of the Schützenberger covering, a *semi-monomial* linear representation. This is the same type of block-matrix structure we have obtained here. Consequently we can obtain an analogous decomposition of the weighted automaton: every Pólya series is recognized by a weighted finite automaton that is a composition of a sequential function followed by a co-deterministic weighted automaton. (A weighted automaton is *co-deterministic* if there is unique final state and for every $x \in X$ and $q \in Q$ there is at most one $p \in Q$ with $E(p, x, q) \neq 0$.) As the construction is very similar to the one in [Sak09, Chapter V.2.2], we omit the details.

If X is a singleton, then series recognized by an unambiguous weighted automaton have a particularly simple shape. In this way we will recover the full univariate result of Pólya, Benzaghou, and Bézivin [Pó21, Ben70, Bé87].

Proposition 5.6. *Suppose $X = \{x\}$ consists of a single element, let $\mathcal{A}$ be an unambiguous weighted automaton with weights in R, and let $S \in R\langle\langle X\rangle\rangle$ be the series it recognizes. Then there exist a finite set $F \subseteq \mathbb{Z}_{\geq 0}$, an element $d \in \mathbb{Z}_{\geq 0}$, and for each $r \in [0, d-1]$ elements $a_r \in R$ and $b_r \in R \setminus \{0\}$ such that*

$$S(x^{kd+r}) = a_r b_r^k \qquad \text{for all } k \in \mathbb{Z}_{\geq 0} \text{ and } r \in [0, d-1] \text{ with } xd + r \notin F.$$

Proof. We may assume that $\mathcal{A}$ is trim. Then, for any two states p, q and any $n \geq 0$ there exists at most one path from p to q labeled by x^n.

Let m be the maximal length of an acyclic path in $\mathcal{A}$, that is, a path that does not visit any vertex twice. Then any cycle, that is, a path whose only repeated vertices are the first and the last one, has length bounded by $m + 1$. Let d be a common multiple of all lengths of cycles in $\mathcal{A}$, e.g., $d = (m+1)!$. Let $r \in \{0, \dots, d-1\}$ and let us consider the claim for $S(x^{kd+r})$ with $k \geq 0$. If $S(x^{kd+r}) = 0$ for all but finitely many $d \geq 1$, the claim holds with $a_r = 0$. Otherwise, let $k_0 \geq 1$ with $k_0 d + r > m$ and $S(x^{k_0 d+r}) \neq 0$. The (unique) accepting path labeled by $x^{k_0 d+r}$ must contain a cycle, and so is of the form pcq with c a cycle of length l dividing d, and p, q paths. Let $e \in \mathbb{N}$ with $d = le$, let $a_r := S(x^{pq})$ and let b_r be the product of the weights along c^e. For all $n \geq 0$, the path $pc(c^e)^n q$ is the unique accepting path for $x^{(k_0+n)d+r}$. Hence $S(x^{(k_0+n)d+r}) = a_r b_r^n$ for all $n \geq 0$. Thus $S(x^{kd+r}) = (a_r b_r^{-k_0}) b_r^k$ for all $k \geq k_0$. □

6. Proof of (b) ⇔ (c)

The following proof very closely follows [Lot02, Proposition 1.3.5], where the same result is proved for deterministic automata without weights. A language $\mathcal{L} \subseteq X^*$ is a code if the elements of $\mathcal{L}$ are a basis of a free submonoid of X^*.

Proposition 6.1 (Reutenauer)**.** *If a rational series $S \in R\langle\langle X\rangle\rangle$ is recognized by an unambiguous weighted automaton with weights in R, then S is unambiguous over R.*

Proof. Let $\mathcal{A}$ be an unambiguous weighted automaton that recognizes S. We may without restriction assume that $\mathcal{A}$ is trim. Then, for any two states p, q and any word $w \in X^*$ there exists at most one path from p to q labeled by w.

For states p, $q \in Q$ and a set $P \subseteq Q$ define

$$S_{p,P,q} = \sum_{\substack{p_1,\dots,p_{l-1}\in P \\ p=p_0,\, p_l=q \\ a_1,\dots,a_l\in X \\ l\geq 1}} E(p_0,a_1,p_1)\cdots E(p_{l-1},a_l,p_l)a_1\cdots a_l.$$

In words, the sum is taken over all non-empty paths from p to q with the property that all states strictly in-between are in P. Since $\mathcal{A}$ is unambiguous, the words in $\operatorname{supp}(S_{p,P,q})$ are in bijective correspondence with non-empty paths from p to q.

Then

$$S = \sum_{p,q\in Q} I(p)S_{p,Q,q}T(q) + \sum_{p\in Q} I(p)T(p),$$

and the finite sum on the left is unambiguous because $\mathcal{A}$ is unambiguous.

It suffices to show that each $S_{p,P,q}$ is unambiguous, and we do so by induction on $|P|$. If $P = \emptyset$, then $S_{p,P,q}$ is a polynomial and hence unambiguous. If $r \notin P$, then

$$S_{p,P\cup\{r\},q} = S_{p,P,q} + S_{p,P,r}S^*_{r,P,r}S_{r,P,q}.$$

Note that $\operatorname{supp}(S_{r,P,r})$ consist of the words labeling first returns of r, that is, non-empty paths starting and ending at r that do not pass through r in-between. Using that $\mathcal{A}$ is unambiguous, it is easily seen that the words in $\operatorname{supp}(S_{r,P,r})$ are a code. Hence $S^*_{r,P,r}$ is unambiguous. Similarly, we see that the products and the sum are unambiguous, by looking at when a path passes through r. □

The converse of the previous implication is also easy to see.

Lemma 6.2. *If $S \in R\langle\langle X\rangle\rangle$ is unambiguous rational, then there exists an unambiguous weighted automaton with weights in R that recognizes S.*

Proof. A suitable weighted automaton can inductively be constructed from an unambiguous rational decomposition of S. □

7. Proof of (c) ⇒ (d)

A clever proof of (c) ⇒ (d) of Theorem 1.2 is given by Reutenauer in the proof of [Reu79, Proposition 4, (iii) ⇒ (ii)]. We opt to give an alternative, somewhat longer but very straightforward, proof of the same result.

Lemma 7.1. *Let A, $B \in \mathbb{Z}\langle\langle X\rangle\rangle$, and let $\mathcal{L}$, $\mathcal{K}$ be rational languages with* $\operatorname{supp}(A) \subseteq \mathcal{L}$ *and* $\operatorname{supp}(B) \subseteq \mathcal{K}$.

(1) *Suppose* $\mathcal{K} \cap \mathcal{L} = \emptyset$. *Then* $C = A + B$ *is a rational series with* $\operatorname{supp}(C) \subseteq \mathcal{K} \cup \mathcal{L}$ *and*

$$C(w) = \begin{cases} A(w) & \text{if } w \in \mathcal{L}, \\ B(w) & \text{if } w \in \mathcal{K}. \end{cases}$$

(2) *Suppose* $\mathcal{L}\mathcal{K}$ *is unambiguous. Then* $C = A\mathbb{1}_{\mathcal{K}} + \mathbb{1}_{\mathcal{L}}B$ *is a rational series with* $\operatorname{supp}(C) \subseteq \mathcal{L}\mathcal{K}$. *For* $w = uv$ *with* $u \in \mathcal{L}$, $v \in \mathcal{K}$,

$$C(w) = A(u) + B(v).$$

(3) *Suppose that* $\mathcal{L}$ *is a code. Then*

$$C = (1 - \mathbb{1}_{\mathcal{L}^*}A)\Big((\mathbb{1}_{\mathcal{L}} + A)^* - \mathbb{1}_{\mathcal{L}^*}\Big)$$

is a rational series with $\operatorname{supp}(C) \subseteq \mathcal{L}^*$. *For* $w = w_1 \cdots w_l$ *with* $w_1, \ldots, w_l \in \mathcal{L}$,

$$C(w) = A(w_1) + \cdots + A(w_l).$$

Proof. Throughout, we use that the characteristic series $\mathbb{1}_{\mathcal{L}} \in \mathbb{Z}\langle\langle X \rangle\rangle$ of a rational language $\mathcal{L}$ is rational.

(1) Clear.

(2) For $w \in X^*$ we have $A\mathbb{1}_{\mathcal{K}}(w) = \sum_{w=uv} A(u)\mathbb{1}_{\mathcal{K}}(v)$. A term $A(u)\mathbb{1}_{\mathcal{K}}(v)$ is nonzero if and only if $u \in \operatorname{supp}(A) \subseteq \mathcal{L}$ and $v \in \mathcal{K}$. Since $\mathcal{L}\mathcal{K}$ is unambiguous there is at most one such term. Thus $A\mathbb{1}_{\mathcal{K}}(w) = A(u)$ if $w \in \mathcal{L}\mathcal{K}$ with $w = uv$ where $u \in \mathcal{L}$, $v \in \mathcal{K}$, and $A\mathbb{1}_{\mathcal{K}}(w) = 0$ if $w \notin \mathcal{L}\mathcal{K}$. An analogous claim holds for $\mathbb{1}_{\mathcal{L}}B$.

(3) For $w \in \mathcal{L}^*$ there are uniquely determined $w_1, \ldots, w_l \in \mathcal{L}$ with $w = w_1 \cdots w_l$. Then

$$(\mathbb{1}_{\mathcal{L}} + A)^*(w) = \sum_{k=0}^{l} \sum_{1 \le i_1 < \cdots < i_k \le l} A(w_{i_1}) \cdots A(w_{i_k}).$$

For $D = (\mathbb{1}_{\mathcal{L}} + A)^* - \mathbb{1}_{\mathcal{L}^*}$ we obtain an analogous sum with $k \in [1, l]$.

Now, $\mathbb{1}_{\mathcal{L}^*}A(1) = 0$ since $1 \notin \mathcal{L}$ and, for $j \ge 1$,

$$\mathbb{1}_{\mathcal{L}^*}A(w_1 \cdots w_j) = \sum_{i=0}^{j} \mathbb{1}_{\mathcal{L}^*}(w_1 \cdots w_i)A(w_{i+1} \cdots w_j) = A(w_j).$$

Therefore

$$\begin{aligned} \mathbb{1}_{\mathcal{L}^*}AD(w) &= \sum_{j=1}^{l} A(w_j)D(w_{j+1} \cdots w_l) = \\ &= \sum_{j=1}^{l} A(w_j) \sum_{k=1}^{l-j} \sum_{j+1 \le i_1 < \cdots < i_k \le l} A(w_{i_1}) \cdots A(w_{i_k}) \\ &= \sum_{k=2}^{l} \sum_{1 \le i_1 < \cdots < i_k \le l} A(w_{i_1}) \ldots A(w_{i_k}). \end{aligned}$$

Thus $(1 - \mathbb{1}_{\mathcal{L}^*}A)D(w) = \sum_{k=1}^{l} A(w_k)$. □

A series $a \in \mathbb{Z}\langle\langle X\rangle\rangle$ is *linearly bounded* if there exists $C \geq 0$ such that $|a(w)| \leq C|w|$ for all nonempty words $w \in X^*$.

Proposition 7.2. *Let $S \in R\langle\langle X\rangle\rangle$ be an unambiguous rational series. Then there exist $\lambda_1, \ldots, \lambda_k \in R \setminus \{0\}$, linearly bounded rational series $a_1, \ldots, a_k \in \mathbb{Z}\langle\langle X^*\rangle\rangle$, and a rational language $\mathcal{L}$ such that $\operatorname{supp}(a_i) \subseteq \mathcal{L}$ for all $i \in [1,k]$ and*

$$S(w) = \begin{cases} \lambda_1^{a_1(w)} \cdots \lambda_k^{a_k(w)} & \text{if } w \in \mathcal{L}, \\ 0 & \text{if } w \notin \mathcal{L}. \end{cases}$$

Proof. The claim is trivially true if S is a polynomial. We show that the property is preserved under unambiguous $+$, $\cdot$, and * constructions.

Let S, T be rational series such that there exist $\lambda_1, \ldots, \lambda_k \in R \setminus \{0\}$, linearly bounded rational series $a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{Z}\langle\langle X\rangle\rangle$, and rational languages $\mathcal{L}$, $\mathcal{K}$ such that $\operatorname{supp}(a_i) \subseteq \mathcal{L}$, $\operatorname{supp}(b_i) \subseteq \mathcal{K}$, and

$$S(w) = \begin{cases} \lambda_1^{a_1(w)} \cdots \lambda_k^{a_k(w)} & \text{if } w \in \mathcal{L}, \\ 0 & \text{if } w \notin \mathcal{L}; \end{cases}$$

$$T(w) = \begin{cases} \lambda_1^{b_1(w)} \cdots \lambda_k^{b_k(w)} & \text{if } w \in \mathcal{K}, \\ 0 & \text{if } w \notin \mathcal{K}. \end{cases}$$

(We can assume that the λ_i's are the same, as we can always extend the set of constants, and set $a_i = 0$, respectively, $b_i = 0$, if λ_i does not appear in the expression for S, respectively, T.)

We first consider $S + T$ with $\mathcal{L} \cap \mathcal{K} = \emptyset$. Then $\mathcal{L} \cup \mathcal{K}$ is a rational language, $S + T(w) = 0$ if $w \notin \mathcal{L} \cup \mathcal{K}$, and

$$(S+T)(w) = S(w) + T(w) = \begin{cases} S(w) = \lambda_1^{a_1(w)} \cdots \lambda_k^{a_k(w)} & \text{if } w \in \mathcal{L}, \\ T(w) = \lambda_1^{b_1(w)} \cdots \lambda_k^{b_k(w)} & \text{if } w \in \mathcal{K}. \end{cases}$$

Since $\mathcal{L} \cap \mathcal{K} = \emptyset$, we get $(S+T)(w) = \lambda_1^{a_1(w)+b_1(w)} \cdots \lambda_k^{a_k(w)+b_k(w)}$ for all $w \in \mathcal{L} \cup \mathcal{K}$. Clearly $a_i + b_i$ is a linearly bounded rational series, and $\operatorname{supp}(a_i + b_i) \subseteq \mathcal{L} \cup \mathcal{K}$.

Now consider ST with $\mathcal{L}\mathcal{K}$ unambiguous. Then $\mathcal{L}\mathcal{K}$ is a rational language, and for $w = uv$ with $u \in \mathcal{L}$, $v \in \mathcal{K}$,

$$(ST)(w) = S(u)T(v) = \lambda_1^{a_1(u)+b_1(v)} \cdots \lambda_k^{a_k(u)+b_k(v)}.$$

Define series c_i by $c_i(uv) = a_i(u) + b_i(v)$ if $w = uv \in \mathcal{L}\mathcal{K}$ with $u \in \mathcal{L}$, $v \in \mathcal{K}$, and $c_i(w) = 0$ for $w \notin \mathcal{L}\mathcal{K}$. Clearly c_i is linearly bounded. Since $\mathcal{L}\mathcal{K}$ is unambiguous, Lemma 7.1 implies that c_i is rational.

Now suppose that $\mathcal{L} = \operatorname{supp}(S)$ is a code and consider S^*. Then $\mathcal{L}^*$ is a rational language. For $w \in \mathcal{L}^*$ there exist uniquely determined $w_1, \ldots, w_l \in \mathcal{L}$ with $w = w_1 \cdots w_l$. We have

$$S^*(w) = S^*(w_1 \cdots w_l) = S(w_1) \cdots S(w_l) = \lambda_1^{a_1(w_1)+\cdots+a_1(w_l)} \cdots \lambda_k^{a_k(w_1)+\cdots+a_k(w_l)}.$$

Define

$$c_i(w_1 \cdots w_l) = a_i(w_1) + \cdots + a_i(w_l).$$

and $c_i(w) = 0$ if $w \notin \mathcal{L}^*$. Then c_i is linearly bounded and, by Lemma 7.1, again rational. □

8. Hadamard sub-invertibility

It is known that every unambiguous rational series is Hadamard sub-invertible, and every Hadamard sub-invertible rational series is a Pólya series [BR11, Exercise 3.1 of Chapter 6]. This is particularly easy for $K = \mathbb{Q}$. For arbitrary fields, the same argument works but requires a theorem of Roquette; hence we give the proof in full.

Lemma 8.1. *Every unambiguous rational series is Hadamard sub-invertible.*

Proof. Every noncommutative polynomial is Hadamard sub-invertible. If S, T are Hadamard sub-invertible, then it is easy to see that the *unambiguous* sums, products, and star operations preserve this property. □

Lemma 8.2. *Every Hadamard sub-invertible series is a Pólya series.*

Proof. Let $S \in R\langle\langle X \rangle\rangle \subseteq K\langle\langle X \rangle\rangle$ be a Hadamard sub-invertible rational series. If $\operatorname{char} K > 0$, let k be the (finite) prime field of K; if $\operatorname{char} K = 0$, let $k = \mathbb{Z}$. It is immediate from the definition of a rational series that there exists a finitely generated k-subalgebra A of K containing all coefficients of S. Since $\sum_{w \in \operatorname{supp}(S)} S(w)^{-1} w$ is also rational, we may moreover assume that A also contains all $S(w)^{-1}$ with $S(w) \neq 0$. Hence, the nonzero coefficients of S are contained in $A^\times$. The group $A^\times$ is finitely generated by a theorem of Roquette [Lan83, Corollary 7.5] in characteristic 0, and a slightly easier argument in positive characteristic [Lan83, Corollary 7.3]. □

9. Putting it all together

The proofs of Theorems 1.1 and 1.2 in the case where $R = K$ is a field are now a formality. To obtain the more general result for domains, we first need to extend Proposition 5.3 to completely integrally closed domains.

Proposition 9.1. *Let R be a completely integrally closed domain. Every rational Pólya series in $R\langle\langle X \rangle\rangle$ is unambiguous rational (over R).*

Proof. Let K be the quotient field of R. If $S \in R\langle\langle X\rangle\rangle$ is a rational Pólya series, Proposition 5.3 together with Proposition 6.1 shows that S is unambiguous rational as a series over K. What remains to be shown is that we can obtain an unambiguous rational decomposition of S in such a way that all the component series have their coefficients in R.

For a domain A, let $U_0(A) = A\langle X\rangle$. For $k \geq 1$, inductively define $U_k(A)$ as the set of all series obtained as unambiguous sums (that is, having pairwise disjoint support) of series of the form

$$S = \lambda w_0 S_1^* w_1 S_2^* w_2 \cdots S_l^* w_l, \tag{7}$$

with $l \geq 0$, with $0 \neq \lambda \in A$, with $w_0, \ldots, w_l \in X^*$, with $S_1, \ldots, S_l \in \bigcup_{k'=0}^{k-1} U_{k'}(A) \setminus \{0\}$ and satisfying the following conditions:

(1) The operations S_i^* are unambiguous, that is $\mathrm{supp}(S_i)$ is a code, for all $i \in [1, l]$, and
(2) the products in (7) are unambiguous, that is, for every word $w \in \mathrm{supp}(S)$ and every $i \in [1, l]$, there exists a unique word $w_i' \in \mathrm{supp}(S_i^*)$ such that $w = w_0 w_1' w_1 \cdots w_l' w_l$.

Let $\mathcal{U}(A)$ denote the set of all unambiguous sums of elements of $\bigcup_{k\geq 0} U_k(A)$. By construction, each series in $\mathcal{U}(A)$ is unambiguous rational. Note that $A\langle X\rangle \subseteq \mathcal{U}(A)$ and that $\mathcal{U}(A)$ is closed under unambiguous sums and the unambiguous star operation. Moreover, by distributivity we see that $\mathcal{U}(A)$ is closed under unambiguous products. Thus $\mathcal{U}(A)$ is the set of all unambiguous rational series over A.

To conclude the proof of the proposition, we show $U_k(K) \cap R\langle\langle X\rangle\rangle = U_k(R)$ by induction on k. For $k = 0$ the claim is trivial because $K\langle X\rangle \cap R\langle\langle X\rangle\rangle = R\langle X\rangle$. Suppose now $k \geq 1$ and the claim has been established for $k' < k$. Let $S \in U_k(K) \cap R\langle\langle X\rangle\rangle$. Decomposing along unambiguous sums, it suffices to consider S as in (7). Let $i \in [1, l]$ and let a_i be a nonzero coefficient of S_i. For each $j \in [1, l] \setminus \{i\}$ pick an arbitrary nonzero coefficient a_j of S_j. Then $\lambda a_1 \cdots a_{i-1} a_i^m a_{i+1} \cdots a_l$ is a coefficient of S for every $m \geq 0$, and hence contained in R. Since R is completely integrally closed, we conclude $a_i \in R$. Thus $S_i \in U_{k'}(K) \cap R\langle\langle X\rangle\rangle = U_{k'}(R)$ for some $k' < k$. Since λ appears as coefficient of $w_0 w_1 \cdots w_l$ (taking the empty word in each S_j^*), we also must have $\lambda \in R$. Thus $S \in U_k(R)$. □

Proof of Theorem 1.2. The equivalence (b) ⇔ (c) is shown in Section 6. For $R = K$ a field, the implication (a) ⇒ (b) follows by Proposition 5.3. More generally, for R a completely integrally closed domain, the implication (a) ⇒ (c) is shown in Proposition 9.1. Next, the implication (c) ⇒ (d) follows from Proposition 7.2. The implication (d) ⇒ (a) is trivial.

Finally, the implication (c) ⇒ (e) follows from Lemma 8.1, and (e) ⇒ (a) holds by Lemma 8.2. □

Remark 9.2. Every completely integrally closed domain is integrally closed, and a noetherian domain is completely integrally closed if and only if it is integrally closed. Krull domains, and thus in particular factorial domains such as $\mathbb{Z}$, are completely integrally closed. The ring of all algebraic integers is a non-noetherian, completely integrally closed domain.

To illustrate that some condition needs to be imposed on the domain R, the following example gives a Pólya series over an integrally closed, but not completely integrally closed, domain R that is *not* unambiguous rational over R.

Example 9.3. Let $R = \mathbb{Z}[y^i z : i \geq 1] \subseteq \mathbb{Z}[y, z]$ and $S = \sum_{i\geq 1} y^i z x^i \in R[\![x]\!]$. Then S is a Pólya series, and indeed, over $\mathbb{Z}[y, z]$ is unambiguous rational as $S = zyx(yx)^*$.

However, suppose S were unambiguous rational over R. Then Proposition 5.6 applies. In particular, there exist $d > 0$, $r \geq 0$, and f, $g \in R$ such that $y^{dk+r} z = f g^k$ for every $k \geq 0$. However, in R the element $y^i z$ is an irreducible element for each $i \geq 1$, so we must have $g = \pm 1$ and $y^{dk+r} z = \pm f$ for all $k \geq 0$, a contradiction.

The ring R is not completely integrally closed, because $(yz)y^i \in R$ for all $i \geq 0$, but $y \notin R$. However it is integrally closed: Note that $R \subseteq \mathbb{Z}[yz, y]$ and the latter ring is factorial, hence integrally closed, so that the integral closure of R must be contained in $\mathbb{Z}[yz, y]$. Let $a \in K$ be integral over R. Then $a \in \mathbb{Z}[yz, y]$ and hence $a = a' + a''$ with $a' \in \mathbb{Z}[y]$ and $a'' \in R$. Then a' is integral over R. Hence there exist $m \geq 1$ and $b_0, \ldots, b_{m-1} \in R$ such that

$$(a')^m + b_{m-1}(a')^{m-1} + \cdots + b_0 = 0.$$

Taking this equation modulo z, we see that $a' \in \mathbb{Z}[y]$ is integral over $\mathbb{Z}$, forcing $a' \in \mathbb{Z}$. Hence $a \in R$.

The previous example is no coincidence; more generally the following holds.

Lemma 9.4. *Suppose R is integrally closed but not completely integrally closed. Then there exists a rational Pólya series $S \in R\langle\langle X \rangle\rangle$ such that S is* not *unambiguous rational.*

Proof. Since R is not integrally closed, there exists $0 \neq a \in R$ and $b \in K \setminus R$ such that $ab^i \in R$ for all $i \geq 0$. Let $S = \sum_{i\geq 0} ab^i x^i$. We show that S is not unambiguous rational. Suppose to the contrary that it is. Then it is recognized by an unambiguous weighted automaton with weights in R, and Proposition 5.6 shows that there exist $m > 0$, $n \geq 0$ and c, $d \in R \setminus \{0\}$ such that $ab^{mk+n} = cd^k$ for all $k \geq 0$. Then $(b^m d^{-1})^k = ca^{-1}b^{-n}$ is constant for all $k \geq 0$. Substituting $k = 0$ and $k = 1$, we see $b^m d^{-1} = 1$. Thus b is a root of the polynomial $t^m - d \in R[t]$, hence integral over R. By hypothesis $b \in R$, a contradiction. □

Finally, we deduce the classical theorem for the univariate theorem as a special case of our result.

Proof of Theorem 1.1. The series S is rational by [BR11, Proposition 6.1.1]. By Theorem 1.2, there exists an unambiguous weighted automaton $\mathcal{A}$ on the alphabet $X = \{x\}$ recognizing S. The claim now follows from Proposition 5.6. □

10. Determinizability

We finish by showing (a) ⇒ (b) of Theorem 1.3. The approach in this section is inspired by [Moh97, Theorem 9]. Mohri shows that a deterministic weighted automaton over the tropical semiring $(\mathbb{R}, \max, +)$ has bounded variation; and that an unambiguous weighted automaton with bounded variation is determinizable.

Definition 10.1. Let $(G, \cdot)$ be a group. A map $\ell\colon G \to \mathbb{R}_{\geq 0}$ is a *length function* if
(1) $\ell(1_G) = 0$.
(2) $\ell(gh) \leq \ell(g) + \ell(h)$ for all g, $h \in G$.
(3) $\ell(g) = \ell(g^{-1})$ for all $g \in G$.

If K has an absolute value $|\cdot|$, and $G \leq K^\times$, then $\ell(g) = |\log(|g|)|$ defines a length function. On $(\mathbb{Z}^r, +)$ we have a length function $(a_1, \ldots, a_r) \mapsto |a_1| + \cdots + |a_r|$. This induces a length function ℓ on any finitely generated free abelian group, since $G/G_{\mathrm{tor}} \cong \mathbb{Z}^r$. This length function satisfies

$$|\{\, g \in G : \ell(g) \leq C \,\}| < \infty \quad \text{for all } C \geq 0. \tag{$*$}$$

There is a metric $\mathsf{d}\colon X^* \times X^* \to \mathbb{Z}_{\geq 0}$, given by

$$\mathsf{d}(u, v) = |u| + |v| - 2|\mathrm{lgcd}(u, v)|,$$

where $\mathrm{lgcd}(u, v)$ is the longest common prefix of u and v.

Definition 10.2. Let G be a group and $\ell\colon G \to \mathbb{R}_{\geq 0}$ a length function. A function $f\colon X^* \to G_0$ has *bounded ℓ-variation* if, for every $c \geq 0$, there exists $C \geq 0$ such that for all u, v with $f(u) \neq 0$ and $f(v) \neq 0$,

$$\mathsf{d}(u, v) \leq c \quad \text{implies} \quad \ell(f(u)f(v)^{-1}) \leq C.$$

Lemma 10.3. *Let $\mathcal{A}$ be a deterministic weighted automaton, and let S be the series recognized by $\mathcal{A}$. Let $G \leq K^\times$ be such that*

- *$S(w) \in G_0$ for $w \in X^*$.*
- *all edge and terminal weights of $\mathcal{A}$ are contained in G_0.*

If $\ell\colon G \to \mathbb{R}_{\geq 0}$ is a length function, then S has bounded ℓ-variation.

Proof. Define

$$C = \max\Big\{\, \ell(E(p, a, q)),\, \ell(T(q)) : p, q \in Q, a \in X \text{ with } T(q) \neq 0, E(p, a, q) \neq 0 \,\Big\}.$$

Let w, $w' \in X^*$. Assume $S(w) \neq 0$ and $S(w') \neq 0$, as otherwise there is nothing to show.

We may suppose

$$w = u_1 \cdots u_k v_1 \cdots v_l \quad \text{and} \quad w' = u_1 \cdots u_k v'_1 \dots v'_m,$$

with u_i, v_i, $v'_i \in X$ and $\mathsf{d}(w, w') = l + m$. Let c and c' denote the accepting paths labeled by w and w'. Since $\mathcal{A}$ is deterministic, we must have

$$\begin{aligned} c &= (p_0, u_1, p_1) \cdots (p_{k-1}, u_k, p_k)(p_k, v_1, q_1) \cdots (q_{l-1}, v_l, q_l), \quad \text{and} \\ c' &= (p_0, u_1, p_1) \cdots (p_{k-1}, u_k, p_k)(p_k, v'_1, q'_1) \cdots (q'_{l-1}, v_m, q'_m), \end{aligned}$$

with states p_i, q_i, q'_i. For notational convenience, set $p_k = q_0 = q'_0$.

Now

$$\begin{aligned} \ell\left(\frac{S(w)}{S(w')}\right) &= \ell\left(\frac{I(p_0)E(c)T(q_l)}{I(p_0)E(c')T(q'_m)}\right) \\ &\le \sum_{i=1}^{l} \ell(E(q_{i-1}, v_i, q_i)) + \sum_{i=1}^{m} \ell(E(q'_{i-1}, v'_i, q'_i)) + \ell(T(q_l)) + \ell(T(q'_m)) \\ &\le (2 + m + l)C = (\mathsf{d}(w, w') + 2)C. \end{aligned}$$

If $w \neq w'$, then $\mathsf{d}(w, w') \ge 2$. Choosing $C' = C + 1$, we have $\ell(S(w)S(w')^{-1}) \le C'\mathsf{d}(w, w')$. □

Lemma 10.4. *Let S be a Pólya series. Let $\ell\colon G \to \mathbb{R}_{\ge 0}$ be a length function satisfying (∗). Suppose that S has bounded ℓ-variation. If (u, μ, v) is a minimal linear representation of S, its linear hull has dimension at most 1.*

Proof. Let $w_1, \dots, w_n \in X^*$ be such that the vectors $b_i = \mu(w_i)v$ form a basis of $K^{n\times 1}$. Let B be the matrix whose i-th column is b_i. Let $\Omega = u\mu(X^*) \subseteq K^{1\times n}$. We have to show that Ω can be covered by finitely many 1-dimensional vector spaces. Since B is invertible, it suffices to show the same for $\Omega B = \{\, u\mu(w)B : w \in X^* \,\}$.

Since S has bounded ℓ-variation, there exists $C \ge 0$ such that for every $w \in X^*$ and $i, j \in [1, n]$

$$\left|\ell\Big(S(ww_i)S(ww_j)^{-1}\Big)\right| \le C,$$

whenever $S(ww_i) \neq 0$ and $S(ww_j) \neq 0$. By our assumptions, this implies that the ratio $S(ww_i)S(ww_j)^{-1}$ can only take finitely many values (when $S(ww_j) \neq 0$). Noting that $S(ww_i) = u\mu(w)b_i$ is the i-th coordinate of $u\mu(w)B$, we conclude that we can cover ΩB by finitely many 1-dimensional vector spaces. □

Proof of Theorem 1.3. (a) ⇒ (b) Let $\mathcal{A}$ be a deterministic weighted automaton that recognizes S. There exists a finitely generated $G \le K^\times$ such that $S(w) \in G_0$ for all $w \in X^*$. This follows from Theorem 1.2, but can also be easily seen directly: Since $\mathcal{A}$ is deterministic, and in particular unambiguous, each coefficient is a product of

some weights of the automaton. By enlarging G if necessary, we can further assume that all edge and terminal weights are contained in G_0.

Let $\ell\colon G \to \mathbb{R}_{\geq 0}$ be a length function on G satisfying $(*)$. By Lemma 10.3, the series S has bounded ℓ-variation. Lemma 10.4 implies the claim.

(b) $\Rightarrow$ (a) By Proposition 3.14. □

References

[AL79] A. Arnold and M. Latteux. A new proof of two theorems about rational transductions. *Theoret. Comput. Sci.*, 8(2):261–263, 1979. doi:10.1016/0304-3975(79)90049-5.

[Ben70] B. Benzaghou. Algèbres de Hadamard. *Bull. Soc. Math. France*, 98:209–252, 1970. URL: http://www.numdam.org/item?id=BSMF_1970__98__209_0.

[BG06] E. Bombieri and W. Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006. doi:10.1017/CBO9780511542879.

[BGW00] A. L. Buchsbaum, R. Giancarlo, and J. R. Westbrook. On the determinization of weighted finite automata. *SIAM J. Comput.*, 30(5):1502–1531, 2000. doi:10.1137/S0097539798346676.

[Bou72] N. Bourbaki. *Elements of mathematics. Commutative algebra*. Hermann, Paris; Addison-Wesley Publishing Co., Reading, Mass., 1972. Translated from the French.

[BR11] J. Berstel and C. Reutenauer. *Noncommutative rational series with applications*, volume 137 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2011.

[Bé87] J.-P. Bézivin. Suites récurrentes linéaires en caractéristique non nulle. *Bull. Soc. Math. France*, 115(2):227–239, 1987.

[DM12] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing I. *Proc. Lond. Math. Soc. (3)*, 104(5):1045–1083, 2012. doi:10.1112/plms/pdr040.

[EG15] J.-H. Evertse and K. Győry. *Unit equations in Diophantine number theory*, volume 146 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015. doi:10.1017/CBO9781316160749.

[Eis95] D. Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry. doi:10.1007/978-1-4612-5350-1.

[Eve84] J.-H. Evertse. On sums of S-units and linear recurrences. *Compositio Math.*, 53(2):225–244, 1984.

[HS00] M. Hindry and J. H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. An introduction. doi:10.1007/978-1-4612-1210-2.

[Lan83] S. Lang. *Fundamentals of Diophantine geometry*. Springer-Verlag, New York, 1983. doi:10.1007/978-1-4757-1810-2.

[Lot02] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2002. doi:10.1017/CBO9781107326019.

[LS06] S. Lombardy and J. Sakarovitch. Sequential? *Theoret. Comput. Sci.*, 356(1-2):224–244, 2006. doi:10.1016/j.tcs.2006.01.028.

[Moh97] M. Mohri. Finite-state transducers in language and speech processing. *Comput. Linguist.*, 23(2):269–311, 1997.

[MR17] M. Mohri and M. D. Riley. A disambiguation algorithm for weighted automata. *Theoret. Comput. Sci.*, 679:53–68, 2017. doi:10.1016/j.tcs.2016.08.019.

[Pó21] G. Pólya. Arithmetische Eigenschaften der Reihenentwicklungen rationaler Funktionen. *J. Reine Angew. Math.*, 151:1–31, 1921. doi:10.1515/crll.1921.151.1.

[Reu79] C. Reutenauer. On Pólya series in noncommuting variables. In *Fundamentals of computation theory (Proc. Conf. Algebraic, Arith. and Categorical Methods in Comput. Theory, Berlin/Wendisch-Rietz, 1979)*, volume 2 of *Math. Res.*, pages 391–396. Akademie-Verlag, Berlin, 1979.

[Reu80] C. Reutenauer. *Séries rationelles et algèbres syntactiques*. Thèse de Doctorat d'etat, L'université Pierre et Marie Curie (Paris VI), 1980.

[Reu96] C. Reutenauer. A survey of noncommutative rational series. In *Formal power series and algebraic combinatorics (New Brunswick, NJ, 1994)*, volume 24 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 159–169. Amer. Math. Soc., Providence, RI, 1996.

[Sak09] J. Sakarovitch. *Elements of automata theory*. Cambridge University Press, Cambridge, 2009. Translated from the 2003 French original by Reuben Thomas. doi:10.1017/CBO9781139195218.

[Sta19] T. Stacks project authors. The stacks project. https://stacks.math.columbia.edu, 2019.

[vdPS82] A. J. van der Poorten and H. P. Schlickewei. The growth condition for recurrence sequences. *Macquarie University Math. Rep.*, 82–0041, 1982.

[vdPS91] A. J. van der Poorten and H. P. Schlickewei. Additive relations in fields. *J. Austral. Math. Soc. Ser. A*, 51(1):154–170, 1991.

Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1

Email address: jpbell@uwaterloo.ca

Email address: dsmertni@uwaterloo.ca