On recognizable and rational formal power series in partially commuting variables *

Manfred Droste[1] and Paul Gastin[2]

[1] Institut für Algebra, Technische Universität Dresden, D-01062 Dresden, droste@math.tu-dresden.de

[2] LITP, Université Paris 7, 2 place Jussieu, F-75251 Paris Cedex 05, Paul.Gastin@litp.ibp.fr

Abstract. We will describe the recognizable formal power series over arbitrary semirings and in partially commuting variables, i.e. over trace monoids. We prove that the recognizable series are certain rational power series, which can be constructed from the polynomials by using the operations sum, product and a restricted star which is applied only to series for which the elements in the support all have the same connected alphabet. The converse is true if the underlying semi-ring is commutative. Moreover, if in addition the semiring is idempotent then the same result holds with a star restricted to series for which the elements in the support have connected (possibly different) alphabets. It is shown that these assumptions over the semiring are necessary. This provides a joint generalization of Kleene's, Schützenberger's and Ochmański's theorems.

1 Introduction

In the theory of automata and formal languages, Kleene's foundational theorem on the coincidence of regular and rational languages in free monoids has been extended in many ways. Schützenberger [15] investigated formal power series over arbitrary semirings (e.g., like the natural numbers) and the free monoid, i.e. in noncommuting variables, and showed that the recognizable formal power series coincide with the rational ones. This was the starting point for a large amount of work on formal power series, cf. [14,9,2,8] for surveys. The concept of recognizable formal power series has also been defined for arbitrary monoids instead of the free monoid, but it was clear and has been stressed by several authors (cf., e.g. [14]) that in general then the recognizable and the rational series do not coincide.

On the other hand, Mazurkiewicz [10,11] introduced an important mathematical model for the behaviour of concurrent systems: trace monoids (or free partially commutative monoids), see also [3,1,4–6] for their well-developed theory. They are monoids whose generators are partially commutative. Again, their recognizable languages do not coincide with the rational ones, but by

* This research was partly carried out during a stay of the first author in Paris and another stay of the second author in Dresden.

Ochmański's theorem [12] they coincide with the c-rational languages where the iteration is restricted to connected languages.

It is the aim of this paper to investigate recognizable formal power series over trace monoids, thereby obtaining a generalization of both Schützenberger's and Ochmański's results.

We denote by $K\langle\langle\mathbb{M}\rangle\rangle$ the set of all formal power series over the semiring K and the free partially commutative monoid $\mathbb{M}$. It is known that in general the recognizable series in $K\langle\langle\mathbb{M}\rangle\rangle$ form a proper subclass of the rational ones. We therefore define the subclasses of *c-rational* and *mc-rational* series. We say that a series S is *connected*, if each element of its support is connected, and S is *mono-alphabetic*, if all elements of its support have the same set of generators. The c-rational series are obtained from the polynomials by allowing the operations sum, product, and star, but the latter applied only to proper and connected series. The mc-rational series are constructed in the same way, but using star only for series which are proper, mono-alphabetic and connected. In view of Ochmański's result, one might expect that the recognizable series in $K\langle\langle\mathbb{M}\rangle\rangle$ coincide with the c-rational ones. However, we will show that this fails in general even for the semiring $(\mathbb{N}, +, \times)$. Our main result is the following:

Theorem 1. *Let $\mathbb{M}$ be a trace monoid and K a semiring.*
(a) Each recognizable series in $K\langle\langle\mathbb{M}\rangle\rangle$ is mc-rational.
(b) If K is commutative, each mc-rational series in $K\langle\langle\mathbb{M}\rangle\rangle$ is recognizable.
(c) If K is commutative and idempotent, each c-rational series in $K\langle\langle\mathbb{M}\rangle\rangle$ is recognizable.

The fact that the recognizable series in $K\langle\langle\mathbb{M}\rangle\rangle$ are closed under the product operation was proved before already by Fliess [7], but only for very specific semirings K (strong Fatou semirings or the Boolean semiring). By Theorem 1(b), this holds for arbitrary commutative semirings, and we show by example that the commutativity of K is needed for this.

Theorem 1(b,c) is proved in section 3. There we also show that if the star S^* of a recognizable proper series S is connected, then it is also recognizable. This gives another closure property of the recognizable series under the star-operation. Part (a) of Theorem 1 is proved in section 4, and in section 5 we give examples and discuss the relationship with Schützenberger's and Ochmański's results. For lack of space, most proofs are not contained in this extended abstract.

It seems a very interesting research road to investigate which other results from the theory of formal power series over non-commuting variables can be extended to series over partially commuting variables, i.e. over trace monoids.

2 Background

Here we recall the necessary notation and background for formal power series and of trace theory. For more details, we refer the reader to [14,2,4,6].

Let M be any monoid and $K = (K, +, \cdot, 0, 1)$ any semiring, i.e., $(K, +, 0)$ is a commutative monoid, $(K, \cdot, 1)$ is a monoid, multiplication distributes over addition, and $0 \cdot x = x \cdot 0 = 0$ for each $x \in K$. If multiplication is commutative, we say that K is *commutative*. If the addition is idempotent, then the semiring is called *idempotent*. For instance, the semiring $(\mathbb{R} \cup \{\infty\}, min, +, \infty, 0)$ is both commutative and idempotent.

Mappings S from M into K are called *formal power series*. They are denoted as formal sums $S = \sum_{m \in M} (S, m).m$ where $(S, m) = S(m) \in K$. The set $supp(S) = \{m \in M \mid (S, m) \neq 0\}$ is called the *support* of S, and if it is finite, then S is called a *polynomial*. The collection of all formal power series is denoted by $K\langle\langle M\rangle\rangle$, and its subset of all polynomials by $K\langle M\rangle$. We consider elements of K also as polynomials in the natural way, having a non-zero entry only at $1 \in M$. If $L \subseteq M$, we define the *characteristic series* of L by $1_L = \sum_{m \in L} 1 \cdot m$.

Let $n \geq 1$ and $[n] = \{1, \ldots, n\}$. We let $K^{n \times n}$ be the monoid of all $(n \times n)$-matrices over K (with matrix multiplication as usual). A series $S \in K\langle\langle M\rangle\rangle$ is called *recognizable*, if there exists an integer $n \geq 1$, a monoid morphism $\mu : M \longrightarrow K^{n \times n}$ and vectors $\lambda \in K^{1 \times n}, \gamma \in K^{n \times 1}$ such that

$$(S, m) = \lambda \cdot (\mu m) \cdot \gamma = \sum_{i,j \in [n]} \lambda_i (\mu m)_{ij} \gamma_j$$

for each $m \in M$. In this case, the triple (λ, μ, γ) is called a *representation* of S, and we often shortly write $S = (\lambda, \mu, \gamma)$ to denote this. If $i, j \in [n]$, we also abbreviate $(\mu m)_{ij} =: \mu m_{ij}$. We let $K^{rec}\langle\langle M\rangle\rangle$ denote the set of all recognizable formal power series.

With componentwise addition, $K\langle\langle M\rangle\rangle$ becomes a commutative monoid. Now, the *(Cauchy) product* of two series S, S' in $K\langle\langle M\rangle\rangle$ is the series defined for $m \in M$ by $(S \cdot S', m) = \sum_{m = m_1 \cdot m_2} (S, m_1) \cdot (S, m_2)$ provided the sum is defined (e.g. when the sum is finite). With this, $K\langle\langle M\rangle\rangle$ is a semiring. The powers $S^n (n \geq 0)$ are defined in the natural way. We call S *proper*, if $(S, 1) = 0$, and then we put, in the natural way, $S^* = \sum_{n \geq 0} S^n$, the *star* (or iteration) of S, and $S^+ = \sum_{n \geq 1} S^n$, provided it is defined. We let $K^{rat}\langle\langle M\rangle\rangle$ denote the smallest subset of $K\langle\langle M\rangle\rangle$ which contains all polynomials and is closed under the operations sum, product and star, where the latter is only applied to proper series. Its elements are called *rational* formal power series. Now Schützenberger's theorem states the following equivalence between recognizable and rational series over the free monoid.

Theorem 2 (Schützenberger, [15]). *Let Σ be any finite set and K any semiring. Then*

$$K^{rec}\langle\langle \Sigma^*\rangle\rangle = K^{rat}\langle\langle \Sigma^*\rangle\rangle.$$

From this, Kleene's theorem on the coincidence of regular and rational languages follows by considering the Boolean semiring $\mathbb{B} = \{0, 1\}$ (with $1 + 1 = 1 \cdot 1 = 1$) and noting that a language $L \subseteq \Sigma^*$ is regular iff its characteristic series $1_L \in \mathbb{B}\langle\langle \Sigma^*\rangle\rangle$ is recognizable, and similarly for rationality.

Later we will also need the *Hadamard product* $S \odot T$ of two series $S, T \in K\langle\langle M\rangle\rangle$. It is defined by $(S \odot T, m) = (S, m) \cdot (T, m)$ for all $m \in M$.

Next we recall basic notions from trace theory. A pair (Σ, I) is called a *trace alphabet*, if Σ is a finite set and I is an irreflexive symmetric binary *independence* relation on Σ. Let $\sim$ denote the smallest congruence on Σ^* containing $\{(ab, ba) : a\ I\ b\}$. The quotient monoid $\mathbb{M} = \mathbb{M}(\Sigma, I) := \Sigma^*/\sim$ is called the *trace monoid* (or free partially commutative monoid) over (Σ, I). If $w \in \Sigma^*$, we let $[w]$ denote the equivalence class of w in $\mathbb{M}$. Also, let $\alpha(w)$ be the set of all letters of Σ occurring in w, called the *alphabet* of w. Since equivalent words have the same alphabet, we may put $\alpha([w]) = \alpha(w)$. If $A, B \subseteq \Sigma$, we write $A\ I\ B$ to denote that $a\ I\ b$ for all $a \in A, b \in B$. We also write $w\ I\ A$ or $[w]\ I\ A$ to abbreviate that $\alpha(w)\ I\ A$, similarly, $w\ I\ w'$ for $\alpha(w)\ I\ \alpha(w')$, etc. A subset $\Delta \subseteq \Sigma$ is called *connected*, if it cannot be split $\Delta = A \cup B$ into two non-empty subsets such that $A\ I\ B$. Again, w and $[w]$ are connected, if $\alpha(w)$ is connected. A language $L \subseteq \mathbb{M}$ or $L \subseteq \Sigma^*$ is called *connected*, if each of its elements is connected, and *mono-alphabetic*, if $\alpha(m) = \alpha(m')$ for all $m, m' \in L$. Then the collection of all *c-rational* languages in $\mathbb{M}$ (respectively, in Σ^*) is defined as the smallest set of languages of $\mathbb{M}$ (respectively, of Σ^*) containing all finite languages and which is closed under the operations union, product and star, where the latter is applied only to connected languages. The following characterizes the recognizable languages of $\mathbb{M}$ (recall that a language $L \subseteq \mathbb{M}$ is recognizable iff it is accepted by some finite $\mathbb{M}$-automaton, or, equivalently, iff its syntactic monoid is finite).

Theorem 3 (Ochmański, [12,4,6]). *Let (Σ, I) be any trace alphabet and $\mathbb{M}$ its trace monoid. Then a language $L \subseteq \mathbb{M}$ is recognizable iff it is c-rational.*

Again, one should note that the Kleene's theorem mentioned above is a special case of Theorem 3 since when the independence relation is empty, the trace monoid $\mathbb{M}(\Sigma, \emptyset)$ is the free monoid Σ^* and in this case all languages are connected, hence rational sets are also c-rational.

The goal of this paper is a common generalization of Theorems 2 and 3, that is, a characterization of the recognizable formal power series in $K\langle\langle\mathbb{M}\rangle\rangle$ where K is a semiring and $\mathbb{M}$ a trace monoid. Let $S \in K\langle\langle\mathbb{M}\rangle\rangle$. We say that S is *connected*, if $supp(S)$ is a connected language in $\mathbb{M}$, and *mono-alphabetic*, if $supp(S)$ is mono-alphabetic. In the latter case, we put $\alpha(S) = \alpha(m)$ if $S \neq 0$ and $m \in supp(S)$. Now let $K^{mc-rat}\langle\langle\mathbb{M}\rangle\rangle$ (mono-alphabetic-connected rational) be the smallest subset of $K\langle\langle\mathbb{M}\rangle\rangle$ which contains all polynomials and is closed under the operations sum, product and star, where the latter gets applied only to proper, mono-alphabetic and connected series. Similarly, we let $K^{c-rat}\langle\langle\mathbb{M}\rangle\rangle$ (connected rational) be the collection of series obtained from the polynomials by allowing the operations sum, product and star, where now star is applied to all proper and connected series. Similarly, we define connected series in $K\langle\langle\Sigma^*\rangle\rangle$ and the collection of *mc-rational* series in $K\langle\langle\Sigma^*\rangle\rangle$.

3 Mc-rational series are recognizable

In this section, let (Σ, I) be a trace alphabet and $\mathbb{M} = \mathbb{M}(\Sigma, I)$ its trace monoid. We will prove Theorem 1(b,c). This will require a more particular notion of representations which we introduce first.

Definition 4. Let $S = (\lambda, \mu, \gamma) \in K\langle\langle\mathbb{M}\rangle\rangle$ be a recognizable series with $\mu : \mathbb{M} \longrightarrow K^{n\times n}$. The representation (λ, μ, γ) is alphabetic, if there exist two functions $\overleftarrow{\alpha}, \overrightarrow{\alpha} : [n] \longrightarrow \mathcal{P}(\Sigma)$ such that for all $u \in \mathbb{M}$, the following three conditions are satisfied:
(1) Whenever $\mu u_{ij} \neq 0$, then $\overleftarrow{\alpha}(j) = \overleftarrow{\alpha}(i) \cup \alpha(u)$ and $\overrightarrow{\alpha}(i) = \overrightarrow{\alpha}(j) \cup \alpha(u)$;
(2) whenever $\lambda_i \neq 0$, then $\overleftarrow{\alpha}(i) = \emptyset$;
(3) whenever $\gamma_j \neq 0$, then $\overrightarrow{\alpha}(j) = \emptyset$.
We call $(\lambda, \mu, \gamma; \overleftarrow{\alpha}, \overrightarrow{\alpha})$ an alphabetic representation of S. Here, $\overleftarrow{\alpha}(k)$ describes the *past alphabet* of k and $\overrightarrow{\alpha}(k)$ the *future alphabet* of k. We say that k is *initial*, if $\overleftarrow{\alpha}(k) = \emptyset$, and k is *final*, if $\overrightarrow{\alpha}(k) = \emptyset$.

We will often use the fact that if (λ, μ, γ) is alphabetic and $\mu u_{ij} \neq 0$, then i initial implies that $\overleftarrow{\alpha}(j) = \alpha(u)$, and j final implies $\overrightarrow{\alpha}(i) = \alpha(u)$. Moreover, if $u \neq 1$, then i initial implies $\mu u_{ki} = 0$, and j final implies $\mu u_{jk} = 0$, for any k.

Proposition 5. *Let $S \in K\langle\langle\mathbb{M}\rangle\rangle$ be a recognizable series. Then there exists an alphabetic representation of S.*

First we want to show that the product of two recognizable series in $K\langle\langle\mathbb{M}\rangle\rangle$ is again recognizable. For more particular semirings K (strong Fatou semirings or the Boolean semiring), the result has been obtained already by Fliess [7, Prop. 2.2.14 and 2.2.15]. Our proof will not use the full notion of alphabetic representation, since it can be based either on the past alphabets (the function $\overleftarrow{\alpha}$) or the future alphabets, only. The full notion of alphabetic representation will come into use when we deal with iteration.

Theorem 6. *Let K be a commutative semiring and let $S_1, S_2 \in K\langle\langle\mathbb{M}\rangle\rangle$ be two recognizable series. Then their product $S = S_1 \cdot S_2$ is also recognizable.*

Proof. Let $(\lambda^1, \mu_1, \gamma^1)$ be a representation of S_1 and let $(\lambda^2, \mu_2, \gamma^2; \overleftarrow{\alpha}, \overrightarrow{\alpha})$ be an alphabetic representation of S_2 (Proposition 5). We assume that $\mu_i : \mathbb{M} \longrightarrow K^{n_i\times n_i}$ for $i = 1, 2$, and let $n = n_1 \cdot n_2$. Subsequently we identify $[n]$ with $[n_1] \times [n_2]$. Next, we define $\mu : \Sigma^* \longrightarrow K^{n\times n}$ by

$$\mu(a)_{(i_1,i_2)(j_1,j_2)} = \delta_{i_2,j_2} I(a, i_2)\mu_1(a)_{i_1,j_1} + \delta_{i_1,j_1}\mu_2(a)_{i_2,j_2}$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad I(u, i) = \begin{cases} 1 & \text{if } u \; I \; \overleftarrow{\alpha}(i) \\ 0 & \text{otherwise} \end{cases}$$

Note that $I(a, j_2)\mu_2(a)_{i_2,j_2} = 0$, hence at most one of the two terms is non-zero.

One can prove that $\mu(a) \cdot \mu(b) = \mu(b) \cdot \mu(a)$ for all $(a, b) \in I$. Hence, μ factorizes to a morphism $\mu : \mathbb{M} \longrightarrow K^{n\times n}$. Next we claim that this factorization is given by the explicit formula

$$\mu(w)_{(i_1,i_2)(j_1,j_2)} = \sum_{w=uv} I(u, i_2)\mu_1(u)_{i_1,j_1}\mu_2(v)_{i_2,j_2}$$

Finally, define $\lambda \in K^{1\times n}, \gamma \in K^{n\times 1}$ by $\lambda_{(i_1,i_2)} = \lambda^1_{i_1}\lambda^2_{i_2}$, $\gamma_{(k_1,k_2)} = \gamma^1_{k_1}\gamma^2_{k_2}$. We can verify that $S = (\lambda, \mu, \gamma)$ which proves the theorem.

The following result shows that a mono-alphabetic recognizable series has an alphabetic representation $(\lambda, \mu, \gamma; \overleftarrow{\alpha}, \overrightarrow{\alpha})$ with an even more specific form. For this, let $e_1 = (1, 0, \ldots, 0) \in K^{1\times n}$ and $e_n = (0, \ldots, 0, 1)^t \in K^{n\times 1}$.

Proposition 7. *Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be recognizable, proper and mono-alphabetic with $\alpha(S) = A$. Then there exists an alphabetic representation $(e_1, \mu, e_n; \overleftarrow{\alpha}, \overrightarrow{\alpha})$ of S with $\overrightarrow{\alpha}(1) = \overleftarrow{\alpha}(n) = A$.*

We will now prove the following essential closure property of recognizable series. Note that Theorem 1(b) follows easily from Theorems 6 and 8.

Theorem 8. *Let K be a commutative semiring and let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be a proper, connected, mono-alphabetic and recognizable series. Then, S^* is recognizable.*

The proof of this theorem is based on a rather involved construction. Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be a proper, recognizable, connected and mono-alphabetic series with $\alpha(S) = A$. Let $S = (e_1, \mu, e_n; \overleftarrow{\alpha}, \overrightarrow{\alpha})$ be an alphabetic representation with $\overrightarrow{\alpha}(1) = \overleftarrow{\alpha}(n) = A$ (Proposition 7). Let $m \geq 1$. We identify $[n^m]$ with the set $[n]^m$ of all m-tuples with entries from $[n]$. We use $\tilde{\imath}$ as abbreviation for such an m-tuple $(i_1, \ldots, i_m)$, similarly $\tilde{\jmath}, \tilde{k}$. Now we define functions $\mu^0, \ldots, \mu^m : \Sigma^* \longrightarrow K^{n^m\times n^m}$ by

$$\mu^0 a_{\tilde{\imath}\tilde{\jmath}} = \begin{cases} \mu a_{i_1 n} & \text{if } \tilde{\jmath} = (i_2, \ldots, i_m, 1) \\ 0 & \text{otherwise} \end{cases}$$

$$\mu^p a_{\tilde{\imath}\tilde{\jmath}} = \begin{cases} \mu a_{i_p j_p} & \text{if } j_l = i_l \text{ for all } l \neq p \\ 0 & \text{otherwise} \end{cases} \quad (p \geq 1)$$

Also, let

$$H_{\tilde{\imath}} = \begin{cases} 1 & \text{if } \overrightarrow{\alpha}(i_p) \cup \overleftarrow{\alpha}(i_p) = A = \alpha(S) \text{ for all } p,\ \overrightarrow{\alpha}(i_1) \neq \emptyset \text{ and} \\ & \quad \overrightarrow{\alpha}(i_p)\ I\ \overleftarrow{\alpha}(i_q) \text{ for all } p < q \\ 0 & \text{otherwise} \end{cases}$$

Let $H \in K^{n^m\times n^m}$ be given by $H_{\tilde{\imath}\tilde{\jmath}} = H_{\tilde{\imath}} \cdot H_{\tilde{\jmath}}$, and define $\mu^* : \Sigma^* \longrightarrow K^{n^m\times n^m}$ by $\mu^* = H \odot (\mu^0 + \cdots + \mu^m)$, where $(H \odot \mu^p)(w)_{\tilde{\imath}\tilde{\jmath}} = H_{\tilde{\imath}\tilde{\jmath}} \cdot \mu^p w_{\tilde{\imath}\tilde{\jmath}}$ for any $w \in \Sigma^*$ and $\tilde{\imath}, \tilde{\jmath} \in [n]^m$.

Theorem 8 results clearly from the following two essential results.

Proposition 9. *Let K be a commutative semiring and assume that $m \geq |A|$. Then $\mu^*(ab) = \mu^*(ba)$ for all $a, b \in \Sigma$ such that $a\ I\ b$.*

Hence μ^* factorizes to a morphism from $\mathbb{M}$ to $K^{n^m\times n^m}$, and we have:

Proposition 10. *Let K be a commutative semiring and assume that $m \geq |A|$. Then $S^* = (\lambda_{\tilde{1}}, \mu^*, \gamma_{\tilde{1}})$ where $\lambda_{\tilde{1}}, \gamma_{\tilde{1}}$ are the row respectively column vectors which have a 1 only at entry $\tilde{1} = (1, \ldots, 1)$, and 0 otherwise.*

Next we wish to derive a further closure properties of $K^{rec}\langle\langle \mathbb{M} \rangle\rangle$.

Definition 11. Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ or $S \in K\langle\langle \Sigma^* \rangle\rangle$ and $A \subseteq \Sigma$. Then the restriction of S to A is the series S_A defined by

$$(S_A, w) = \begin{cases} (S, w) & \text{if } \alpha(w) = A \\ 0 & \text{otherwise} \end{cases}$$

First we show that the restriction preserves both recognizability and mc-rationality of series.

Proposition 12. *Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be recognizable. Then S_A is also recognizable.*

Proposition 13. *Let $S \in K\langle\langle \Sigma^* \rangle\rangle$ or $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be mc-rational. Then S_A is also mc-rational.*

The following lemma generalizes a result of Pighizzini [13] for trace languages.

Lemma 14. *Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be proper and $A \subseteq \Sigma$ be nonempty. Then $(S^*)_A = Z^+ X$ where $X = \sum_{B \subset A} (S^*)_B$ and $Z = (X \cdot S)_A$.*

Next we derive another sufficient condition which implies that the star of a recognizable series is again recognizable and, also, that the star of an mc-rational series is again mc-rational.

Theorem 15.

1. *Let K be a commutative semiring and $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be proper and recognizable such that S^* is connected. Then S^* is recognizable.*
2. *Let K be any semiring and $S \in K\langle\langle \Sigma^* \rangle\rangle$ or $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be proper and mc-rational such that S^* is connected. Then S^* is mc-rational.*

For positive semirings, the condition S^* connected is stronger than S connected. This latter condition is actually sufficient to obtain the closure properties stated in Theorem 15 when the semiring is commutative and idempotent. This is an easy consequence of Theorem 1(a) and of Theorem 17 for which the following lemma is crucial.

Lemma 16. *Let K be a commutative and idempotent semiring. Let $S \in K\langle\langle \mathbb{M} \rangle\rangle$ be a connected series and let $B, C \subseteq \Sigma$ be independent subsets of the alphabet. Then, $(S^*)_{B \cup C} = (S^*)_B \cdot (S^*)_C$.*

Theorem 17. *Let K be a commutative and idempotent semiring. A series in $K\langle\langle \mathbb{M} \rangle\rangle$ is mc-rational iff it is c-rational.*

Proof. One direction is clear and for the converse, it suffices to show that the star of an mc-rational connected series S is still mc-rational. We will first show by induction on the size of $A \subseteq \Sigma$ that if S is an mc-rational connected series then $(S^*)_A$ is mc-rational. The theorem follows directly since $S^* = \sum_{A \subseteq \Sigma} (S^*)_A$.

Clearly, $(S^*)_\emptyset = 1$ is mc-rational. Now, assume $A \neq \emptyset$ and let $A_1, \ldots, A_n$ be the connected components of A: $A = A_1 \cup \cdots \cup A_n$ and $A_i \; I \; A_j$ for $i \neq j$. By Lemma 16, we obtain $(S^*)_A = (S^*)_{A_1} \cdots (S^*)_{A_n}$ and we are reduced to the case A connected. Now, using Lemma 14 we obtain $(S^*)_A = Z^+ X$ where $X = \sum_{B \subset A} (S^*)_B$ and $Z = (X \cdot S)_A$. Then X is mc-rational by induction hypothesis. By Proposition 13, it follows that Z is also mc-rational. Since we have assumed A connected, we deduce that $(S^*)_A = Z \cdot Z^* \cdot X$ is mc-rational.

Note that Theorem 1(c) follows from Theorem 1(b) and Theorem 17.

4 Recognizable series are mc-rational

Thoughout this section, let K be an arbitrary (possibly non-commutative) semiring and (Σ, I) a trace alphabet. We will prove that all recognizable series in $K\langle\langle \mathbb{M} \rangle\rangle$ are mc-rational. This uses the concept of lexicographic normal forms of traces and LNF-representations of series which we introduce first. For this, fix any linear order $\leq$ on Σ. We extend this to the lexicographic linear order, also denoted by $\leq$, on Σ^*. We say that a word w is the lexicographic normal form of $[w]$, if it is the smallest element of $[w]$ with respect to $\leq$. Then LNF is the set of all words which are lexicographic normal forms. Note that LNF is closed under prefixes (and suffixes). Now let $\mathcal{A}_{\mathrm{LNF}} = (Q, \Sigma, \delta, q_0, Q)$ be the minimal (reduced) automaton for LNF.

Definition 18. We will call a morphism $\mu : \Sigma^* \longrightarrow K^{n \times n}$ an LNF-morphism, if there exists a function $\pi : [n] \longrightarrow Q$ such that for all $a \in \Sigma$ and all $i, j \in [n]$, $\mu a_{ij} \neq 0$ implies $\pi(i) \xrightarrow{a} \pi(j)$ in $\mathcal{A}_{\mathrm{LNF}}$. Then any representation (λ, μ, γ) with an LNF-morphism μ of a series $S \in K\langle\langle \Sigma^* \rangle\rangle$ will be called an LNF-representation of S.

Proposition 19. *Let $S' \in K\langle\langle \Sigma^* \rangle\rangle$ be recognizable. Then $S = S' \odot 1_{\mathrm{LNF}}$ has an* LNF-*representation.*

Next we note that for any $n \geq 1$ there is a canonical isomorphism Φ between the semiring of $n \times n$-matrices $K\langle\langle \Sigma^* \rangle\rangle^{n \times n}$ and the semiring of formal power series $K^{n \times n}\langle\langle \Sigma^* \rangle\rangle$, given by $(\Phi(A), w) = ((A_{ij}, w))$ if $A = (A_{ij}) \in K\langle\langle \Sigma^* \rangle\rangle^{n \times n}$. Subsequently, we will often identify A with its image $\Phi(A)$.

We will also use the following result.

Lemma 20 (Ochmański, [12,4]). *Let $w \in \Sigma^*$ be a word such that $w, w^2 \in$* LNF. *Then w is connected.*

Proposition 21. *Let $\mu : \Sigma^* \longrightarrow K^{n \times n}$ be an* LNF-*morphism, and let $M = \sum_{a \in \Sigma} \mu a \cdot a \in K^{n \times n}\langle \Sigma^* \rangle$. Then the entries of M^* are mc-rational series.*

Proof. We first show, by induction on the length of w, that $(M^*, w) = \mu w$ for any word w. Indeed, clearly $(M^*, 1) = 1 = \mu 1$ and $(M^*, wa) = (1 + M^* M, wa) = (M^* M, wa) = (M^*, w)(M, a) = \mu w \cdot \mu a = \mu(wa)$.

By lack of space we only give the proof for $n = 1$, which already shows several connections between all the results. Hence, assume that $n = 1$. Then $M \in K\langle \Sigma^* \rangle$ is proper and mc-rational. Now, let $w \in \Sigma^*$. If $(M^*, w) = \mu w \neq 0$, since μ is an LNF-morphism, we have a path $\pi(1) \xrightarrow{w} \pi(1)$ in $\mathcal{A}_{\mathrm{LNF}}$. Therefore, $w, w^2 \in \mathrm{LNF}$ and by Ochmański's lemma 20, w is connected. Hence M^* is connected and so, by Theorem 15, mc-rational.

Theorem 22. *Let $S \in K\langle\langle \Sigma^* \rangle\rangle$ be recognizable. Then $S \odot 1_{\mathrm{LNF}}$ is mc-rational.*

Proof. By Proposition 19 we can choose an LNF-representation (λ, μ, γ) of $S' = S \odot 1_{\mathrm{LNF}}$. Let $M = \sum_{a \in \Sigma} \mu a \cdot a$. We have seen in the proof of Proposition 21 that $(M^*, w) = \mu w$ for any word w.

Now, λ and γ are vectors with entries in K, and M^* has only mc-rational series as entries by Proposition 21. Hence $\lambda M^* \gamma \in K\langle\langle \Sigma^* \rangle\rangle$ is an mc-rational series. Finally, observe that for any word w,

$$(\lambda M^* \gamma, w) = (\sum_{i,j} \lambda_i (M^*)_{ij} \gamma_j, w) = \sum_{i,j} \lambda_i ((M^*)_{ij}, w) \gamma_j)$$

$$= \sum_{i,j} \lambda_i \mu w_{ij} \gamma_j = \lambda \mu w \gamma = (S', w).$$

Therefore $S \odot 1_{\mathrm{LNF}} = S' = \lambda M^* \gamma$ is mc-rational.

Corollary 23. *Let $S \in K\langle\langle \Sigma^* \rangle\rangle$ be recognizable with supp$(S) \subseteq \mathrm{LNF}$. Then S is mc-rational.*

Let M, N be two monoids and $h : M \longrightarrow N$ be a morphism. Then $h^{-1} : K\langle\langle N \rangle\rangle \longrightarrow K\langle\langle M \rangle\rangle$ given by $(h^{-1}(S), w) = (S, h(w))$ $(w \in N)$ is a semiring morphism. Moreover, if $S = (\lambda, \mu, \gamma) \in K^{rec}\langle\langle N \rangle\rangle$, then $(h^{-1}(S), w) = (S, h(w)) = \lambda \mu h(w) \gamma$, hence (cf. [14, p.32])

$$h^{-1}(S) = (\lambda, \mu \circ h, \gamma) \in K^{rec}\langle\langle M \rangle\rangle.$$

Let $\varphi : \Sigma^* \longrightarrow \mathbb{M}$ be the canonical epimorphism. Then φ extends naturally to a mapping, denoted by Φ, from $K\langle\langle \Sigma^* \rangle\rangle$ to $K\langle\langle \mathbb{M} \rangle\rangle$ given by

$$\Phi(S) = \sum_{w \in \Sigma^*} (S, w) \varphi(w) = \sum_{t \in \mathbb{M}} \left(\sum_{w \in \varphi^{-1}(t)} (S, w) \right) .t.$$

As is well-known from general results (cf., e.g., [14, pp.13,14]), Φ is a semiring morphism and if S is proper, then $\Phi(S^*) = \Phi(S)^*$. Furthermore, if S is connected (respectively, mono-alphabetic), then $\Phi(S)$ is also connected (respectively, mono-alphabetic). From this, it is clear that if S is mc-rational, then $\Phi(S)$ is also mc-rational. Now we prove Theorem 1(a).

Theorem 24. *Let $S \in K\langle\langle\mathbb{M}\rangle\rangle$ be recognizable. Then S is mc-rational.*

Proof. Let $S = (\lambda, \mu, \gamma) \in K^{rec}\langle\langle\mathbb{M}\rangle\rangle$. As noted before, $\varphi^{-1}(S) \in K^{rec}\langle\langle\Sigma^*\rangle\rangle$. By Theorem 22, $\varphi^{-1}(S) \odot 1_{\mathrm{LNF}}$ is mc-rational. Hence also $\Phi(\varphi^{-1}(S) \odot 1_{\mathrm{LNF}})$ is mc-rational. Now for each $t \in \mathbb{M}$ we have

$$\begin{aligned}(\Phi(\varphi^{-1}(S) \odot 1_{\mathrm{LNF}}), t) &= \sum_{w \in \varphi^{-1}(t)} (\varphi^{-1}(S) \odot 1_{\mathrm{LNF}}, w) \\ &= \sum_{w \in \varphi^{-1}(t) \cap \mathrm{LNF}} (\varphi^{-1}(S), w) = \sum_{w \in \varphi^{-1}(t) \cap \mathrm{LNF}} (S, \varphi(w)) = (S, t).\end{aligned}$$

Therefore, $S = \Phi(\varphi^{-1}(S) \odot 1_{\mathrm{LNF}})$ is mc-rational.

5 Examples and consequences

Here we will give two examples to show that the assumptions in Theorems 6 and 8 (hence, in Theorem 1(b,c)) are necesssary. We also indicate the relationship with the results of Schützenberger and Ochmański. First, we show that in Theorem 6 the commutativity of K is necessary.

Example 25. Consider the trace alphabet (Σ, I) with $\Sigma = \{a, b\}$ and $a \; I \; b$, and let $K = \mathbb{B}\langle\Sigma^*\rangle$. Let $S = \sum_n a^n.a^n, T = \sum_n b^n.b^n \in K\langle\langle\mathbb{M}\rangle\rangle$. Then S and T are recognizable. Indeed, if $\mu : \Sigma^* \longrightarrow K$ is defined by $\mu(a) = a$ and $\mu(b) = 0$ and $\lambda = \gamma = 1$, then $S = (\lambda, \mu, \gamma)$. However, we can show that $S \cdot T \in K\langle\langle\mathbb{M}\rangle\rangle$ is *not* recognizable.

Secondly, we want to show that in general $K^{rec}\langle\langle\mathbb{M}\rangle\rangle$ is properly contained in $K^{c-rat}\langle\langle\mathbb{M}\rangle\rangle$. That is, we show that the star of a connected recognizable series may not be recognizable. (Thus by Theorem 15, the star of this series will not be connected.)

Example 26. Again consider the trace alphabet (Σ, I) with $\Sigma = \{a, b\}$ and $a \; I \; b$, and let $S = a + b \in \mathbb{N}\langle\mathbb{M}\rangle$. Then, obviously, S is a connected polynomial and $(S^*, t) = \binom{|t|_a + |t|_b}{|t|_a}$ for all $t \in \mathbb{M}$. Hence, $S^* = \sum_{n,m \in \mathbb{N}} \binom{n+m}{n} a^n b^m$. We can prove that S^* is not recognizable.

Let Σ be any finite alphabet. If $I = \emptyset$, the trace monoid $\mathbb{M}(\Sigma, I)$ is isomorphic to Σ^*. Hence, by Theorem 24 we have $K^{rec}\langle\langle\Sigma^*\rangle\rangle \subseteq K^{mc-rat}\langle\langle\Sigma^*\rangle\rangle \subseteq K^{rat}\langle\langle\Sigma^*\rangle\rangle$. Now, using one inclusion of Theorem 2, we obtain $K^{rec}\langle\langle\Sigma^*\rangle\rangle = K^{mc-rat}\langle\langle\Sigma^*\rangle\rangle = K^{rat}\langle\langle\Sigma^*\rangle\rangle$ which is in fact a strengthening of Theorem 2.

Now we show how to deduce and actually strengthen Theorem 3 from our results. The following can be proved in the same way as classically for the free monoid (cf. [14,2]).

Proposition 27. *$L \subseteq \mathbb{M}$ is recognizable (resp. rational, c-rational, mc-rational) iff $1_L \in \mathbb{B}\langle\langle\mathbb{M}\rangle\rangle$ is recognizable (resp. rational, c-rational, mc-rational).*

Since the boolean semiring $\mathbb{B}$ is both commutative and idempotent, we deduce from Theorem 1 that a series in $\mathbb{B}\langle\langle\mathbb{M}\rangle\rangle$ is recognizable iff it is c-rational iff it is mc-rational. Using Proposition 27, we deduce that a trace language $L \subseteq \mathbb{M}$ is recognizable iff it is c-rational iff it is mc-rational. The first equivalence is precisely Ochmański's theorem. The second one is a strengthening of a result by Pighizzini [13] which characterizes the recognizable languages as those languages obtained from finite sets of traces using union, concatenation, *restriction to subalphabet* and star restricted to monoalphabetic and connected languages.

References

1. I.J. Aalbersberg and G. Rozenberg. Theory of traces. *Theoretical Computer Science*, 60:1–82, 1988.
2. J. Berstel and Ch. Reutenauer. *Rational Series and Their Languages*, volume 12 of *EATCS Monographs in Theoretical Computer Science*. Springer Verlag, 1988.
3. Ch. Choffrut. Free partially commutative monoids. Rapport LITP 86.20, Université Paris 7 (France), 1986.
4. V. Diekert. *Combinatorics on Traces*. Number 454 in Lecture Notes in Computer Science. Springer Verlag, 1990.
5. V. Diekert and Y. Métivier. Partial commutation and traces. In G. Rozenberg and A. Salomaa, editors, *Handbook on Formal Languages*, volume III. Springer Verlag. To appear.
6. V. Diekert and G. Rozenberg, editors. *Book of Traces*. World Scientific, Singapore, 1995.
7. M. Fliess. Matrices de Hankel. *J. Math. Pures et Appl.*, 53:197–224, 1974.
8. W. Kuich. Semirings and formal power series: Their relevance to formal languages and automata. In *Handbook on Formal Languages*. Springer Verlag, 1997. To appear.
9. W. Kuich and A. Salomaa. *Semirings, Automata, Languages*, volume 6 of *EATCS Monographs in Theoretical Computer Science*. Springer Verlag, 1986.
10. A. Mazurkiewicz. Concurrent program schemes and their interpretations. Tech. rep. DAIMI PB 78, Aarhus University, 1977.
11. A. Mazurkiewicz. Trace theory. In W. Brauer et al., editors, *Advances in Petri Nets'86*, number 255 in Lecture Notes in Computer Science, pages 279–324. Springer Verlag, 1987.
12. E. Ochmański. Regular behaviour of concurrent systems. *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, 27:56–67, Oct 1985.
13. G. Pighizzini. Synthesis of nondeterministic asynchronous automata. In M. Droste and Y. Gurevich, editors, *Semantics of Programming Languages and Model Theory*, number 5 in Algebra, Logic and Applications, pages 109–126. Gordon and Breach Science Publ., 1993.
14. A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series*. Texts and Monographs in Computer Science. Springer Verlag, 1978.
15. M.P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.