Chapter 4: Rational and Recognisable Power Series*

Jacques Sakarovitch

LTCI, ENST/CNRS,
46 rue Barrault, 75634 Paris Cedex 13, France
sakarovitch@enst.fr

* This chapter is adapted from Chaps. III and IV of the book *Elements of Automata Theory*, Jacques Sakarovitch, 2009, ©Cambridge University Press, where missing proofs, detailed examples and further developments can be found.

M. Droste, W. Kuich, H. Vogler (eds.), *Handbook of Weighted Automata*,
Monographs in Theoretical Computer Science. An EATCS Series,
DOI 10.1007/978-3-642-01492-5_4, Springer-Verlag Berlin Heidelberg 2009

1 Introduction

Weighted automata realise power series—in contrast to 'classical' automata which accept languages. There are many good reasons that make power series worth an interest compared to languages, beyond the raw appeal to generalisation that inhabits every mathematician.

First, power series provide a more powerful mean for modelisation, replacing a pure acceptance/rejection mode by a quantification process. Second, by putting automata theory in a seemingly more complicated framework, one benefits from the strength of mathematical structures thus involved and some results and constructions become simpler, both conceptually, and on the complexity level. Let us also mention, as a third example, that in the beginning of the theory, weighted automata were probably considered for their ability of defining languages—via the supports of realised power series—rather than for the power series themselves. In all these instances, what matters is that the choice of the semiring S of multiplicity be as wide as possible and our first aim is to develop as far as possible a theory with a priori no assumption at all on S.

With this in mind, I have chosen as the main thread of this chapter to lay comprehensive bases for the proof of the decidability of the equivalence of deterministic k-tape transducers which is, at least in my opinion, one of the most striking examples of the application of algebra to "machine theory." To that end, I develop in particular the following points:

(a) The definition of rational series over *graded monoids* (in order to deal with direct product of free monoids) and not over free monoids only. A side benefit of the definition of series over arbitrary (graded) monoids is that it makes clearer the distinction between the *rational* and the *recognisable* series.
(b) The *reduction theory* of series over a free monoid and with coefficients in a (skew) field that leads to a procedure for the decidability of equivalence (with a cubic complexity).
(c) As it is natural for series with coefficients in a field, and since the topological machinery is set anyway, the star of series is defined in a slightly more general setting than cycle-free series.
(d) The basics for rational relations with multiplicity, for the weighted generalisation of the often called Kleene–Schützenberger theorem on transducers as well as of the Myhill theorem (on recognisable sets in a product of monoids) or of McKnight theorem (on the inclusion of recognisable set in rational ones in finitely generated monoids).

The core of this chapter pertains to a now classical part of automata theory, originating in the seminal paper of M.P. Schützenberger [46] and having been exposed in several treatises already quoted in Chap. 1: Eilenberg [14], Salomaa and Soittola [45], Berstel and Reutenauer [5], and Kuich and Salomaa [29]. I have not resisted though to include some more recent developments which are the result of my own work with my colleagues M.-P. Béal and S. Lombardy: the derivation of weighted expressions [33], and the connection between conjugacy and equivalence [3, 4].

The presentation given here (but for the last quoted result that is too recent) is adapted from Chaps. III and IV of my book *Elements of Automata Theory* [43], where missing proofs, detailed examples, and further developments can be found. I am grateful to Reuben Thomas who has translated this book from French to English and to Cambridge University Press for allowing me to use the material for this chapter. Finally, I want to acknowledge the always inspiring discussions I have had in the last 10 years with Sylvain Lombardy.

2 Rational Series and Weighted Rational Expressions

In the preceding chapters, the formal power series that have been considered are series over a free monoid with coefficients in a semiring S that is almost always supposed to be *complete* or *continuous*, opening the way to straightforward generalisations of results and methods developed for languages, that are series with multiplicity in the Boolean semiring, and classical automata.

Our first purpose is to build a theory where no assumptions are made on the semiring of coefficients, and as few as possible on the base monoid. There will be some redundancy with Chaps. 1 and 3, but I have preferred to write a comprehensive text that naturally flows rather than to interrupt it with references to results that are always stated under slightly different hypotheses.

In what follows, M is a monoid and S a semiring, a priori arbitrary.

2.1 Series over a Graded Monoid

For any set E, the set of maps from E to S is usually written S^E and canonically inherits from S a structure of semiring when equipped with *pointwise* addition and multiplication. When E is a monoid M, we equip S^M with another multiplication which derives from the *monoid structure* of M, and we thus use different notation and terminology for these maps together with this other semiring structure—indeed, the ones set up in Chap. 1, Sect. 3.

Any map from M to S is a *formal power series* over M with coefficients in S—abbreviated as S-series over M, or even as *series* if there is ambiguity neither on S nor on M. The set of these series is written $S\langle\langle M\rangle\rangle$. If r is a

series, the image of an element m of M under r is written (r, m) rather than $(m)r$ and is called the *coefficient of m in r*.

The *support* of a series r is the subset of elements of M whose coefficient in r is not 0_S. A series with finite support is a *polynomial*; the set of polynomials over M with coefficients in S is written $S\langle M\rangle$.

For all r and r', and all s in S, the following operations on $S\langle\langle M\rangle\rangle$ are defined:

(i) The (left and right) *'exterior' multiplications*[1]:

$$sr \quad \text{and} \quad rs \quad \text{by} \quad \forall m \in M \quad (sr, m) = s(r, m) \text{ and } (rs, m) = (r, m)s.$$

(ii) The pointwise *addition*:

$$r + r' \quad \text{by} \quad \forall m \in M \quad (r + r', m) = (r, m) + (r', m).$$

(iii) The *Cauchy product*:

$$rr' \quad \text{by} \quad \forall m \in M \quad (rr', m) = \sum_{\substack{u,v \in M \\ uv = m}} (r, u)(r', v). \qquad (*)$$

Addition makes $S\langle\langle M\rangle\rangle$ a commutative monoid, whatever S and M; together with the two exterior multiplications, it makes $S\langle\langle M\rangle\rangle$ a left and right *semimodule*[2] on S.

The Cauchy product raises a problem for there could very well exist elements m in M such that the set of pairs (u, v) satisfying $uv = m$ is infinite, and hence there could exist series such that the sum on the right-hand side of $(*)$ is *not defined*. Thus, we cannot ensure, without further assumptions, that the Cauchy product is a binary operation totally defined on $S\langle\langle M\rangle\rangle$. This difficulty can be overcome in at least three ways.

The first is to retreat: we no longer consider $S\langle\langle M\rangle\rangle$ but only the set $S\langle M\rangle$ of polynomials. If r and r' are polynomials, the sum in $(*)$ is infinite but only a finite number of terms are non-zero; the Cauchy product is defined on $S\langle M\rangle$ and makes it indeed a semiring (a semi-algebra on S), a subsemi-algebra of $S\langle\langle M\rangle\rangle$ when that is defined.

The second is to assume that S is *complete*: every sum, even if infinite, is defined on S, and the Cauchy product of two series is defined for any M. This is the case, for example, if S is equal to $\mathbb{B}$, $\mathbb{B}\langle\langle M\rangle\rangle$, $\langle \mathbb{N}^\infty, +, \cdot\rangle$ or $\langle \mathbb{N}^\infty, \mathsf{min}, +\rangle$. The theory of finite automata over a free monoid and with multiplicity in a complete semiring has been developed in Chap. 3 of this book.

The third way, which is ours, aims at being able to define weighted automata, and hence series, *without restriction on S*, and we are led in this case

[1] Which are called *scalar products* in Chap. 1.

[2] For sake of uniformity in this book, I use the terms 'semimodule' and 'semialgebra' whereas in [43] and other publications, I follow the convention of Berstel and Reutenauer [5] and speak of 'module' and 'algebra' (over a semiring).

to make assumptions about M: we suppose for the rest of this chapter that the monoids are *graded*, a condition that we shall describe in the next paragraph and which allows the natural generalisation of the standard construction of formal power series of a single variable.[3] This somewhat different assumption makes it necessary to restate, and sometimes to reprove again, some of the statements already established when S is *complete*.

2.1.1 Graded Monoid

For the Cauchy product to be always defined on $S\langle\langle M\rangle\rangle$, independently of S, it is necessary (and sufficient) that, for all m in M, the set of pairs (u, v) such that $uv = m$ is finite—we will say that m is *finitely decomposable*. However, making $S\langle\langle M\rangle\rangle$ a semiring is not an end in itself: the development of the theory to come is the characterisation of the behaviour of finite automata by means of rational operations—a fundamental theorem—and then not only must sum and product be defined on the series, but so must the star operation, which implies an *infinite sum*. This forces us to have some sort of *topology* on $S\langle\langle M\rangle\rangle$, to which we shall return in the next paragraph.

The construction of series on Σ^*, which generalises that of series of one variable, shows that it is from the *length* of words in Σ^* that we build a topology on $S\langle\langle\Sigma^*\rangle\rangle$. The existence of an *additive length* is the main assumption that we shall make about M. Returning to the initial problem, we then seek an additional condition that ensures that every element is finitely decomposable. For reasons of simplicity, we assume that M is *finitely generated*. This solves the problem, while allowing us to deal with the cases that interest us.

Definition 2.1. *A function* $\varphi\colon M \to \mathbb{N}$ *is a* length *on* M *if:*

(i) $\varphi(m)$ *is* strictly *positive for all* m *other than* 1_M
(ii) $\forall m, n \in M\ \varphi(mn) \leq \varphi(m) + \varphi(n)$

We shall say that a length is a gradation *if it is* additive*; that is, if:*

(iii) $\forall m, n \in M\ \varphi(mn) = \varphi(m) + \varphi(n)$

and that M *is* graded *if it is equipped with a gradation.*

Every free monoid and every Cartesian product of free monoids is graded. The definition implies that $\varphi(1_M) = 0$ and that a finite monoid, more generally a monoid that contains an idempotent other than the identity (for example, a zero), cannot be equipped with a gradation.

Proposition 2.2. *In a finitely generated graded monoid, the number of elements whose length is less than an arbitrary given integer* n *is finite.*

[3] A fourth method exists that takes out of both the first and the third. It involves making an assumption about M (we require it to be an *ordered group*) and considering only a subset of $S\langle\langle M\rangle\rangle$ (those series whose support is well ordered). A reference to that set of series will be made in Sect. 5.3.

In other words, every element of a graded monoid M can only be written in a finite number of different ways as the product of elements of M other than 1_M. We can deduce in particular the following corollary.

Corollary 2.3. *In a finitely generated graded monoid, every element is finitely decomposable.*

Note that a finite monoid is not graded, but that every element in it is nonetheless finitely decomposable. From Corollary 2.3, we deduce the proposition aimed at by Definition 2.1:

Proposition 2.4. *Let M be a finitely generated graded monoid and S a semiring. Then $S\langle\langle M\rangle\rangle$, equipped with the Cauchy product, is a semiring, and what is more, a (left and right) semi-algebra*[4] *on S.*

In the following, M is a graded monoid that is implicitly assumed to be finitely generated. To simplify the notation and in imitation of the free monoid, we will write the length function as a pair of vertical bars, that is, $|m|$ rather than $\varphi(m)$.

From the semiring $S\langle\langle M\rangle\rangle$, one then builds other semirings, by means of classical constructions; let us quote in particular and for further reference the following fundamental isomorphism.

Lemma 2.5. *Let S be a semiring, M a graded monoid, and Q a finite set; then the set of square matrices of dimension Q and with entries in the semiring $S\langle\langle M\rangle\rangle$ is a semiring, isomorphic to that of series over M with coefficient in the semiring of square matrices of dimension Q and with entries in S; that is, $S\langle\langle M\rangle\rangle^{Q\times Q} \cong S^{Q\times Q}\langle\langle M\rangle\rangle$.*

Remark 2.6. A notion that is often considered in relationship with gradation is *equidivisibility*. A monoid M is *equidivisible* if whenever $mn = pq$ with m, n, p, and q in M, there exists u such that $mu = p$ and $n = uq$ or $m = pu$ and $un = q$. There is then a theorem by F.W. Levi which states that *a graded equidivisible monoid is free* (cf. [30]). This notion is also to be compared with the one of *equisubtractivity* that is considered below.

2.1.2 Topology on $S\langle\langle M\rangle\rangle$

The definition to come of the *star operation*, an infinite sum, calls for the definition of a *topology* on $S\langle\langle M\rangle\rangle$.

Since $S\langle\langle M\rangle\rangle = S^M$ is the *set of maps* from M to S, it is naturally equipped with the *product topology* of the topology on S. If this topology on S is defined by a *distance*, the product topology on $S\langle\langle M\rangle\rangle$ coincides, as M is countable, with the *simple convergence topology*:

[4] If S is a ring, $S\langle\langle M\rangle\rangle$ is even what is classically called a *graded algebra*, which is the origin of the terminology chosen for graded monoids.

$$r_n \ \textit{converges to } r, \quad \text{if and only if,}$$
$$\text{for all } m \text{ in } M, \quad (r_n, m) \textit{ converges to } (r, m).$$

We shall reexamine the topology question using only the notion of distance, more in line with intuition and explain how to define a distance between two series under the assumption that M is graded. The foregoing reference to simple convergence topology was nevertheless worthwhile, as it made clear that *the basis of the topology on* $S\langle\langle M\rangle\rangle$ *is the topology on* S.

Distance on $S\langle\langle M\rangle\rangle$

A *distance* on a set E is a map $\mathbf{d}$ which relates to every pair (x, y) of elements of E a *positive real number* $\mathbf{d}(x, y)$, called the *distance from* x *to* y (or *between* x and y), which satisfies the following properties:

- Symmetry: $\mathbf{d}(x, y) = \mathbf{d}(y, x)$
- Positivity: $\mathbf{d}(x, y) > 0$ if $x \neq y$ and $\mathbf{d}(x, x) = 0$
- Triangular inequality: $\mathbf{d}(x, y) \leq \mathbf{d}(x, z) + \mathbf{d}(y, z)$

When this triangular inequality can be replaced by

- $\forall x, y, z \in E$ $\mathbf{d}(x, y) \leq \mathsf{max}\{\mathbf{d}(x, z), \mathbf{d}(y, z)\}$

the distance $\mathbf{d}$ is called *ultrametric*.

A sequence $\{x_n\}_{n\in\mathbb{N}}$ of elements of E *converges* to x if the distance between x_n and x becomes arbitrarily small as n grows; that is, more formally,

$$\forall \eta > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N \quad \mathbf{d}(x_n, x) \leq \eta.$$

Such an element x is *unique*; it is called the *limit* of the sequence $\{x_n\}_{n\in\mathbb{N}}$ and we write $x = \mathsf{lim}_{n\to+\infty} x_n$, or simply $x = \mathsf{lim}\, x_n$ if there is no ambiguity. We say that $\mathbf{d}$ *equips* E *with a topology*.

Remark 2.7. We can always assume that a distance is a real number less than or equal to 1. If that is not the case, then by taking

$$\mathbf{f}(x, y) = \mathsf{inf}\{\mathbf{d}(x, y), 1\},$$

we obtain a distance $\mathbf{f}$ on E that defines *the same topology*; that is, a distance for which *the same sequences* will converge to *the same limits*.

Remark 2.8. Whatever E is, we can choose a trivial distance function which is 1 for every pair of distinct elements. This is equivalent to saying that two distinct elements are never 'close' to each other, and that the only convergent sequences are those that are eventually *stationary*. We then say that E is equipped with the *discrete topology*.

We are confronted with two situations which seem fundamentally different. The first is that of a semiring S such as $\mathbb{B}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{N}^\infty$, etc., whose elements are 'detached' from each other. The natural topology on these semirings is the discrete topology. The second is that of semirings such as $\mathbb{Q}$, $\mathbb{Q}_+$, $\mathbb{R}$, etc., or even later $S\langle\langle M\rangle\rangle$ itself, which can act as a semiring of coefficients for series on another monoid; that is, semirings on which there is a priori a distance which can be arbitrarily small. On these semirings as well, we can choose a discrete topology, but it is more satisfactory to preserve their 'native' topology. By means of the definition of a distance and the topological notions derived from it, we treat these two situations in the same way.

We first assume that S is equipped with a distance $\mathbf{c}$ which is bounded by 1. The length function on M allows us *to put an ordering* on the elements of M and we set

$$\mathbf{d}(r, r') = \frac{1}{2} \sum_{n \in \mathbb{N}} \left(\frac{1}{2^n} \max\{\mathbf{c}\big((r, m), (r', m)\big) \mid |m| = n\} \right).$$

We then verify that $\mathbf{d}$ is indeed a distance on $S\langle\langle M\rangle\rangle$, ultrametric when $\mathbf{c}$ is, and that the topology defined on $S\langle\langle M\rangle\rangle$ by $\mathbf{d}$ is, as stated, the *simple convergence* topology; that is, the following property.

Property 2.9. A sequence $\{r_n\}_{n\in\mathbb{N}}$ of series of $S\langle\langle M\rangle\rangle$ converges to r, if and only if, for all m in M the sequence of coefficients (r_n, m) converges to (r, m).

Furthermore, choosing a topology on a semiring only really makes sense if the constituent operations of the semiring, addition and multiplication, are consistent with the topology—we say they are *continuous*—that is, if the limit of a sum (resp. of a product) is the sum (resp. the product) of the limits. We say in this case that not only is the semiring equipped with a topology, but that it is a *topological semiring*. We easily verify that if S is topological, then so is $S\langle\langle M\rangle\rangle$. In other words, if $\{r_n\}_{n\in\mathbb{N}}$ and $\{r'_n\}_{n\in\mathbb{N}}$ are two convergent sequences of elements of $S\langle\langle M\rangle\rangle$, we have

$$\lim(r_n + r'_n) = (\lim r_n) + (\lim r'_n) \quad \text{and} \quad \lim(r_n r'_n) = (\lim r_n)(\lim r'_n).$$

Note that conversely the fact that the sequence $\{r_n + r'_n\}_{n\in\mathbb{N}}$ or $\{r_n r'_n\}_{n\in\mathbb{N}}$ converges *says nothing* about whether $\{r_n\}_{n\in\mathbb{N}}$ or $\{r'_n\}_{n\in\mathbb{N}}$ converges or not.

If S is a topological semiring, then so is $S^{Q\times Q}$ and the isomorphism quoted in Lemma 2.5 is moreover a *bi-continuous* bijection.

Summable Families

Let T be a semiring[5] equipped with a distance which makes it a topological semiring. We thus know precisely what means that an infinite sequence $\{t_n\}_{n\in\mathbb{N}}$

[5] We have temporarily changed the symbol we use for a semiring on purpose: T will not only play the role of S but also of $S\langle\langle M\rangle\rangle$ in what follows.

converges to a limit t when n tends to infinity. We must now give an equally precise meaning to the sum of an infinite family $\{t_i\}_{i\in I}$ and it turns out to be somewhat harder. The difficulty arises from the fact that we want a sort of *associativity–commutativity* extended 'to infinity', and hence to ensure that the result and its existence does not depend on an arbitrary order put on the set I of indices.

We shall therefore define an 'absolute' method of summability, and a family will be described as 'summable' if we can find an increasing sequence of finite sets of indices, a sort of 'kernels', such that not only do partial sums on these sets tend to a limit, but above all that any sum on a finite set containing one of these kernels stays close to this limit. More precisely, we take the following definition.

Definition 2.10. *A family $\{t_i\}_{i\in I}$ of elements of T indexed by an arbitrary set I is called* summable *if there exists t in T such that, for all positive η, there exists a finite subset J_η of I such that, for all finite subsets L of I which contain J_η, the distance between t and the sum of $\{t_i\}$ for i in L is less than η; that is,*

$$\exists t \in T,\ \forall \eta > 0,\ \exists J_\eta \text{ finite},\ J_\eta \subset I,\ \forall L \text{ finite},\ J_\eta \subseteq L \subset I$$
$$\mathbf{d}\Big(\sum_{i\in L} t_i, t\Big) \leq \eta.$$

The element t thus defined is unique *and is called the* sum *of the family $\{t_i\}_{i\in I}$.*

The sum just defined is obviously equal to the usual sum if I is finite, and we write

$$t = \sum_{i\in I} t_i.$$

From the definition of a summable family, we easily deduce an associativity property restricted to *finite groupings*, but that repeats infinitely.

Property 2.11. Let $\{t_i\}_{i\in I}$ be a summable family with sum t in T. Let K be a set of indices and $\{J_k\}_{k\in K}$ a *partition* of I where all the J_k are *finite* (that is, $I = \bigcup_{k\in K} J_k$ and the J_k are pairwise disjoint). Set $s_k = \sum_{i\in J_k} t_i$ for every k in K. Then the family $\{s_k\}_{k\in K}$ is summable with sum t.

As in the preceding chapters, we say that a family of series $\{r_i\}_{i\in I}$ is *locally finite* if for every m in M there is only a finite number of indices i such that (r_i, m) is different from 0_S.

Property 2.12. A locally finite family of power series is summable.

This simple property is a good example of what the topological structure placed on $S\langle\langle M\rangle\rangle$ imposes and adds. That we can *define a sum* for a locally

finite family of series is trivial: pointwise addition is defined for each m, independently of any assumption about M. To say that the family is *summable* is to add extra information: it ensures that partial sums converge to the result of pointwise addition.

For every series r, the family of series $\{(r, m)m \mid m \in M\}$, where m is identified with its characteristic series, is locally finite, and we have

$$r = \sum_{m \in M} (r, m)m,$$

which is the usual notation that is thus justified. We also deduce from this notation that $S\langle M\rangle$ is *dense* in $S\langle\langle M\rangle\rangle$. Property 2.12 extends beyond locally finite families and generalises to a proposition which links the summability of a family of series and that of families of coefficients.

Property 2.13. A family $\{r_i\}_{i \in I}$ of $S\langle\langle M\rangle\rangle$ is summable with sum r if and only if for each m in M, the family $\{(r_i, m)\}_{i \in I}$ of elements of S is summable with sum (r, m).

2.2 Rational Series

We are now ready to define the star operation on a series. We must nevertheless introduce here an assumption on the semiring, somehow an axiom of *infinite distributivity*. After that, the definition of rational series comes easily, the double definition indeed, one as a closure under rational operations and one by means of rational expressions which opens the way to effective computations.

2.2.1 Star of a Series

We start by considering the problem in arbitrary semirings and not only in the semirings of series.

Let t be an element of a topological semiring T; it is possible for the family $\{t^n\}_{n \in \mathbb{N}}$ to be, or not to be summable. If it is summable, we call its sum the 'star of t' and write it t^*:

$$t^* = \sum_{n \in \mathbb{N}} t^n.$$

Whether t^* is defined depends on t, on T, on the distance on T, or on a combination of all these elements. For example, $(0_T)^* = 1_T$ is defined for all T; if $T = \mathbb{Q}$, we have $(\frac{1}{2})^* = 2$ if $\mathbb{Q}$ is equipped with the natural topology, or undefined if the chosen topology is the discrete topology, while 1^* is not defined in either case.

Lemma 2.14. *Let T be a topological semiring and t an element of T whose star is defined. We have the double equality*

$$t^* = 1_T + tt^* = 1_T + t^*t. \tag{1}$$

Proof. We obviously have $t^{\leq n} = 1_T + tt^{<n} = 1_T + t^{<n}t$. As $\lim t^{<n} = \lim t^{\leq n} = t^*$, and as *addition and multiplication are continuous operations* on T, we obtain (1) by taking the limit of each side of the above equation. □

Remark 2.15. If T is a topological *ring*, and if the star of t is defined, (1) can be written $t^* - tt^* = t^* - t^*t = 1$ or $(1-t)t^* = t^*(1-t) = 1$ and so t^* is the *inverse* of $1 - t$. Hence, the classic identity

$$t^* = \frac{1}{1-t} = 1 + t + t^2 + \cdots, \tag{2}$$

is justified in full generality. It also means that forming the star can be considered as a substitute of taking the inverse in poor structure that has no inverse.

Star of a Proper Series

By reference to polynomials and to series in one variable, we call the *constant term* of a series r of $S\langle\langle M\rangle\rangle$ the coefficient of the neutral element of M in r: $\mathsf{c}(r) = (r, 1_M)$. A power series is called *proper* if its constant term is zero. The sum of two proper series is a proper series; the product of a proper series with any other series is a proper series, *since M is graded.*

If r is proper, the family $\{r^n \mid n \in \mathbb{N}\}$ is locally finite, and thus the star of a proper series of $S\langle\langle M\rangle\rangle$ is defined.

Lemma 2.16 (Arden). *Let r and u be two series of $S\langle\langle M\rangle\rangle$; if r is a proper series, each of the equations*

$$X = rX + u \quad \textit{and} \tag{3}$$
$$X = Xr + u \tag{4}$$

has a unique solution*: the series r^*u and ur^*, respectively.*

Proof. In (1), we replace t by r and multiply on the left (resp. on the right) by u and we obtain that r^*u (resp. ur^*) is a solution of (3) (resp. of (4)). Conversely, if v is a solution of the equation $X = u + rX$, we have

$$v = u + rv \quad \Longrightarrow \quad v = u + ru + r^2v = \cdots = r^{<n}u + r^nv,$$

for all integers n. Since r is proper, and multiplication continuous, we have $\lim r^n = \lim r^n v = 0$, from which follows $v = \lim(r^{<n}u) = (\lim r^{<n})u = r^*u$. □

From which, we deduce the following proposition.

Proposition 2.17. *Let r and u be two proper series of $S\langle\langle M\rangle\rangle$; the following equalities (or identities) hold:*

$$(r+u)^* = r^*(ur^*)^* = (r^*u)^*r^*, \tag{$\boldsymbol{S}$}$$
$$(ru)^* = 1 + r(ur)^*u, \tag{$\boldsymbol{P}$}$$
$$\forall n \in \mathbb{N} \quad r^* = r^{<n}(r^n)^*. \tag{$\boldsymbol{Z}_n$}$$

Following [12], the identity ($\boldsymbol{S}$) is called the *sum star identity* in Chap. 1, ($\boldsymbol{P}$) the *product star identity*.

Remark 2.18. It follows by Lemma 2.5 that a square matrix m of dimension Q with elements in $S\langle\langle M\rangle\rangle$ is a proper series of $S^{Q\times Q}\langle\langle M\rangle\rangle$ if all its elements are proper series; (we say in this case that m is proper), and hence that the identities ($\boldsymbol{S}$), ($\boldsymbol{P}$), and ($\boldsymbol{Z}_n$) are satisfied by proper matrices.

Strong Semirings and Star of an Arbitrary Series

The star of an arbitrary series, not necessarily proper, may or may not be defined. The following proposition allows us to tell the difference between the two cases. First, we make a timely definition to avoid a difficulty.

Definition 2.19. *A topological semiring is* strong *if the product of two summable families is a summable family; that is, if the two families* $\{r_i \mid i \in I\}$ *and* $\{u_j \mid j \in J\}$ *are summable with sum* s *and* t*, respectively, then the family* $\{r_i u_j \mid (i,j) \in I \times J\}$ *is summable with sum* st*.*

All the semirings which we shall consider are strong: semirings equipped with the discrete topology, the sub-semirings of $\mathbb{C}^n$ (equipped with the natural topology), and the positive semirings. We then easily verify the following property.

Property 2.20. The semirings of matrices and the semirings of series on a graded monoid, with coefficients in a strong semiring are strong.

Let r be a series of $S\langle\langle M\rangle\rangle$; the *proper part* of r is the proper series that coincides with r for all the elements m of M other than 1_M. It is convenient to write $r_0 = \mathsf{c}(r)$ for the constant term of r, and r_p for the proper part of r:

$$(r_\mathsf{p} 1_M) = 0_S \qquad \text{and} \qquad \forall m \in M\backslash 1_M \quad (r_\mathsf{p}, m) = (r, m),$$

and we write $r = r_0 + r_\mathsf{p}$ (rather than $r = r_0 1_M + r_\mathsf{p}$). These definitions and notations are taken in view of the following, which generalises to a series with coefficients in an arbitrary strong semiring, a result already established for series with coefficients in a continuous semiring.

Proposition 2.21. *Let* S *be a strong topological semiring and* M *a graded monoid. Let* r *be a series of* $S\langle\langle M\rangle\rangle$*,* r_0 *its constant term and* r_p *its proper part. Then* r^* *is defined if and only if* r_0^* *is defined and in this case we have*

$$r^* = (r_0^* r_\mathsf{p})^* r_0^* = r_0^* (r_\mathsf{p} r_0^*)^*. \tag{5}$$

Proof. The condition is necessary since $(r^n, 1_M) = r_0^n$ and, if r^* is defined, the coefficients of 1_M in $\{r^n\}_{n\in\mathbb{N}}$ form a summable family.

Conversely, assume that $\{r_0^n\}_{n\in\mathbb{N}}$ is summable, with sum r_0^*. For all pairs of integers k and l, set

$$P_{k,l} = \sum_{\substack{i_0,i_1,\dots,i_k \in \mathbb{N} \\ i_0+i_1+\cdots+i_k=l}} r_0^{i_0} r_\mathsf{p} r_0^{i_1} r_\mathsf{p} \cdots r_0^{i_{k-1}} r_\mathsf{p} r_0^{i_k}.$$

By convention, set $P_{0,l} = r_0^l$ and $P_{k,0} = r_\mathsf{p}^k$. We verify by inspection that, for all integers n,

$$r^n = (r_0 + r_\mathsf{p})^n = \sum_{l=0}^{l=n} P_{n-l,l}. \tag{6}$$

By induction on k, we will show that the family

$$F_k = \left\{ r_0^{i_0} r_\mathsf{p} r_0^{i_1} r_\mathsf{p} \cdots r_0^{i_{k-1}} r_\mathsf{p} r_0^{i_k} \;\middle|\; i_0, i_1, \dots, i_k \in \mathbb{N} \right\}$$

is summable in $S\langle\langle M \rangle\rangle$, with sum

$$Q_k = (r_0^* r_\mathsf{p})^k r_0^* = r_0^* (r_\mathsf{p} r_0^*)^k.$$

The ingredients of the proof are depicted in Fig. 1.

In fact, the hypothesis on r_0 ensures the property for $k = 0$, and also that the family $G = \{{r_0}^n r_\mathsf{p} \mid n \in \mathbb{N}\}$ is summable in $S\langle\langle M \rangle\rangle$, with sum ${r_0}^* r_\mathsf{p}$. The family F_{k+1} is the product of the families G and F_k and the assumption that S, and hence $S\langle\langle M \rangle\rangle$ is strong gives us the conclusion.

Hence, we deduce that, for each k, the family $\{P_{k,l} \mid l \in \mathbb{N}\}$ is summable, with sum Q_k. The family $\{Q_k \mid k \in \mathbb{N}\}$ is locally finite, hence summable, with sum

$$u = \sum_{k=0}^{\infty} Q_k = (r_0^* r_\mathsf{p})^* r_0^* = r_0^* (r_\mathsf{p} r_0^*)^*.$$

We can now easily finish the proof by showing that the 'doubly indexed' family $\{P_{k,l} \mid k, l \in \mathbb{N}\}$ is summable, with sum u. Equation (6) and Property 2.11 then ensure that the family $\{r^n \mid n \in \mathbb{N}\}$ is summable with sum u. □

The case of *cycle-free series* (see Chap. 1 and 3) falls in the scope of Proposition 2.21. In the same spirit as Remark 2.18, we note that (5) holds for every matrix m such that the star of its matrix of constant terms is defined. A particularly interesting case of this is where the matrix of constant terms is a strict upper triangular, another case of cycle-free series.

Proposition 2.22 (Bloom–Ésik [7]). *Let S be a strong topological semiring and M a graded monoid. Let r and u be series of $S\langle\langle M \rangle\rangle$ with constant terms r_0 and u_0, respectively, and such that r_0^*, u_0^*, and $(r_0 + u_0)^*$ are defined. Then the identities $(\boldsymbol{S})$, $(\boldsymbol{P})$, and $(\boldsymbol{Z}_n)$ hold for r and u.*

In other words, with the terminology of Chap. 1, and if one skips the question of the definition of star, if S is a Conway semiring, so is $S\langle\langle M \rangle\rangle$.

Remark 2.23. Along the line of Remark 2.15, it holds that if S is a ring, a series of $S\langle\langle M \rangle\rangle$ is invertible, if and only if its constant term is invertible.

For the rest of the chapter, S is a strong topological semiring.

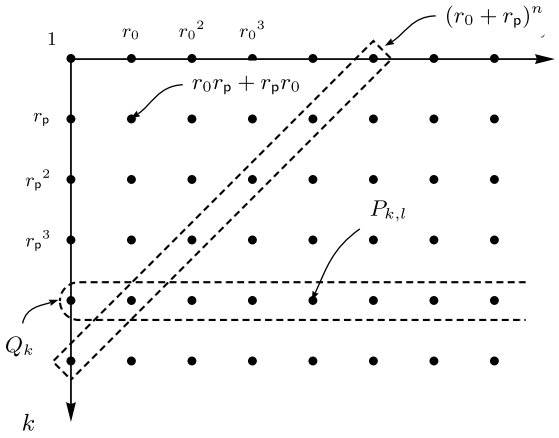


Fig. 1. A graphical representation of Proposition 2.21

2.2.2 The Family of Rational Series

We first characterise rational series 'from above' with the definition of rational operations and of closed families, and then inductively 'from below', with the definition of weighted rational expressions.

S-Rational Operations

The *rational operations* on $S\langle\langle M\rangle\rangle$ are:

(i) The *S-algebra operations*, that is:
- The two *exterior multiplications* by the elements of S
- The *addition*
- The *product*

(ii) The *star* operation, which is not defined everywhere.

Point (ii) leads us to tighten the notion of closure: a subset $\mathcal{E}$ of $S\langle\langle M\rangle\rangle$ is *closed under star* if s^* belongs to $\mathcal{E}$ for every series s in $\mathcal{E}$ such that s^* is defined.

A subset of $S\langle\langle M\rangle\rangle$ is *rationally closed* if it is closed under the rational operations; that is, if it is a subsemi-algebra of $S\langle\langle M\rangle\rangle$ closed under the star operation. The intersection of any family of rationally closed subsets is rationally closed, and thus the *rational closure* of a set $\mathcal{E}$ is the *smallest* rationally closed subset which contains $\mathcal{E}$, written $S\mathrm{Rat}\,\mathcal{E}$.

Definition 2.24. *A series of* $S\langle\langle M\rangle\rangle$ *is* S-rational *if it belongs to the rational closure of* $S\langle M\rangle$*, the set of polynomials on* M *with coefficients in* S*. The set of* S*-rational series (over* M *with coefficients in* S*) is written* $S\mathrm{Rat}\,M$*.*

If the monoid M is implied by the context, we shall say *S-rational series*, or just *rational series*, if S is also understood.

Example 2.25.

(i) Let M be the one-generator free monoid $\{x\}^*$ and S be a field $\mathbb{F}$. Then $\mathbb{F}\mathrm{Rat}\, x^*$ is exactly the set of series developments of ($\mathbb{F}$-)rational functions (that is, quotients of two polynomials) and this is where the name *rational*—rather the more common *regular* (for expressions and languages)—comes from.
(ii) If $S = \mathbb{B}$, we simply write $\mathrm{Rat}\, M$ for $\mathbb{B}\mathrm{Rat}\, M$ and its elements are the *rational subsets* of M.
(iii) If $S = \mathbb{N}$ and $M = \Sigma^* \times \Delta^*$, $\mathbb{N}\mathrm{Rat}\, \Sigma^* \times \Delta^*$ is the set of rational relations from Σ^* to Δ^* with multiplicity in $\mathbb{N}$, which we shall consider later.

Characteristic Series and Unambiguous Rational Sets

The notions introduced so far allow for a precise definition of unambiguity[6] (for rational sets) and some illustrative computations. For brevity, let us denote by $\underline{P}$ the *characteristic series* of a subset P of M (rather than by $\mathrm{char}(P)$ as in Chap. 1).

Definition 2.26. *Set $S = \mathbb{N}$ and let P and Q be two subsets of M.*

(i) The union $P \cup Q$ *is* unambiguous *if and only if* $\underline{(P \cup Q)} = \underline{P} + \underline{Q}$.
(ii) The product PQ *is* unambiguous *if and only if* $\underline{(PQ)} = \underline{P}\,\underline{Q}$.
(iii) The star *of P is* unambiguous *if and only if* $\underline{P^*} = (\underline{P})^*$.

A subset of M is unambiguously rational *if it belongs to the* unambiguous rational closure of finite subsets *of M. The family of unambiguous rational subsets of M is written* $\mathrm{URat}\, M$.

Then $P \in \mathrm{URat}\, M$ if, and only if $\underline{P} \in \mathbb{N}\mathrm{Rat}\, M$ and then $\underline{P} \in S\mathrm{Rat}\, M$ for any S. It is well known for instance that $\mathrm{URat}\, \Sigma^* = \mathrm{Rat}\, \Sigma^*$ and that $\mathrm{URat}(\Sigma^* \times \Delta^*)$ is strictly contained in $\mathrm{Rat}(\Sigma^* \times \Delta^*)$.

As Σ freely generates Σ^*, we have $\underline{(\Sigma)^*} = \underline{\Sigma^*}$, and thus $\underline{\Sigma^*} = \underline{\varepsilon} + \underline{\Sigma}\,\underline{\Sigma^*} = \underline{\varepsilon} + \underline{\Sigma^*}\,\underline{\Sigma}$ which gives $(\underline{\varepsilon} - \underline{\Sigma})\underline{\Sigma^*} = \underline{\Sigma^*}(\underline{\varepsilon} - \underline{\Sigma}) = \underline{\varepsilon}$, and thus $\underline{\Sigma^*} = (\underline{\varepsilon} - \underline{\Sigma})^{-1}$ if $S = \mathbb{Z}$.

If P is a non-empty prefix-closed subset of Σ^*, the *border of P* is the set:

$$C = P\Sigma \backslash P.$$

As an example, Fig. 2 shows the prefix-closed subset $\{\varepsilon, b, ba\}$ and its border $\{a, bb, baa, bab\}$.

Let P is a non-empty prefix-closed subset of Σ^* and let $h = pa$ with p in P and a in Σ (this is the unique expression of h in this form). There are two, mutually exclusive, possible cases: h is in C or h is in P. Conversely, every word of $P \cup C$ can be written in this way, except ε. Hence, we deduce the equality *between characteristic series*:

[6] A more or less folklore notion; an early reference for unambiguous rational sets is [15].

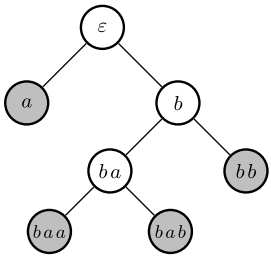


Fig. 2. A prefix-closed subset and its border

$$\underline{C} + \underline{P} = \underline{P\Sigma} + \underline{\varepsilon},$$

which we first rewrite as $\underline{\varepsilon} - \underline{C} = \underline{P}(\underline{\varepsilon} - \underline{\Sigma})$ then by right multiplication by $\underline{\Sigma^*} = (\underline{\varepsilon} - \underline{\Sigma})^{-1}$, as $\underline{\Sigma^*} - \underline{C\Sigma^*} = \underline{P}$. We thus have proved the following lemma.

Lemma 2.27. *Let P be a non-empty prefix-closed subset and $C = P\Sigma \backslash P$ its border. Every word f of $\Sigma^* \backslash P$ can be written uniquely as $f = cg$ with c in C and g in Σ^*.*

Rational S-Expressions

The definition of expressions will provide useful tools and handier ways to deal with rational series. Let $\{\mathsf{0}, \mathsf{1}, +, \cdot, *\}$ be five function symbols. Naturally, the functions $+$ and $\cdot$ are binary, $*$ is unary, and $\mathsf{0}$ and $\mathsf{1}$ are nullary (they represent constants). We define, for each s in S, *two unary functions*, also written s.

Definition 2.28. *A* weighted rational expression *over M with weight in S, or rational S-expression over M, is obtained inductively in the following manner:*

(i) $\mathsf{0}$, $\mathsf{1}$, *and m, for all m in M, are rational expressions (the* atomic *expressions).*
(ii) If E *is a rational expression and s is in S, then* $(s\mathsf{E})$ *and* $(\mathsf{E}s)$ *are rational expressions.*
(iii) If E *and* F *are rational expressions, then so are* $(\mathsf{E} + \mathsf{F})$, $(\mathsf{E} \cdot \mathsf{F})$, *and* (E^*).

We write $S\mathsf{RatE}\, M$ for the set of rational S-expressions over M.

Remark 2.29.

(i) We can restrict the atomic expressions, other than $\mathsf{0}$ and $\mathsf{1}$, to be elements g of any given generating set G of M without reducing the power of the definition. That is what we usually do when M is a free monoid Σ^*.
(ii) We could have considered the elements of S to be atoms and not operators, again without changing the power of the definition, and that would simplify somewhat some upcoming equations. The chosen way is, however, more consistent with the upcoming definition of the *derivation* of S-expressions over Σ^*.

We define the *depth* of an expression E, $\mathsf{d}(\mathsf{E})$, as the height of the syntactical tree of the corresponding expression:

$$\begin{gathered}\mathsf{d}(0) = \mathsf{d}(1) = \mathsf{d}(m) = 0, \quad \text{for all } m \text{ in } M,\\ \mathsf{d}((s\mathsf{E})) = \mathsf{d}((\mathsf{E}s)) = \mathsf{d}((\mathsf{E}^*)) = 1 + \mathsf{d}(\mathsf{E}),\\ \mathsf{d}((\mathsf{E}+\mathsf{F})) = \mathsf{d}((\mathsf{E}\cdot\mathsf{F})) = 1 + \mathsf{max}\big(\mathsf{d}(\mathsf{E}), \mathsf{d}(\mathsf{F})\big).\end{gathered}$$

The *constant term* of an expression E, $\mathsf{c}(\mathsf{E})$, is defined by induction on the depth of E; it is an element of S, computed by the following equations:

$$\begin{gathered}\mathsf{c}(1) = 1_S, \qquad \mathsf{c}(0) = \mathsf{c}(m) = 0_S \quad \text{for all } m \text{ in } M,\\ \mathsf{c}((s\mathsf{E})) = s\mathsf{c}(\mathsf{E}), \qquad \mathsf{c}((\mathsf{E}s)) = \mathsf{c}(\mathsf{E})s,\\ \mathsf{c}((\mathsf{E}+\mathsf{F})) = \mathsf{c}(\mathsf{E}) + \mathsf{c}(\mathsf{F}), \qquad \mathsf{c}((\mathsf{E}\cdot\mathsf{F})) = \mathsf{c}(\mathsf{E})\mathsf{c}(\mathsf{F}), \qquad \text{and}\\ \mathsf{c}((\mathsf{E}^*)) = \mathsf{c}(\mathsf{E})^* \quad \text{if the right-hand side is defined in } S.\end{gathered}$$

A rational S-expression may represent an element of $S\langle\langle M\rangle\rangle$ or not, the distinction between the two cases being made by the constant term, exactly as for the star of an arbitrary series and using that result. We shall say that an expression in $S\mathsf{Rat}\mathsf{E}\, M$ is *valid* if its constant term is defined. The series *denoted by a valid expression* E, which we write $|\mathsf{E}|$, is defined by induction on the depth of E by the equations

$$\begin{gathered}|0| = 0_S, \qquad |1| = 1_M, \qquad |m| = m \quad \text{for all } m \text{ in } M,\\ |(s\mathsf{E})| = s|\mathsf{E}|, \qquad |(\mathsf{E}s)| = |\mathsf{E}|s,\\ |(\mathsf{E}+\mathsf{F})| = |\mathsf{E}| + |\mathsf{F}|, \qquad |(\mathsf{E}\cdot\mathsf{F})| = |\mathsf{E}||\mathsf{F}|, \qquad \text{and} \qquad |(\mathsf{E}^*)| = |\mathsf{E}|^*.\end{gathered}$$

We verify both that these equations are well defined and that they are consistent, in the sense that the *constant term of the expression* E is the *constant term of the series* $|\mathsf{E}|$, *in parallel*, and *in the same induction*, using Proposition 2.21. In other words, and in order to define $|\mathsf{E}|$, we shall also have proved the following.

Property 2.30. For all valid S-expressions E in $S\mathsf{Rat}\mathsf{E}\, M$, $\mathsf{c}(\mathsf{E}) = (|\mathsf{E}|, 1_M)$.

Example 2.31. Take $M = \{a, b\}^*$ and $S = \mathbb{Q}$. The $\mathbb{Q}$-expression $(a^* + (-1b^*))^*$ is valid, as is $\mathsf{E}_1 = (\frac{1}{6}a^* + \frac{1}{3}b^*)^*$ since $\mathsf{c}(\frac{1}{6}a^* + \frac{1}{3}b^*) = \frac{1}{2}$, and hence $\mathsf{c}(\mathsf{E}_1) = 2$ is defined; $(a^* + b^*)^*$ is not valid.

The set of series denoted by valid S-expressions is rationally closed, and every rationally closed subset of $S\langle\langle M\rangle\rangle$ that contains every element of M (and thus $S\langle M\rangle$) contains every series denoted by a valid S-expression, which proves the following proposition.

Proposition 2.32. *A series of $S\langle\langle M\rangle\rangle$ is S-rational if and only if it is denoted by a valid rational S-expression over M.*

3 Weighted Automata

An *automaton over M with weight (or* with multiplicity*) in S*, or *S-automaton*[7] *over M* is a *graph* labelled with elements of $S\langle\langle M\rangle\rangle$, associated with two maps from the set of vertices to $S\langle\langle M\rangle\rangle$. We develop and complete this definition. We build on the identification of a graph with its incidence matrix and the proofs will be performed systematically with matrix computations. The essence of an automaton, however, remains that of a graph and the behaviour of on automaton is defined in the language of graphs. We also continue to use the graph representation and its vocabulary to aid intuition.

3.1 The Behaviour of a Weighted Automaton

An automaton $\mathfrak{A}$ over M with weights in S is specified by the choice of the following:[8]

- A non-empty set Q of *states* of $\mathfrak{A}$, also called the *dimension* of $\mathfrak{A}$.
- An element E of $S\langle\langle M\rangle\rangle^{Q\times Q}$, a square matrix of dimension Q with entries in $S\langle\langle M\rangle\rangle$, called the *transition matrix* of $\mathfrak{A}$; we can view each entry $E_{p,q}$ different from 0_S as the label of a unique edge which goes from state p to state q in the graph with vertices Q and we write $p \xrightarrow{x} q$, or $p \xrightarrow[\mathfrak{A}]{x} q$, if $x = E_{p,q}$. (If $E_{p,q} = 0_S$, we consider there to be *no* edge from p to q.)
- Two elements I and T of $S\langle\langle M\rangle\rangle^{Q}$; that is, two functions I and T from Q to $S\langle\langle M\rangle\rangle$: I is the *initial function* and T the *final function* of $\mathfrak{A}$; they can also be seen as vectors of dimension Q: I is a *row vector* and T a *column vector*, called respectively the *initial vector* and *final vector* of $\mathfrak{A}$.

The S-automaton $\mathfrak{A}$ is written, naturally enough,

$$\mathfrak{A} = \langle I, E, T\rangle.$$

We use the familiar conventions to represent S-automata graphically (see figures below); the values of I labelling the incoming arrows and those of T the outgoing arrows.

A *path* in $\mathfrak{A}$ is a sequence of transitions such that the source of each is the destination of the previous one; it can be written

$$c := p_0 \xrightarrow{x_1} p_1 \xrightarrow{x_2} p_2 \xrightarrow{x_3} \cdots \xrightarrow{x_n} p_n.$$

The *label*, or *result* of c, written $|c|$, is the *product* of the labels of the transitions of c. In the above case, $|c| = x_1x_2\cdots x_n$.

A *computation* in $\mathfrak{A}$ is a path to which is added an arrow arriving at the source and one leaving from the destination, with their respective labels. The computation corresponding to the above path is hence

[7] Or *weighted automaton* if S is understood or immaterial.

[8] This definition is a priori more general than the one given in Chap. 3; the two will coincide for finite automata.

$$d := \xrightarrow{I_{p_0}} p_0 \xrightarrow{x_1} p_1 \xrightarrow{x_2} p_2 \xrightarrow{x_3} \cdots \xrightarrow{x_n} p_n \xrightarrow{T_{p_n}} .$$

The *label* or *result* of d, still written $|d|$, is the product of the label of the incoming arrow, that of the path, and that of the outgoing arrow, in that order; in our case: $|d| = I_{p_0} x_1 x_2 \cdots x_n T_{p_n}$.

The definitions we have made for weighted automata are indeed a generalisation of the classical definitions:

(i) An automaton over Σ is a $\mathbb{B}$-automaton over Σ^*; an automaton over M is a $\mathbb{B}$-automaton over M.

(ii) The distinction between *path* and *computation*, which are often used as synonyms, may seem useless. But apart from the fact that it is consistent with our terminology—'path' refers to 'graph' while 'computation' refers to 'automaton', and what distinguishes an automaton from a graph is precisely that initial and final states are taken into account—it was only introduced in order to make precise definitions that incorporate the generality that we have now allowed for I and T. In the majority of cases, the non-zero elements of I and T will be scalar (that is, elements of S), usually equal to 1_S and the two notions will coincide.

(iii) Along the same lines, the disappearance of the notion of a *successful computation* is merely apparent. A state p such that the component I_p is *non-zero* (that is, different from $0_{S\langle\langle M\rangle\rangle}$) can be called *initial*, and a state where T_p is non-zero can be called *final*. We can then say that a computation is successful if its source is an initial state and its destination is a final state.

Definition 3.1. *The* behaviour *of an automaton* $\mathfrak{A} = \langle I, E, T\rangle$ *of finite dimension* Q *is defined if and only if for all* p *and* q *in* Q *the family of labels of paths with source* p *and destination* q *is summable. In this case, the family of labels of computations of* $\mathfrak{A}$ *is summable and its sum is the behaviour of* $\mathfrak{A}$*, written*[9] $|\mathfrak{A}|$*. We also say that* $\mathfrak{A}$ accepts *or* realises *the series* $|\mathfrak{A}|$.

The description of the transitions of an automaton by a matrix is justified by the fact that a walk over a graph corresponds to a matrix multiplication. This is expressed by the following proposition.

Lemma 3.2. *Let* $\mathfrak{A} = \langle I, E, T\rangle$ *be an* S*-automaton over* M *of finite dimension. For every integer* n*,* E^n *is the matrix of the sums of the labels of paths of length* n*.*

Proof. By induction on n. The assertion is true for $n = 1$ (and also for $n = 0$ by convention). The definition of the $(n+1)$st power of E is

$$\forall p, q \in Q \quad \left(E^{n+1}\right)_{p,q} = \sum_{r \in Q} \left(E^n\right)_{p,r} E_{r,q}.$$

[9] Written $\|\mathfrak{A}\|$ in Chap. 3.

Every path of length $n+1$ is the concatenation of a path of length n with a path of length 1, that is, a single transition. We can therefore write[10]

$$\begin{aligned}&\{c \mid c := p \underset{\mathfrak{A}}{\to} q,\ l(c) = n+1\}\\ &\quad = \bigcup_{r \in Q} \{(d,e) \mid d := p \underset{\mathfrak{A}}{\to} r,\ l(d) = n,\ e := r \underset{\mathfrak{A}}{\to} q \in E\},\end{aligned}$$

and hence

$$\begin{aligned}&\sum \{|c| \mid c := p \underset{\mathfrak{A}}{\to} q,\ l(c) = n+1\}\\ &\quad = \sum_{r \in Q} \left(\{|d||e| \mid d := p \underset{\mathfrak{A}}{\to} r,\ l(d) = n,\ e := r \underset{\mathfrak{A}}{\to} q \in E\}\right)\\ &\quad = \sum_{r \in Q} \left[\left(\sum \{|d| \mid d := p \underset{\mathfrak{A}}{\to} r,\ l(d) = n\}\right) E_{r,q}\right].\end{aligned}$$

As $\sum\{|d| \mid d = p \underset{\mathfrak{A}}{\to} r,\ l(d) = n\} = (E^n)_{p,r}$ by the induction hypothesis, the lemma is proved. □

Since the sum of the results of the *computations* of length n is equal by definition to the product $I \cdot E^n \cdot T$, and since the behaviour of $\mathfrak{A}$ is equal to the sum of the results of the computations of all the lengths, the following statement holds.

Corollary 3.3. *Let $\mathfrak{A} = \langle I, E, T\rangle$ be a S-automaton of finite dimension whose behaviour is defined, then E^* is defined and we have $|\mathfrak{A}| = I \cdot E^* \cdot T$.*

Example 3.4. The $\mathbb{N}$-automaton over $\{a,b\}^*$ defined by

$$\mathfrak{B}_1 = \left\langle (1\ 0), \begin{pmatrix} a+b & b \\ 0 & a+b \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$$

is shown in Fig. 3 (left). A simple calculation allows us to determine its behaviour:

$$\forall f \in \Sigma^* \quad (|\mathfrak{B}_1|, f) = |f|_b; \qquad \text{that is} \qquad |\mathfrak{B}_1| = \sum_{f \in \Sigma^*} |f|_b f = u_1.$$

Another $\mathbb{N}$-automaton is shown in Fig. 3 (right)

$$\mathfrak{C}_1 = \left\langle (1\ 0), \begin{pmatrix} a+b & b \\ 0 & 2a+2b \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle.$$

If we use the convention that each word f of Σ^* is considered as a number written in binary, interpreting a as the digit 0 and b as the digit 1, and if we

[10] The length of a path c is here written $l(c)$.

Fig. 3. The $\mathbb{N}$-automata $\mathfrak{B}_1$ and $\mathfrak{C}_1$

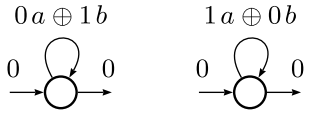


Fig. 4. The $\mathbb{M}$-automaton $\mathfrak{S}_1$

write $\bar{f}$ for the integer represented by the word f, it is easy to verify that $\bar{f}$ is computed by $\mathfrak{C}_1$ in the sense that

$$\forall f \in \Sigma^* \quad (|\mathfrak{C}_1|, f) = \bar{f}; \qquad \text{that is,} \qquad |\mathfrak{C}_1| = \sum_{f \in \Sigma^*} \bar{f} f.$$

Example 3.5. To illustrate the case where S is *different* from $\mathbb{N}$: let $\mathbb{M} = \langle \mathbb{N}^\infty, \mathsf{min}, +, \infty, 0 \rangle$ be the 'tropical' semiring (cf. Chap. 1, Sect. 2). The $\mathbb{M}$-automaton $\mathfrak{S}_1$ over $\{a, b\}^*$ and defined by

$$\mathfrak{S}_1 = \left\langle (0\ 0), \begin{pmatrix} 0a + 1b & \infty \\ \infty & 1a + 0b \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle$$

is shown in Fig. 4. Clearly, the support of $|\mathfrak{S}_1|$ is all of $\{a, b\}^*$ and the coefficient in $|\mathfrak{S}_1|$ of an arbitrary word f of $\{a, b\}^*$ is $\mathsf{min}\{|f|_a, |f|_b\}$.

Remark 3.6. The behaviour of an automaton was defined by returning to the essence of an 'automaton': a procedure for describing computations. With this definition, the behaviour of the two automata in Fig. 5(a), (b) are not defined although in the first case the family $\{I \cdot E^n \cdot T\}_{n \in \mathbb{N}}$ is summable since all its terms are zero, and in the second E^* is defined since $E^2 = 0$.

Such a definition of the behaviour is more 'robust' than one that would be based on the transition matrix and its star only. For instance, it is invariant under the decomposition of a transition into a strictly longer path. Figure 6

Fig. 5. Two $\mathbb{Z}$-automata with behavioural problems

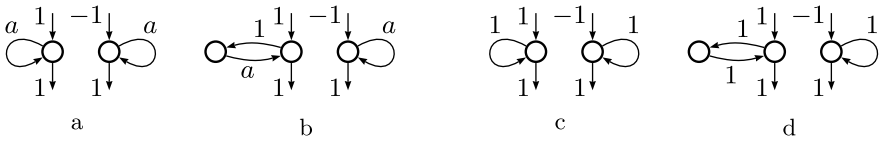


Fig. 6. Advocating for a path-based definition of the behaviour of automata

illustrates this point: as the automaton in (a) is obviously equivalent to the one in (b), those in (c) and (d) should also be equivalent.

In Chap. 3, the behaviour of an automaton is defined under the assumption that the automaton is cycle-free. Under the same assumption, the behaviour—as defined here—is always well defined and, by Corollary 3.3, equal to the one defined in Chap. 3.

Remark 3.7. On the other hand, these examples also lead us to note that the transition between each pair of states p and q must be *unique*, and labelled $E_{p,q}$; otherwise, we would be able to 'decompose' these entries in such a way that the family of labels of paths would no longer be summable.

From Lemma 2.5 and Proposition 2.21, we deduce a sort of generalisation of the same Proposition 2.21.

Proposition 3.8. *Let S be a strong topological semiring and M a graded monoid. The behaviour of an S-automaton over M, $\mathfrak{A} = \langle I, E, T\rangle$ is defined if and only if the behaviour of the S-automaton $\mathfrak{A}_0 = \langle I, E_0, T\rangle$ is defined, where E_0 is the matrix of constant terms of entries of E, and in that case we have*

$$|\mathfrak{A}| = I \cdot ({E_0}^* \cdot E_{\mathfrak{p}})^* \cdot {E_0}^* \cdot T.$$

The example of Fig. 5(b) shows that it is not sufficient that ${E_0}^*$ be defined, nor even that E_0 be nilpotent[11] for the behaviour of $\mathfrak{A}$ be defined. On the other hand, the behaviour of $\mathfrak{A}$ is defined when E_0 is *strict upper triangular* since in this case the number of computations in $\mathfrak{A}_0$ is *finite*. And this is the case (up to a renaming of the states) if the automaton is cycle-free.

Definition 3.9. *A S-automaton over M, $\mathfrak{A} = \langle I, E, T\rangle$, is* finite *if:*

(i) The dimension of $\mathfrak{A}$ is finite.
(ii) The coefficients of E, I and T are polynomials; that is, have finite support.

3.2 The Fundamental Theorem of Automata

One hesitates to say of a proposition, 'here is *the fundamental theorem*'. However, this seems justified for the one that follows: it states completely generally,

[11] That is, there exists an n such that ${E_0}^n = 0$.

at least under the current assumption that M is a finitely generated graded monoid and S a strong topological semiring that what one can 'do' with a finite automaton is precisely what one can 'do' with rational operations.

Theorem 3.10. *A series of $S\langle\langle M\rangle\rangle$ is rational if and only if it is the behaviour of some finite S-automaton over M.*

Remark 3.11. Theorem 3.10 is usually called Kleene's theorem, and again in this handbook (cf. Chap. 3). When M is a free monoid Σ^*, there is no possibility to distinguish between *rational* and *recognisable* sets or series, but at the level of speech. When M is not free, recognisable sets or series take their own quality and become a distinct family from the one of rational sets, or series. We thus have two distinct results: the first one (Theorem 3.10) that states that in any graded monoid the elements of *one* certain family—for which there is no reason to coin *two* different names—may have two distinct characterisations: by rational expressions and by finite automata and another one (Theorem 4.6 below) that states that two families of sets or series, which are distinct in general, coincide in the case of free monoids.

Since every language of Σ^* is the behaviour of an *unambiguous* automaton (of a deterministic one indeed)—we quoted above that $\mathrm{URat}\,\Sigma^* = \mathrm{Rat}\,\Sigma^*$—we then have the following.

Proposition 3.12. *The characteristic series of a rational language of Σ^* is a S-rational series, for any semiring S.*

3.2.1 Proper Automata

We can make Theorem 3.10 both more precise and more general, closer to the properties used in the proof. For this, we need to define a restricted class of S-automata.

Definition 3.13. *An S-automaton over M, $\mathfrak{A} = \langle I, E, R\rangle$, is* proper *if:*

(i) The matrix E is proper.
(ii) The entries of I and T are scalar*; that is, $I \in S^{1\times Q}$ and $T \in S^{Q\times 1}$.*

It follows from Proposition 3.8 that the behaviour of a proper automaton is well defined; the following result adds the converse.

Proposition 3.14. *Every S-automaton $\mathfrak{A}$ over M whose behaviour is defined is equivalent to a proper automaton whose entries, other than the scalar entries of the initial and final vectors, are linear combinations of proper parts of the entries of $\mathfrak{A}$.*

Proof. We first show that $\mathfrak{A} = \langle I, E, T\rangle$ is equivalent to an automaton $\mathfrak{B} = \langle J, F, U\rangle$ where the entries of J and U are scalar. We set

$$J = \left(1 \boxed{\begin{matrix} 0 \end{matrix}} 0\right), \qquad F = \begin{pmatrix} 0 & \boxed{I} & 0 \\ \boxed{0} & E & \boxed{T} \\ 0 & \boxed{0} & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 0 \\ \boxed{0} \\ 1 \end{pmatrix}. \tag{7}$$

Every path in $\mathfrak{B}$ is a path or a computation in $\mathfrak{A}$ and the behaviour of $\mathfrak{B}$ is defined if and only if that of $\mathfrak{A}$ is, and in that case E^* is defined.[12] We verify by induction that, for every integer n greater than or equal to 2,

$$F^n = \begin{pmatrix} 0 & I \cdot E^{n-1} & I \cdot E^{n-2} \cdot T \\ 0 & E^n & E^{n-1} \cdot T \\ 0 & 0 & 0 \end{pmatrix}. \tag{8}$$

We have $J \cdot U = J \cdot F \cdot U = 0$, $J \cdot F^{n+2} \cdot U = I \cdot E^n \cdot T$, hence $J \cdot F^* \cdot U = I \cdot E^* \cdot T$ and $\langle J, F, U \rangle$ is equivalent to $\mathfrak{A}$.

Next, starting from an automaton $\mathfrak{B} = \langle J, F, U \rangle$ whose initial and final vectors are scalar, we set

$$F = F_0 + F_{\mathfrak{p}}.$$

The behaviour of $\mathfrak{B}$ is defined if and only if the behaviour of the automaton $\langle J, F_0, U \rangle$ is defined, and in this case F_0^* is defined, also. We then have

$$|\mathfrak{B}| = J \cdot F^* \cdot U = J \cdot H^* \cdot V,$$

with $H = F_0^* \cdot F_{\mathfrak{p}}$ and $V = F_0^* \cdot U$. Since F_0^* is an element of $S^{Q \times Q}$, the entries of H are linear combinations (with coefficients in S) of entries of $F_{\mathfrak{p}}$ and the entries of V are scalar. □

3.2.2 Standard Automata

It is convenient to define an even more restricted class of automata and to show that an automaton of that class can be canonically associated with every S-expression.

Definition 3.15. *An S-automaton $\mathfrak{A} = \langle I, E, T \rangle$ is* standard *if the initial vector I has a single non-zero coordinate i, equal to 1_S, and if this unique initial state i is not the destination of any transition whose label is non-zero.*

In matrix terms, this means that $\mathfrak{A}$ can be written

$$\mathfrak{A} = \left\langle \left(1 \boxed{\begin{matrix} 0 \end{matrix}}\right), \begin{pmatrix} 0 & \boxed{K} \\ \boxed{0} & \boxed{F} \end{pmatrix}, \begin{pmatrix} c \\ \boxed{U} \end{pmatrix} \right\rangle. \tag{9}$$

[12] The automaton $\mathfrak{B}$ is the *normalised automaton* $\mathfrak{A}'$ built in Chap. 3 (proof of Theorem 2.11).

The definition does not forbid the initial state i from also being final, that is, the scalar c is not necessarily zero. If $\mathfrak{A}$ is not only standard but also *proper*, c is the *constant term* of $|\mathfrak{A}|$. The proof of Proposition 3.14 itself proves the following proposition.

Proposition 3.16. *Every S-automaton $\mathfrak{A}$ over M whose behaviour is defined is equivalent to a* standard *proper automaton whose entries, other than the scalar entries of the initial and final vectors, are linear combinations of proper parts of the entries of $\mathfrak{A}$.*

We now define *operations* on standard automata (as in Chap. 3, Sect. 2.2) that are parallel to the *rational operations*. Let $\mathfrak{A}$ (as in (9)) and $\mathfrak{A}'$ (with obvious translation) be two proper standard automata; the following standard S-automata are defined:

- $$s\mathfrak{A} = \left\langle \begin{pmatrix} 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & sK \\ 0 & F \end{pmatrix}, \begin{pmatrix} sc \\ U \end{pmatrix} \right\rangle \quad \text{and}$$
$$\mathfrak{A}s = \left\langle \begin{pmatrix} 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & K \\ 0 & F \end{pmatrix}, \begin{pmatrix} cs \\ Us \end{pmatrix} \right\rangle$$
- $$\mathfrak{A} + \mathfrak{A}' = \left\langle \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & K & K' \\ 0 & F & 0 \\ 0 & 0 & F' \end{pmatrix}, \begin{pmatrix} c+c' \\ U \\ U' \end{pmatrix} \right\rangle$$
- $$\mathfrak{A} \cdot \mathfrak{A}' = \left\langle \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & K & cK' \\ 0 & F & H \\ 0 & 0 & F' \end{pmatrix}, \begin{pmatrix} cc' \\ V \\ U' \end{pmatrix} \right\rangle$$
where $H = (U \cdot K') \cdot F'$ and $V = Uc' + (U \cdot K') \cdot U'$
- $$\mathfrak{A}^* = \left\langle \begin{pmatrix} 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & c^*K \\ 0 & G \end{pmatrix}, \begin{pmatrix} c^* \\ Uc^* \end{pmatrix} \right\rangle$$
which is defined if and only if c^* is defined, and where $G = U \cdot c^*K + F$.

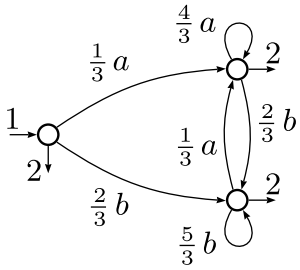


Fig. 7. The $\mathbb{Q}$-automaton $\mathcal{S}_{\mathsf{E}_1}$

By construction, $s\mathfrak{A}$, $\mathfrak{A}s$, $\mathfrak{A}+\mathfrak{A}'$, $\mathfrak{A}\cdot\mathfrak{A}'$, and $\mathfrak{A}^*$ are all proper. Straightforward computations show that $|s\mathfrak{A}| = s|\mathfrak{A}|$, $|\mathfrak{A}s| = |\mathfrak{A}|s$, $|\mathfrak{A}+\mathfrak{A}'| = |\mathfrak{A}| + |\mathfrak{A}'|$, $|\mathfrak{A}\cdot\mathfrak{A}'| = |\mathfrak{A}||\mathfrak{A}'|$ and $|\mathfrak{A}^*| = |\mathfrak{A}^*|$.

With every valid rational S-expression E, we thus canonically associate, by induction on the depth of E, a proper standard S-automaton $\mathcal{S}_{\mathsf{E}}$ that we call *the* standard automaton of E. Let $\ell(\mathsf{E})$ denote the *literal length* of E, that is, the number of atoms different from $\mathsf{0}$ and $\mathsf{1}$ in E. The following proposition holds.

Proposition 3.17. *If* E *is a valid rational* S*-expression, then* $|\mathcal{S}_{\mathsf{E}}| = |\mathsf{E}|$ *and the dimension of* $\mathcal{S}_{\mathsf{E}}$ *is* $\ell(\mathsf{E}) + 1$.

Example 3.18 (Example 2.31 continued). Figure 7 shows the $\mathbb{Q}$-automaton $\mathcal{S}_{\mathsf{E}_1}$ associated with the rational expression $\mathsf{E}_1 = (\frac{1}{6}a^* + \frac{1}{3}b^*)^*$ by the construction described above.

3.2.3 Statement and Proof of the Fundamental Theorem

Definition 3.19. *We will say that a family of series is* proper *if it contains the proper part of each of its elements.*[13]

In particular, the polynomials form a proper family of $S\langle\langle M\rangle\rangle$.

Theorem 3.20. *Let* $\mathcal{C}$ *be a proper family of series of* $S\langle\langle M\rangle\rangle$. *A series* s *of* $S\langle\langle M\rangle\rangle$ *belongs to* $S\mathrm{Rat}\,\mathcal{C}$ *if and only if* s *is the behaviour of a* proper standard S*-automaton over* M *of finite dimension whose (non-scalar) entries are finite linear combinations of elements of* $\mathcal{C}$.

Proof. The proof of Theorem 3.20 splits in the "if" and "only if" parts, which by Proposition 2.32, essentially amount to show respectively that given a proper automaton we can compute an equivalent valid rational expression and conversely that given a valid rational expression we can compute an equivalent automaton.

[13] As opposed to all the series in the family being proper.

We write $\mathcal{D}$ for the family of behaviours of proper standard S-automata whose entries are linear combinations of elements of $\mathcal{C}$. We first show that $\mathcal{D}$ contains 0_S, behaviour of the standard automaton $\langle 1_S, 0_S, 0_S\rangle$ of dimension 1 and 1_S, behaviour of $\langle 1_S, 0_S, 1_S\rangle$, as well as every element in $\mathcal{C}$: for r in $\mathcal{C}$, r_p is in $\mathcal{C}$ since $\mathcal{C}$ is a proper family and it holds:

$$r = \begin{pmatrix} 1_S & 0_S \end{pmatrix} \cdot \begin{pmatrix} 0_S & r_\mathsf{p} \\ 0_S & 0_S \end{pmatrix}^* \cdot \begin{pmatrix} r_0 \\ 1_S \end{pmatrix}.$$

If $\mathfrak{A}$ and $\mathfrak{A}'$ are two proper standard S-automata whose entries are linear combinations of elements of $\mathcal{C}$, the above constructions $s\mathfrak{A}$, $\mathfrak{A}s$, $\mathfrak{A} + \mathfrak{A}'$, $\mathfrak{A} \cdot \mathfrak{A}'$ and $\mathfrak{A}^*$ show that $\mathcal{D}$ is rationally closed.

Conversely, we start from a proper automaton $\mathfrak{A} = \langle I, E, T\rangle$ whose behaviour is thus defined and equal to $|\mathfrak{A}| = I \cdot E^* \cdot T$. This part then amounts to prove that the entries of the star of a proper matrix E belong to the rational closure of the entries of E, a classical statement established in general under different hypotheses (e.g. [12]). Since we have to reprove it anyway, we choose a slightly different method. We write $|\mathfrak{A}| = I \cdot V$ with $V = E^* \cdot T$. Since E is proper and by Lemmas 2.5 and 2.16, V is *the unique solution* of

$$X = E \cdot X + T \tag{10}$$

and we have to prove that all entries of the vector V belong to the rational closure of the entries of E. Lemma 2.16 already states that the property holds if $\mathfrak{A}$ is of dimension 1. For $\mathfrak{A}$ of dimension Q, we write (10) as a system of $\|Q\|$ equations:

$$\forall p \in Q \quad V_p = \sum_{q \in Q} E_{p,q} V_q + T_p. \tag{11}$$

We choose (arbitrarily) one element q in Q and by Lemma 2.16 again, it comes:

$$V_q = E_{q,q}^* \Big[\sum_{p \in Q\setminus\{q\}} E_{q,p} V_p + T_q \Big],$$

an expression for V_q that can be substituted in every other equation of the system (11), giving a new system

$$\forall p \in Q\setminus\{q\} \quad V_p = \sum_{r \in Q\setminus\{q\}} [E_{p,r} + E_{p,q} E_{q,q}^* E_{q,r}] V_r + E_{p,q} E_{q,q}^* T_q + T_p.$$

And the property is proved by induction hypothesis. □

The fundamental theorem states the equality of two families of series (infinite objects), but its proof is better understood as the description of two algorithms. Here, we have chosen on one hand the construction of the standard automaton of an expression and on the other hand the algorithm known as the *state elimination method* for the computation of an expression denoting

the behaviour of an automaton. In the latter case, the result depends on the order of elimination (the choice of the state q in (11)). The relationship between the possible different results is given by the following Proposition 3.21. We shall say that two (S-)expressions E and F are *equivalent modulo an identity* $\boldsymbol{I}$ if E can be transformed into F by using instances of $\boldsymbol{I}$ and of the so-called 'natural identities' which express that the expressions are interpreted in a semiring (associativity, distributivity of $\cdot$ over $+$, commutativity of $+$).

Proposition 3.21. *Let $\mathfrak{A}$ be an S-automaton of dimension Q. The expressions denoting $|\mathfrak{A}|$ and obtained by the state elimination method with distinct orders on Q are all equivalent modulo the identities $\boldsymbol{S}$ and $\boldsymbol{P}$.*

3.3 Conjugacy and Covering of Automata

After the definition of any structure, one looks for *morphisms* between objects of that structure, and weighted automata are no exception. Moreover, morphisms of graphs and, therefore, of classical Boolean automata, are not less classical, and one waits for their generalisation to weighted automata. Taking into account multiplicity proves, however, to be not so simple. In the sequel, all automata are supposed to be of finite dimension.

3.3.1 From Conjugacy to Covering

We choose to describe the morphisms of weighted automata, which we call *coverings*, via the notion of *conjugacy*, borrowed from the theory of symbolic dynamical systems.

Definition 3.22. *An S-automaton $\mathfrak{A} = \langle I, E, T\rangle$* is conjugate to *an S-automaton $\mathfrak{B} = \langle J, F, U\rangle$ if there exists a matrix X with entries in S such that*

$$IX = J, \qquad EX = XF, \qquad \text{and} \qquad T = XU.$$

The matrix X is the transfer matrix *of the conjugacy and we write $\mathfrak{A} \overset{X}{\Longrightarrow} \mathfrak{B}$.*

In spite of the idea conveyed by the terminology, the conjugacy relation *is not an equivalence* but a *pre-order* relation. Suppose that $\mathfrak{A} \overset{X}{\Longrightarrow} \mathfrak{C}$ holds; if $\mathfrak{C} \overset{Y}{\Longrightarrow} \mathfrak{B}$, then $\mathfrak{A} \overset{XY}{\Longrightarrow} \mathfrak{B}$, but if $\mathfrak{B} \overset{Y}{\Longrightarrow} \mathfrak{C}$ then $\mathfrak{A}$ is not necessarily conjugate to $\mathfrak{B}$, and we write $\mathfrak{A} \overset{X}{\Longrightarrow} \mathfrak{C} \overset{Y}{\Longleftarrow} \mathfrak{B}$ or even $\mathfrak{A} \overset{X}{\Longrightarrow}\overset{Y}{\Longleftarrow} \mathfrak{B}$. This being well understood, we shall speak of "conjugate automata" when the orientation does not matter.

As $JF^nU = IXF^nU = IEXF^{n-1}U = \cdots = IE^nXU = IE^nT$ for every integer n, the following proposition holds.

Proposition 3.23. *Two conjugate automata are equivalent.*

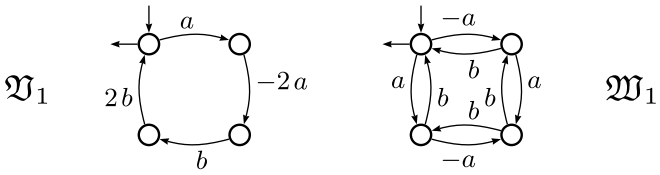


Fig. 8. Two conjugate $\mathbb{Z}$-automata

Example 3.24. It is easily checked that the $\mathbb{Z}$-automaton $\mathfrak{V}_1$ of Fig. 8 is conjugate to the $\mathbb{Z}$-automaton $\mathfrak{W}_1$ of the same figure with the transfer matrix X_1:

$$X_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Let φ be an *equivalence* relation on Q or what is the same, let $\varphi\colon Q \to R$ be a surjective map and H_φ the $Q \times R$-matrix where the (q, r) entry is 1 if $\varphi(q) = r$, 0, otherwise. Since φ is a map, each row of H_φ contains exactly one 1 and since φ is surjective, each column of H_φ contains at least one 1. Such a matrix is called an *amalgamation matrix* [31, Definition 8.2.4].

Definition 3.25. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two S-automata of dimension Q and R, respectively. We say that $\mathfrak{B}$ is* a S-quotient *of $\mathfrak{A}$ and conversely that $\mathfrak{A}$ is* a S-covering *of $\mathfrak{B}$ if there exists a surjective map $\varphi\colon Q \to R$ such that $\mathfrak{A}$ is conjugate to $\mathfrak{B}$ by H_φ.*

The notion of S-quotient is *lateralised* since the conjugacy relation is not symmetric. Somehow, it is the price we pay for extending the notion of morphism to S-automata. Therefore, the *dual* notions *co-S-quotient* and *co-S-covering* are defined in a natural way.

Definition 3.26. *With the above notation, we say that $\mathfrak{B}$ is* a co-S-quotient *of $\mathfrak{A}$ and conversely that $\mathfrak{A}$ is* a co-S-covering *of $\mathfrak{B}$ if there exists a surjective map $\varphi\colon Q \to R$ such that $\mathfrak{B}$ is conjugate to $\mathfrak{A}$ by ${}^{\mathrm{t}}H_\varphi$.*

We also write $\varphi\colon \mathfrak{A} \to \mathfrak{B}$ and call φ, by way of metonymy, *a S-covering, or a co-S-covering from $\mathfrak{A}$ onto $\mathfrak{B}$.*

Example 3.27. Consider the $\mathbb{N}$-automaton $\mathfrak{C}_2$ of Fig. 9 and the map φ_2 from $\{j, r, s, u\}$ to $\{i, q, t\}$ such that $j\varphi_2 = i$, $u\varphi_2 = t$ and $r\varphi_2 = s\varphi_2 = q$, then

$$H_{\varphi_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and φ_2 is an $\mathbb{N}$-covering from $\mathfrak{C}_2$ onto $\mathfrak{V}_2$ and a co-$\mathbb{N}$-covering from $\mathfrak{C}_2$ onto $\mathfrak{V}'_2$.

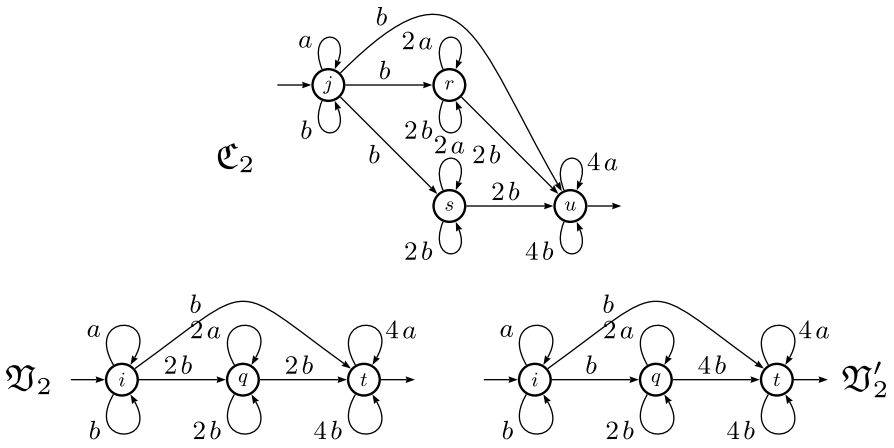


Fig. 9. $\mathfrak{C}_2$ is an $\mathbb{N}$-covering of $\mathfrak{V}_2$ and a co-$\mathbb{N}$-covering of $\mathfrak{V}_2$

3.3.2 Minimal S-Quotient

Let us first express that in a S-covering $\varphi\colon \mathfrak{A} \to \mathfrak{B}$ the image is somewhat immaterial and only counts the map equivalence of φ. From any amalgamation matrix H_φ, we construct a matrix K_φ by transposing H_φ and by arbitrarily cancelling certain entries in such a way that K_φ is row monomial (with exactly one 1 per row); K_φ is not uniquely determined by φ, but also depends on the choice of a 'representative' in each class for the map equivalence of φ. Whatever K_φ, the product $K_\varphi H_\varphi$ is the identity matrix of dimension R (as the matrix representing $\varphi^{-1}\varphi$). Easy matrix computations establish the following.

Proposition 3.28. *Let $\mathfrak{A} = \langle I, E, T\rangle$ and $\mathfrak{B} = \langle J, F, U\rangle$ be two S-automata of dimension Q and R, respectively. A surjective map $\varphi\colon Q \to R$ is* a S-covering *if and only if $\mathfrak{A}$ satisfies the two equations:*

$$H_\varphi \cdot K_\varphi \cdot E \cdot H_\varphi = E \cdot H_\varphi, \tag{12}$$

and

$$H_\varphi \cdot K_\varphi \cdot T = T. \tag{13}$$

In which case, $\mathfrak{B}$ satisfies

$$F = K_\varphi \cdot E \cdot H_\varphi, \qquad J = I \cdot H_\varphi \qquad \text{and} \qquad U = K_\varphi \cdot T. \tag{14}$$

Theorem 3.29. *Let $\mathfrak{A}$ be a S-automaton of finite dimension over M. Among all the S-quotients of $\mathfrak{A}$ (resp. among all the co-S-quotients of $\mathfrak{A}$), there exists one, unique up to isomorphism and effectively computable from $\mathfrak{A}$, which has a minimal number of states and of which all these S-automata are S-coverings (resp. co-S-coverings).*

Proof. A surjective map $\varphi\colon Q \to R$ defines a S-covering $\varphi\colon \mathfrak{A} \to \mathfrak{B}$ if (12) and (13) (which do not involve $\mathfrak{B}$) are satisfied.

To prove the existence of a minimal S-quotient, it suffices to show that if $\varphi\colon Q \to R$ and $\psi\colon Q \to P$ are two maps that define S-coverings, the map $\omega\colon Q \to V$ also defines a S-covering, where $\omega = \varphi \vee \psi$ is the map whose map equivalence is the upper bound of those of φ and ψ; that is, the finest equivalence which is coarser than the map equivalences of φ and ψ. In other words, there exist $\varphi'\colon R \to V$ and $\psi'\colon P \to V$ such that $\omega = \varphi\varphi' = \psi\psi'$ and each class modulo $\omega = \varphi \vee \psi$ can be seen at the same time as a union of classes modulo φ and as a union of classes modulo ψ. It follows that

$$E \cdot H_\omega = E \cdot H_\varphi \cdot H_{\varphi'} = E \cdot H_\psi \cdot H_{\psi'}; \tag{15}$$

and if two states p and r of Q are congruent modulo ω, there exists q such that $p\varphi = q\varphi$ and $q\psi = r\psi$ (in fact, a sequence of states q_i, etc.). The rows p and q of $E \cdot H_\varphi$ are equal, and the rows q and r of $E \cdot H_\psi$ are equal; hence, by (15), the rows p and r of $E \cdot H_\omega$ are equal, also.

To compute this minimal S-quotient, we can proceed by successive refinements of partitions, exactly as for the computation of the minimal automaton of a language from a deterministic automaton which recognises the language.

In what follows, the maps φ_i are identified with their map equivalences; the image is irrelevant. A state r of Q is identified with the row vector of dimension Q, characteristic of r, and treated as such. For example, $r\varphi = s\varphi$ can be written $r \cdot H_\varphi = s \cdot H_\varphi$.

The maps φ_0 have the same map equivalence as T, that is,

$$r \cdot H_{\varphi_0} = s \cdot H_{\varphi_0} \quad \Longleftrightarrow \quad r \cdot T = s \cdot T,$$

which can also be written

$$H_{\varphi_0} \cdot K_{\varphi_0} \cdot T = T, \tag{16}$$

and the same equation holds for every map finer than φ_0. For each i, φ_{i+1} is finer than φ_i and, by definition, r and s are joint in φ_i (that is, $r \cdot H_{\varphi_i} = s \cdot H_{\varphi_i}$) and disjoint in φ_{i+1} if $r \cdot E \cdot H_{\varphi_i} \neq s \cdot E \cdot H_{\varphi_i}$. Let j be the index such that $\varphi_{j+1} = \varphi_j$, that is, such that

$$r \cdot H_{\varphi_j} = s \cdot H_{\varphi_j} \quad \Longrightarrow \quad r \cdot E \cdot H_{\varphi_j} = s \cdot E \cdot H_{\varphi_j}, \tag{17}$$

which can be rewritten

$$H_{\varphi_j} \cdot K_{\varphi_j} \cdot E \cdot H_{\varphi_j} = E \cdot H_{\varphi_j}. \tag{18}$$

By (16) and (18), φ_j is a S-covering.

Conversely, every S-covering ψ satisfies (13) and is hence finer than φ_0. Then for all i, if ψ is finer than φ_i, it must also be finer than φ_{i+1}. In fact, if r and s are joint in ψ, it follows that $r \cdot H_\psi = s \cdot H_\psi$, and hence also $r \cdot H_{\varphi_i} = s \cdot H_{\varphi_i}$ since φ_i is coarser than ψ, and hence r and s are joint in φ_{i+1}: ψ is finer than φ_j, which is thus the coarsest S-covering. □

Remark 3.30. Even if the minimal S-quotient of a S-automaton and the minimal automaton of a language are computed with the *same* algorithm, they are nevertheless fundamentally different: the second automaton is canonically associated with the language, whereas the first is associated with the S-automaton we started from, and not with its behaviour.

Remark 3.31. The above construction applies of course if $S = \mathbb{B}$, and thus shows that the notion of minimal ($\mathbb{B}$-)quotient is well defined even for a *nondeterministic automaton* (as we just wrote, this minimal quotient is not associated with the recognised language anymore). Moreover, it can be checked that two Boolean automata are *bisimilar* if and only if their minimal $\mathbb{B}$-quotients are isomorphic (cf. [2]).

3.3.3 From Covering to Conjugacy

We have defined quotients (and co-quotients) as a special case of conjugacy. Under some supplementary hypothesis—that is naturally met in cases that are important to us: $\mathbb{N}$, $\mathbb{Z}$, etc.—it can be established that a kind of converse holds and that any conjugacy can basically be realised by the composition of an inverse co-covering and a covering.

In order to state these results, we need two further definitions. A matrix is *non-degenerate* if it contains no zero row nor zero column. We call *a circulation matrix* a diagonal invertible matrix.

Theorem 3.32 ([3]). *Let $\mathfrak{A}$ be a $\mathbb{Z}$-automaton conjugate to a $\mathbb{Z}$-automaton $\mathfrak{B}$ by a non-negative and non-degenerate transfer matrix X. Then there exists a $\mathbb{Z}$-automaton $\mathfrak{C}$ that is a co-$\mathbb{Z}$-covering of $\mathfrak{A}$ and a $\mathbb{Z}$-covering of $\mathfrak{B}$.*

We can free ourselves from the two hypotheses on the transfer matrix if we allow a further conjugacy by a circulation matrix.

Theorem 3.33 ([3]). *Let $\mathfrak{A}$ be a $\mathbb{Z}$-automaton conjugate to a $\mathbb{Z}$-automaton $\mathfrak{B}$ by a transfer matrix X. Then there exists two $\mathbb{Z}$-automata $\mathfrak{C}$ and $\mathfrak{D}$ and a circulation matrix D, such that $\mathfrak{C}$ is a co-$\mathbb{Z}$-covering of $\mathfrak{A}$, $\mathfrak{D}$ a $\mathbb{Z}$-covering of $\mathfrak{B}$ and $\mathfrak{C}$ is conjugate to $\mathfrak{D}$ by D.*

Example 3.34 (Example 3.24 continued). The $\mathbb{Z}$-automata $\mathfrak{X}_1$ of Fig. 10 is a co-$\mathbb{Z}$-covering of $\mathfrak{V}_1$, $\mathfrak{Y}_1$ is a $\mathbb{Z}$-covering of $\mathfrak{W}_1$, and $\mathfrak{X}_1$ is conjugate to $\mathfrak{Y}_1$ by the circulation matrix where the only -1 entry is at state 1.

The proof of Theorem 3.33 involves indeed two properties. Let us say first that a semiring *has property* (SU) if *every element is a sum of units*. The semiring $\mathbb{N}$, the ring $\mathbb{Z}$, and all fields have property (SU). In any semiring with (SU), every matrix X can be written as $X = CDR$ where C is a co-amalgamation, R an amalgamation, and D a circulation matrix. In $\mathbb{Z}$, the dimension of D will be the sum of the absolute value of the entries of X.

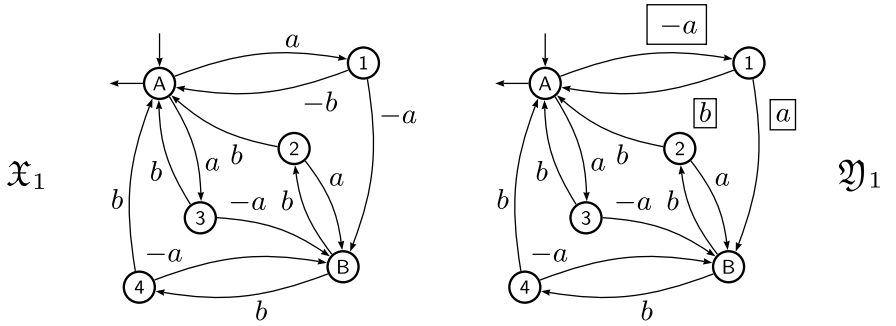


Fig. 10. The co-covering and covering of $\mathfrak{V}_1$ and $\mathfrak{W}_1$

Having secured the existence of C, D, and R, the second step consists in building the automata $\mathfrak{C}$ and $\mathfrak{D}$ that will fit in. To that end, we say that a commutative monoid is *equisubtractive* if for all p, q, r, and s such that $p+q=r+s$ there exist x, y, z, and t such that $p=x+y$, $q=z+t$, $r=x+z$ and $s=y+t$. A semiring is equisubtractive if it is so as a monoid for addition.

The semirings $\mathbb{N}$ and $\mathbb{Z}$ are equisubtractive, and if S is equisubtractive, then so are $S\langle\Sigma^*\rangle$ and $S\langle\langle\Sigma^*\rangle\rangle$. The construction of $\mathfrak{C}$ and $\mathfrak{D}$ will then follow from the following property.

Lemma 3.35. *Let S be an equisubtractive semiring and let $t_1, t_2, \ldots, t_n, s_1, s_2, \ldots, s_m$ be elements of S such that*

$$t_1 + t_2 + \cdots + t_n = s_1 + s_2 + \cdots + s_m.$$

There exists an $n \times m$ matrix G with entries in S such that the sum of the entries of each row i is equal to t_i and the sum of the entries of each column j is equal to s_j.

Another consequence of the definition of equi-subtractive semiring and of Lemma 3.35 is to allow a sort of converse to Theorem 3.29. The existence of a minimal S-covering implies a kind of Church–Rosser property: if we have two diverging arrows, that is, the *upper part* of a commutative diagram, we can construct the lower part of it. The following proposition states that it is possible to complete a commutative diagram when the *lower part* of it is known.

Proposition 3.36 ([43, 3]). *Let S be an* equisubtractive *semiring and let $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ be three S-automata.*

(a) If $\mathfrak{A}$ and $\mathfrak{B}$ are S-coverings of $\mathfrak{C}$ (resp. co-S-coverings of $\mathfrak{C}$), there exists a S-automaton $\mathfrak{D}$ which is a S-covering (resp. a co-S-covering) of both $\mathfrak{A}$ and $\mathfrak{B}$.

(b) If $\mathfrak{A}$ is a S-covering of $\mathfrak{C}$ and $\mathfrak{B}$ is a co-S-covering of $\mathfrak{C}$, there exists a S-automaton $\mathfrak{D}$ which is both a co-S-covering of $\mathfrak{A}$ and a S-covering of $\mathfrak{B}$.

4 Recognisable Series and Representations

As in the last section, S denotes a strong topological semiring and M a graded monoid, a priori arbitrary. We shall now consider another family of series of $S\langle\langle M\rangle\rangle$, other than $S\mathrm{Rat}\, M$, but that coincide with it when M is a free monoid Σ^*: this is the Kleene–Schützenberger theorem (Theorem 4.6). We first define these series by means of representations. We then consider the Hadamard product of series, which is a weighted generalisation of intersection. In a third subsection, by considering the series over a Cartesian product of monoids, we briefly sketch the prolegomena to a theory of weighted relations. This allows us, among other things, to establish the weighted generalisation of results on the morphic image of rational sets (Theorem 4.35).

4.1 The Family of Recognisable Series

An *S-representation of M of dimension Q* is a *morphism* μ from M to the semiring of square matrices of dimension Q with entries in S. By definition, in fact so that we can multiply the matrices, the dimension Q is *finite*. An S-representation of M (of dimension Q) is also the name we give a *triple* (λ, μ, ν) where, as before,

$$\mu : M \to S^{Q\times Q}$$

is a morphism and where λ and ν are two vectors:

$$\lambda \in S^{1\times Q} \quad \text{and} \quad \nu \in S^{Q\times 1};$$

that is, λ is a *row* vector and ν a *column* vector of dimension Q, with entries in S. Such a representation defines a map from M to S by

$$\forall m \in M \quad m \mapsto \lambda \cdot m\mu \cdot \nu;$$

that is, the *series* r:

$$r = \sum_{m\in M} (\lambda \cdot m\mu \cdot \nu)m.$$

A series r of $S\langle\langle M\rangle\rangle$ is *realised* or *recognised* by the representation (λ, μ, ν). We also say that (λ, μ, ν) *realises* or *recognises* the series r.

Definition 4.1. *A series of $S\langle\langle M\rangle\rangle$ is* S-recognisable *if it is recognised by an S-representation. The set of S-recognisable series over M is written $S\mathrm{Rec}\, M$.*

Example 4.2 (Example 3.4 continued). Take $S = \mathbb{N}$ and $M = \{a, b\}^*$. Let $(\lambda_1, \mu_1, \nu_1)$ be the representation defined by

$$a\mu_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad b\mu_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \lambda_1 = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad \text{and} \qquad \nu_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For all f in $\{a, b\}^*$, we verify that $\lambda_1 \cdot f\mu_1 \cdot \nu_1 = |f|_b$, hence the series $u_1 = \sum_{f\in\Sigma^*} |f|_b f$ is $\mathbb{N}$-recognisable.

Remark 4.3. It is not difficult to check that Definition 4.1 coincides, for $S = \mathbb{B}$, with the definition of the *recognisable subsets* of a monoid *as the sets that are saturated by a congruence of finite index* [14]. If r is a $\mathbb{B}$-recognisable series over M, realised by the representation (λ, μ, ν), then $\mu : M \to \mathbb{B}^{Q\times Q}$ is a morphism from M *to a finite monoid*. The series r of $\mathbb{B}\langle\langle M\rangle\rangle$, $r = \sum_{m\in M}(\lambda \cdot m\mu \cdot \nu)m$, can be seen as the *subset* $r = P\mu^{-1}$ of M where $P = \{p \in \mathbb{B}^{Q\times Q} \mid \lambda \cdot p \cdot \nu = 1_{\mathbb{B}}\}$. Conversely, a morphism α from M into a finite monoid N is a morphism from M into the monoid of Boolean matrices of dimension N (the representation of N by right translations over itself) and the $\mathbb{B}$-representation that realises any subset recognised by α easily follows.

These definitions and the following two properties of $S\mathrm{Rec}\, M$ do not involve multiplication in $S\langle\langle M\rangle\rangle$, and are hence valid without even requiring that M be graded.

Proposition 4.4. *Every finite linear combination, with coefficients in S, of S-recognisable series over M is an S-recognisable series.*

Proof. Let r and u be two S-recognisable series over M, respectively recognised by the S-representations (λ, μ, ν) and (η, κ, ζ). For all s in S, the series sr is recognised by the representation $(s\lambda, \mu, \nu)$, the series rs by the representation $(\lambda, \mu, \nu s)$, and the series $r + u$ by the representation (δ, π, ξ) defined by the following block decomposition:

$$\delta = \begin{pmatrix}\lambda & \eta\end{pmatrix}, \qquad m\pi = \begin{pmatrix} m\mu & 0 \\ 0 & m\kappa \end{pmatrix}, \qquad \xi = \begin{pmatrix}\nu \\ \zeta\end{pmatrix}. \quad \square$$

Let $\varphi\colon S \to T$ be a morphism of semirings which extends to a morphism $\varphi\colon S\langle\langle M\rangle\rangle \to T\langle\langle M\rangle\rangle$ by $(r\varphi, m) = (r, m)\varphi$ for all r in $S\langle\langle M\rangle\rangle$ and all m in M. If (λ, μ, ν) is a representation of the series r of $S\langle\langle M\rangle\rangle$, then $(\lambda\varphi, \mu\varphi, \nu\varphi)$ is a representation of $r\varphi$. That is:

Proposition 4.5. *Let $\varphi\colon S \to T$ be a morphism of semirings. The image under φ of an S-recognisable series over M is a T-recognisable series over M.*

We can now get to our main point.

Theorem 4.6 (Kleene–Schützenberger). *Let S be a strong topological semiring, and Σ a finite alphabet. A series of $S\langle\langle\Sigma^*\rangle\rangle$ is S-rational if and only if it is S-recognisable. That is,*

$$S\mathrm{Rec}\,\Sigma^* = S\mathrm{Rat}\,\Sigma^*.$$

We prove the two inclusions one at a time:

$$S\mathrm{Rec}\,\Sigma^* \subseteq S\mathrm{Rat}\,\Sigma^* \quad \text{and} \quad S\mathrm{Rat}\,\Sigma^* \subseteq S\mathrm{Rec}\,\Sigma^*. \tag{19}$$

Each of the inclusions is obtained from the Fundamental Theorem together with the freeness of Σ^* and the finiteness of Σ. This is used in both cases by means of the following result.

Lemma 4.7. *Let S be a semiring and Σ a finite alphabet. Let Q be a finite set and $\mu\colon \Sigma^* \to S^{Q\times Q}$ a morphism. We set*

$$X = \sum_{a\in\Sigma} (a\mu)a.$$

Then for all f in Σ^*, we have $(X^*, f) = f\mu$.

Proof. The matrix X is a proper series of $S^{Q\times Q}\langle\langle\Sigma^*\rangle\rangle$, and hence X^* is defined. We first prove, by induction on the integer n, that

$$X^n = \sum_{f\in\Sigma^n} (f\mu)f,$$

an equality trivially verified for $n = 0$, and true by definition for $n = 1$. It follows that

$$X^{n+1} = X^n \cdot X = \left(\sum_{f\in\Sigma^n} (f\mu)f\right) \cdot \left(\sum_{a\in\Sigma} (a\mu)a\right) = \sum_{(f,a)\in\Sigma^n\times\Sigma} (f\mu \cdot a\mu)fa$$
$$= \sum_{(f,a)\in\Sigma^n\times\Sigma} (fa)\mu fa = \sum_{g\in\Sigma^{n+1}} (g\mu)g,$$

since, for each integer n, Σ^{n+1} is in bijection with $\Sigma^n \times \Sigma$ as Σ^* is freely generated by Σ. For the same reason, Σ^* is the *disjoint* union of the Σ^n, for n in $\mathbb{N}$, and it follows, for all f in Σ^*, that

$$(X^*, f) = \left(X^{|f|}, f\right) = f\mu. \quad \square$$

Proof (of Theorem 4.6). Each of the two inclusions (19) is proved in the form of a property.

Property 4.8. If Σ is finite, S-recognisable series on Σ^* are S-rational.

Proof. Let (λ, μ, ν) be a representation which recognises a series r; that is, $(r, f) = \lambda \cdot f\mu \cdot \nu$, for all f in Σ^*. Let $\langle\lambda, X, \nu\rangle$ be the automaton defined by

$$X = \sum_{a\in\Sigma} (a\mu)a.$$

By Lemma 4.7, we have

$$r = \sum_{f\in\Sigma^*} (\lambda \cdot f\mu \cdot \nu)f = \lambda \cdot \left(\sum_{f\in\Sigma^*} (f\mu)f\right) \cdot \nu = \lambda \cdot X^* \cdot \nu.$$

By the Fundamental Theorem, the series r belongs to the rational closure of the entries of X. These entries are finite linear combinations of elements of Σ since Σ *is finite*: r belongs to $S\mathrm{Rat}\,\Sigma^*$. $\square$

Property 4.9. The S-rational series on Σ^* are S-recognisable.

Proof. By Theorem 3.20, the series r is the behaviour of a *proper* finite S-automaton $\langle I, X, T\rangle$, such that the entries of X are finite linear combinations of elements of Σ (and those of I and T are scalar). We can therefore write $X = \sum_{a\in\Sigma}(a\mu)a$ where $a\mu$ is the matrix of coefficients of the letter a in X. By Lemma 4.7, we have

$$\forall f \in \Sigma^* \quad (r, f) = (I \cdot X^* \cdot T, f) = I \cdot f\mu \cdot T,$$

and the series r is recognised by the *representation* (I, μ, T). □

The two inclusions (19) prove the theorem. □

4.2 Other Products on Recognisable Series

The two products that we shall now consider, the Hadamard and shuffle products are defined on general series—the second one for series on a free monoid—but it is their effect on recognisable series which will interest us, and we first define a product on *representations*.

4.2.1 Tensor Product of S-Representations

The *tensor product* of matrices has been defined in Chap. 1. Let A be a matrix of dimension $P \times P'$ and B a matrix of dimension $R \times R'$ (with entries in the same semiring S); the tensor product of A by B written $A \otimes B$ is a matrix of dimension $(P \times R) \times (P' \times R')$ defined by

$$\forall p \in P,\ \forall p' \in P',\ \forall r \in R,\ \forall r' \in R' \quad A \otimes B_{(p,r),(p',r')} = A_{p,p'}B_{r,r'}.$$

If S is *commutative*, the tensor product is also. We shall need the tensor product to be commutative under more general assumptions. We shall say that two sub-semirings U and V of a non-commutative semiring S are *commutable* if every element of U commutes with every element of V. For example, the *centre* of S and any sub-semiring of S are commutable. As another example, $1_T \times T$ and $T \times 1_T$ are two commutable sub-semirings[14] in $T \times T$. The following result has already been quoted (Chap. 1, Theorem 4.7).

Lemma 4.10. *Let A, B, C, and D be four matrices with entries in S, respectively of dimension $P \times Q$, $P' \times Q'$, $Q \times R$, and $Q' \times R'$, and such that all the entries of B commute with those of C. Then*

$$(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D).$$

It then follows:

[14] On the other hand, we shall not say that two matrices A and B are *commutable* to mean that all the entries of A commute with those of B; this would be too easily confused with the fact that the two matrices *commute*, that is, $AB = BA$.

Proposition 4.11 (Tensor product of representations). *Let U and V be two commutable sub-semirings of S. Let M and N be two arbitrary monoids and $\mu : M \to U^{Q\times Q}$ and $\kappa : N \to V^{R\times R}$ two representations. The map $\mu \otimes \kappa$, defined for all (m, n) in $M \times N$ by*

$$(m, n)\,[\mu \otimes \kappa] = m\mu \otimes n\kappa$$

is a representation of $M \times N$ in $S^{(Q\times R)\times(Q\times R)}$.

Proof. For all (m, n) and (m', n') in $M \times N$, we have

$$\begin{aligned}\big((m, n)[\mu \otimes \kappa]\big) \cdot \big((m', n')[\mu \otimes \kappa]\big) &= (m\mu \otimes n\kappa) \cdot (m'\mu \otimes n'\kappa)\\ &= (m\mu \cdot m'\mu) \otimes (n\kappa \cdot n'\kappa)\\ &= (mm')\mu \otimes (nn')\kappa = (mm', nn')[\mu \otimes \kappa],\end{aligned}$$

since under the proposition's assumptions, all the entries of $m'\mu$ commute with those of $n\kappa$. □

4.2.2 Hadamard Product

The Hadamard product is to series (sets with multiplicity) what intersection is to sets, which only really makes sense if the semiring of coefficients is commutative. In the same way that the recognisable subsets of an arbitrary monoid are closed under intersection, we have the following.

Theorem 4.12. *Let S be a commutative semiring and M an arbitrary monoid. Then $S\mathrm{Rec}\, M$ is closed under the Hadamard product.*

Under the more precise assumptions of Proposition 4.11, we can state a more general result.

Theorem 4.13 (Schützenberger). *Let U and V be two commutable sub-semirings of S and M a monoid. The Hadamard product of a U-recognisable series over M and a V-recognisable series over M is an S-recognisable series over M.*

More precisely, if (λ, μ, ν) recognises r and (η, κ, ζ) recognises u, then $r \odot u$ is recognised by $(\lambda \otimes \eta, \mu \otimes \kappa, \nu \otimes \zeta)$.

Proof. First note that, since the map $m \mapsto (m, m)$ is a morphism from M to $M \times M$, Proposition 4.11 implies that the map $m \mapsto m\mu \otimes m\kappa$ is also a morphism, and we also write it $\mu \otimes \kappa$.

Let r be a series over M recognised by the U-representation (λ, μ, ν) and u be a series over M recognised by the V-representation (η, κ, ζ). By definition, we have for all m in M,

$$(r \odot u, m) = (\lambda \cdot m\mu \cdot \nu)(\eta \cdot m\kappa \cdot \zeta) = (\lambda \cdot m\mu \cdot \nu) \otimes (\eta \cdot m\kappa \cdot \zeta),$$

the second equality expressing the product of two coefficients of S as the tensor product of two 1×1 matrices. Under the assumptions of the theorem, we can apply Lemma 4.10 (three times) and obtain

$$(r \odot u, m) = (\lambda \otimes \eta) \cdot (m\mu \otimes m\kappa) \cdot (\nu \otimes \zeta) = (\lambda \otimes \eta) \cdot (m[\mu \otimes \kappa]) \cdot (\nu \otimes \zeta)\,.$$

Again, according to these assumptions, $\mu \otimes \kappa$ is an S-representation, the series $r \odot u$ is recognisable, and is recognised by the stated representation. □

As a consequence of Theorem 4.6, *the Hadamard product of two S-rational series on Σ^* is an S-rational series* (if S is a commutative semiring). Moreover, the tensor product of representations of Σ^* translates directly into a construction on S-automata over Σ^* whose labels are linear combinations of letters of Σ, which is the natural generalisation of the Cartesian product of automata, and which we can call the *Hadamard product* of S-automata.

Example 4.14. The $\mathbb{N}$-automaton $\mathfrak{C}_2$ of Fig. 9 is the Hadamard product of the $\mathbb{N}$-automaton $\mathfrak{C}_1$ of Fig. 3 by itself. Therefore, for every f in Σ^*, it holds $f|\mathfrak{C}_2| = \overline{f}^2$.

4.2.3 Shuffle Product

We now suppose that M is a free monoid Σ^* and that S is commutative (usually $S = \mathbb{N}$ but that is not required). The shuffle product (or Hurwitz product) of two words of Σ^*, and then by linearity of two series in $S\langle\langle\Sigma^*\rangle\rangle$, has been defined at Chap. 1, mostly for ancillary purposes. Let us recall this definition as the interest of which goes far beyond the computations it was used for so far.

Definition 4.15. *For all f and g in Σ^*, the* shuffle *of f and g, written $f \between g$, is an homogeneous polynomial of $S\langle\Sigma^*\rangle$ defined by induction on $|f| + |g|$ by*

$$\forall f \in \Sigma^* \quad f \between \varepsilon = \varepsilon \between f = f,$$
$$\forall f, g \in \Sigma^*,\ \forall a, b \in A \quad fa \between gb = (fa \between g)b + (f \between gb)a,$$

The shuffle is extended 'by linearity' to $S\langle\langle\Sigma^*\rangle\rangle$, that is,

$$\forall r, u \in S\langle\langle\Sigma^*\rangle\rangle \quad r \between u = \sum_{f,g\in\Sigma^*} (r, f)(u, g) f \between g,$$

which is defined since the family of polynomials $f \between g$ for f and g in Σ^* is locally finite.

Example 4.16.

$$ab \between ab = 4aabb + 2abab,$$
$$ab \between ba = abab + 2abba + 2baab + baba \quad \text{and}$$
$$(\varepsilon + a) \between a^* = [a^*]^2.$$

Shuffle is an associative, commutative, and continuous product and makes of $S\langle\langle \Sigma^* \rangle\rangle$ a commutative S-algebra. The shuffle of two words is characterised by the following.

Proposition 4.17. *Let $\chi\colon \Sigma^* \to S\langle \Sigma^* \times \Sigma^* \rangle$ be the morphism (of monoids) defined by $a\chi = (a, \varepsilon) + (\varepsilon, a)$, for all a in Σ^*. It then follows that*

$$\forall h \in \Sigma^* \quad h\chi = \sum_{f,g \in \Sigma^*} (f \shuffle g, h)(f, g).$$

Theorem 4.18. *Let S be a commutative semiring. The shuffle of two S-recognisable series on Σ^* is an S-recognisable series.*

Proof. Let r and u be S-recognisable series on Σ^*, respectively recognised by the S-representations (λ, μ, ν) and (η, κ, ζ). For all h in Σ^*, the definition yields

$$\begin{aligned}
(r \shuffle u, h) &= \sum_{f,g \in \Sigma^*} \big((r, f)(u, g)\big)(f \shuffle g, h) \\
&= \sum_{f,g \in \Sigma^*} \big((\lambda \cdot f\mu \cdot \nu)(\eta \cdot g\kappa \cdot \zeta)\big)(f \shuffle g, h) \\
&= \sum_{f,g \in \Sigma^*} \big((\lambda \otimes \eta) \cdot \big((f, g)[\mu \otimes \kappa]\big) \cdot (\nu \otimes \zeta)\big)(f \shuffle g, h) \\
&= (\lambda \otimes \eta) \cdot \big((h\chi)[\mu \otimes \kappa]\big) \cdot (\nu \otimes \zeta) \quad \text{by Proposition 4.17.}
\end{aligned}$$

By the theorem's assumptions, $\chi \circ [\mu \otimes \kappa]$ is an S-representation; the series $r \shuffle u$ is recognisable. □

A consequence of Theorem 4.6 again, *the shuffle of two S-rational series on Σ^* is an S-rational series* (if S is a commutative semiring). As for the Hadamard product, the construction on representations that underlies the proof of Theorem 4.18 translates into a construction on S-automata over Σ^*, which we can call the *shuffle product* of S-automata.

Formally, if $\mathfrak{A}' = \langle Q', \Sigma, E', I', T' \rangle$ and $\mathfrak{A}'' = \langle Q'', \Sigma, E'', I'', T'' \rangle$ are two proper S-automata over Σ^* whose labels are linear combinations of letters of Σ, the shuffle of $|\mathfrak{A}'|$ and $|\mathfrak{A}''|$ is realised by the S-automaton written $\mathfrak{A}' \shuffle \mathfrak{A}''$ and defined by

$$\mathfrak{A}' \shuffle \mathfrak{A}'' = \langle Q' \times Q'', \Sigma, E, I' \otimes I'', T' \otimes T'' \rangle ,$$

where the set E of transitions is described by

$$\begin{aligned}
E = &\big\{ \big((p', p''), k'a, (q', p'')\big) \bigm| (p', k'a, q') \in E' \text{ and } p'' \in Q'' \big\} \\
&\cup \big\{ \big((p', p''), k''a, (p', q'')\big) \bigm| p' \in Q' \text{ and } (p'', k''a, q'') \in E'' \big\}.
\end{aligned}$$

Example 4.19. The $\mathbb{Z}$-automaton $\mathfrak{W}_1$ of Fig. 8 is the shuffle product of the obvious two state $\mathbb{Z}$-automata that respectively accept $(ab)^*$ and $(-ab)^*$. The equivalence with $\mathfrak{V}_1$ in the same figure yields the identity[15]

$$(ab)^* \between (-ab)^* = \left(-4a^2b^2\right)^*. \tag{20}$$

4.3 Series on a Product of Monoids

Series on a (Cartesian) product of monoids is a major subject in itself and their study could occupy a whole chapter of this book: they are the behaviour of *transducers with multiplicity*, of interest both from a theoretical and applications point of view (cf. Chaps. 7, 11, and 14, for instance). Here, we confine ourselves to few definitions and results stemming from the canonical isomorphisms between several semirings of series and with the aim of being able to state (and to prove) results about the image of series under morphisms and of comparing the families of rational and recognisable series.

4.3.1 The Canonical Isomorphisms

Polynomials or series in several (commutative) variables can be ordered with respect to one or another variable. It is a purely formal exercise to verify that these manipulations generalise to polynomials, or to series, over a product of monoids.

The semialgebras $S\langle\langle M\rangle\rangle$ and $S\langle\langle N\rangle\rangle$ are canonically isomorphic to two sub-S-semi-algebras of $S\langle\langle M\times N\rangle\rangle$: we identify m with $(m, 1_N)$ and n with $(1_M, n)$. This identification enables us to build the two canonical isomorphisms.

Proposition 4.20. *The three S-semi-algebras*

$$S\langle\langle M \times N\rangle\rangle, \qquad [S\langle\langle M\rangle\rangle]\,\langle\langle N\rangle\rangle \qquad \textit{and} \qquad [S\langle\langle N\rangle\rangle]\,\langle\langle M\rangle\rangle$$

are isomorphic. Under these isomorphisms, the three sub-S-semi-algebras

$$S\langle M \times N\rangle, \qquad S\langle M\rangle\langle N\rangle \qquad \textit{and} \qquad [S\langle N\rangle]\,\langle M\rangle$$

correspond.

Remark 4.21. Modulo this canonical embedding *and if S is commutative*, then every element of $S\langle\langle M\rangle\rangle$ commutes with every element of $S\langle\langle N\rangle\rangle$ in $S\langle\langle M\times N\rangle\rangle$.

Definition 4.22. *Let r be in $S\langle\langle M\rangle\rangle$ and u be in $S\langle\langle N\rangle\rangle$. The* tensor product *of r and u, written $r \otimes u$, is the series of $S\langle\langle M \times N\rangle\rangle$ defined by*

$$\forall (m, n) \in M \times N \quad \big(r \otimes u, (m, n)\big) = (r, m)(u, n).$$

[15] Due to M. Petitot (see Sect. 7).

This definition allows the weighted generalisation of a result and is usually credited to Myhill.

Proposition 4.23. *Suppose that S is commutative. A series r of $S\langle\langle M \times N\rangle\rangle$ is recognisable if and only if there exists a finite family $\{j_i\}_{i\in I}$ of series of $S\mathrm{Rec}\, M$ and a finite family $\{u_i\}_{i\in I}$ of series of $S\mathrm{Rec}\, N$ such that*

$$r = \sum_{i\in I} j_i \otimes u_i.$$

Proof. If j is in $S\mathrm{Rec}\, M$, that is, if j is recognised by the representation (λ, μ, ν), the map $(m, n) \mapsto m\mu$ is also a morphism and the series j' of $S\langle\langle M \times N\rangle\rangle$ defined by $(j', (m, n)) = \lambda \cdot m\mu \cdot \nu = (j, m)$ is recognisable. Likewise, if $u \in S\mathrm{Rec}\, N$, the series u' of $S\langle\langle M \times N\rangle\rangle$ defined by $(u', (m, n)) = (u, n)$ is recognisable. Definition 4.22 shows that

$$j \otimes u = j' \odot u',$$

which is thus recognisable and Proposition 4.4—hence, we need S to be commutative—implies that the condition is sufficient.

Conversely, suppose that r is recognised by (λ, μ, ν), a representation of $M \times N$ of dimension Q. By definition of a representation, for all (m, n) in $M \times N$, it holds $(m, n)\mu = (m, 1_N)\mu(1_M, n)\mu$. The map $\mu' \colon M \to S^{Q\times Q}$ defined by $m\mu' = (m, 1_N)\mu$ is a morphism. For each q in Q, let j_q be the series defined by

$$\forall m \in M \quad (j_q, m) = [\lambda \cdot m\mu']_q,$$

which is a recognisable series of $S\langle\langle M\rangle\rangle$. Likewise, $\mu'' \colon N \to S^{Q\times Q}$ defined by $n\mu'' = (1_M, n)\mu$ is a morphism and u_q defined by

$$\forall n \in N \quad (u_q, n) = [n\mu'' \cdot \nu]_q,$$

is a recognisable series of $S\langle\langle N\rangle\rangle$. Since for all (m, n) of $M \times N$, we have

$$\lambda \cdot (m, n)\mu \cdot \nu = \sum_{q\in Q} [\lambda \cdot m\mu']_q [n\mu'' \cdot \nu]_q,$$

it follows that

$$r = \sum_{q\in Q} j_q \otimes u_q. \quad \square$$

4.3.2 Rational Series in a Product

The Fundamental Theorem of (S-)automata for series in $S\langle\langle M \times N\rangle\rangle$ directly yields (weighted and generalised version of a theorem by Elgot and Mezei [16]) the following.

Proposition 4.24. *Let G and H be generating sets of M and N, respectively. A series of $S\langle\langle M \times N\rangle\rangle$ is rational if and only if it is the behaviour of a proper finite S-automaton whose coefficients are S-linear combinations of elements of $(G \times 1_N) \cup (1_M \times H)$.*

Proposition 4.25. *The canonical isomorphism from $S\langle\langle M \times N\rangle\rangle$ to $[S\langle\langle N\rangle\rangle]\,\langle\langle M\rangle\rangle$ sends $S\mathrm{Rat}(M \times N)$ to $[S\mathrm{Rat}\,N]\mathrm{Rat}\,M$.*

Proof. From the inclusion

$$S\langle N\rangle \subseteq S\langle M\rangle N \subseteq S\mathrm{Rat}(M \times N),$$

we deduce successively, by liberal use of the canonical embeddings,

$$\begin{aligned} S\mathrm{Rat}\,N &\subseteq S\mathrm{Rat}(M \times N),\\ [S\mathrm{Rat}\,N]\langle M\rangle &\subseteq S\mathrm{Rat}(M \times N),\\ [S\mathrm{Rat}\,N]\mathrm{Rat}\,M &\subseteq S\mathrm{Rat}(M \times N). \end{aligned}$$

Conversely, let r be in $S\mathrm{Rat}(M \times N)$. There exists a proper S-automaton $\langle I, X, T\rangle$ such that $r = I \cdot X^* \cdot T$ and such that the coefficients of X are finite S-linear combinations of elements of $(M \times 1) \cup (1 \times N)$. We write $X = Y + Z$, in such a way that the coefficients of Y are linear combinations of elements of $M \times 1$ and those of Z are linear combinations of elements of $1 \times N$ (with coefficients in S). The series r is the result of the automaton $\langle I, Z^* \cdot Y, Z^* \cdot T\rangle$ whose coefficients are linear combinations of elements of $M \times 1$, with coefficients in $1 \times S\mathrm{Rat}\,N$. □

The specialisation of this proposition when M is a free monoid gives the weighted version of what is often known as the 'Kleene–Schützenberger theorem for rational relations' (cf. Corollary 4.29). We shall state it after the definition of *weighted relations*.

4.3.3 Weighted Relations

We first need a few more definitions and notation. We write S_{c} for the *centre* of S, that is, the set of elements of S which commute with every element of S—S_{c} is a sub-semiring of S. In any case, 1_S belongs to S_{c}, which is thus never empty.

The *scalar product*[16] of two series r and u in $S\langle\langle M\rangle\rangle$, written (r, u) is defined by

$$(r, u) = \sum_{m \in M} (r, m)(u, m),$$

which may or may not be defined since the family $\{(r, m)(u, m) \mid m \in M\}$ is not necessarily summable. It is defined if r or u is a polynomial. The identification of m with its *characteristic series* $\underline{m}$ makes this notation consistent

[16] Different from what is called the *scalar product* in Chap. 1.

with the notation (r, m) for the coefficient of m in r. Even *if S is not commutative*, but if r *or* u belong to $S_{\mathsf{c}}\langle\langle M\rangle\rangle$, we have $(r, u) = (u, r)$. In this case, the scalar product is even compatible with left and right multiplication by arbitrary elements of S:

$$k(r, u) = (kr, u),$$
$$(r, u)k = (u, r)k = (u, rk) = (rk, u).$$

Definition 4.26. *An* S-relation from M to N, *written* $\theta\colon S\langle\langle M\rangle\rangle \to S\langle\langle N\rangle\rangle$, *or more often* $\theta\colon M \to N$, *is any series* θ *of* $[S_{\mathsf{c}}\langle\langle N\rangle\rangle]\,\langle\langle M\rangle\rangle$.

The image *of every* m *in* M *under* θ *is the series* (θ, m) *of* $S\langle\langle N\rangle\rangle$, *written more simply* $m\theta$.

The image of every r *in* $S\langle\langle M\rangle\rangle$ *under* θ, *denoted* $r\theta$, *is then obtained 'by linearity'. It is defined if and only if the family* $\{(r, m)(\theta, m) \mid m \in M\}$ *is a* summable family *of series of* $S\langle\langle N\rangle\rangle$ *and is its sum.*

The *graph* $\widehat{\theta}$ of an S-relation θ is the series of $S_{\mathsf{c}}\langle\langle M \times N\rangle\rangle$ which corresponds to θ under the canonical isomorphism. The *inverse* of θ, namely θ^{-1}, is the S-relation from N to M, and hence a series of $[S_{\mathsf{c}}\langle\langle M\rangle\rangle]\,\langle\langle N\rangle\rangle$, which has the same graph $\widehat{\theta}$ as θ. It then holds

$$\forall (m, n) \in M \times N \quad (m\theta, n) = \big(\widehat{\theta}, (m, n)\big) = \big(m, n\theta^{-1}\big). \tag{21}$$

Remark 4.27. Instead of assuming that the semiring of coefficients is commutative, we have 'only' imposed the condition that the coefficients of the relation, $\widehat{\theta}$, belong to the *centre* of this semiring. This could seem a rather weak generalisation; in fact, it allows us first and foremost to consider, as S-relations from M to N, the *characteristic relations* of relations from M to N, even if S is not commutative.

Example 4.28. For every series u in $S_{\mathsf{c}}\langle\langle M\rangle\rangle$, and in particular for every characteristic series u, the Hadamard product with u (or S-*intersection with* u) is an S-relation from M to itself, written $\iota_u\colon r\iota_u = r \odot u$ and $r\iota_u$ is defined for all r in $S\langle\langle M\rangle\rangle$.

It is then natural to say that an S-relation from M to N is *rational* if its graph is a S_{c}-rational series of $S\langle\langle M \times N\rangle\rangle$. And the announced specialisation of Proposition 4.25 then reads as the following corollary.

Corollary 4.29 (Kleene–Schützenberger). *An* S*-relation* θ *from* Σ^* *to* N *is rational if and only if there exists an* $(S_{\mathsf{c}}\mathrm{Rat}\, N)$*-representation of* Σ^*, *namely* (λ, μ, ν), *such that for all* f *in* Σ^*, $f\theta = \lambda \cdot f\mu \cdot \nu$, *that is,*

$$S_{\mathsf{c}}\mathrm{Rat}(\Sigma^* \times N) \cong [S_{\mathsf{c}}\mathrm{Rat}\, N]\mathrm{Rec}\, \Sigma^*.$$

Example 4.30. The rational $\mathbb{B}$-relation from $\Sigma^* = \{a, b\}^*$ into itself realised by the transducer of Fig. 11 is also realised by the $[\mathbb{B}\mathrm{Rat}\, \Sigma^*]$-representation of Σ^* of dimension 1 $(1, \mu, 1)$ with $a\mu = aa^*$ and $b\mu = bb^*$.

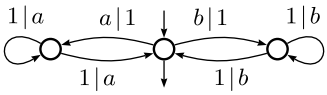


Fig. 11. A transducer to be transformed into a representation

It follows from (21) that the image $r\theta$ of a series r in $S\langle\langle M\rangle\rangle$ by an S-relation θ from M to N is defined if and only if $(r, n\theta^{-1})$ is defined for every n in N, and we have

$$(r\theta, n) = \left(r, n\theta^{-1}\right).$$

Hence, we have the following definition.

Definition 4.31. *We say that an S-relation $\theta\colon M \to N$ is* of finite co-image *if $n\theta^{-1}$ is a* polynomial *for all n.*

The image of any series by a relation of finite co-image is always defined, and this is the case that we shall only consider here. *Regulated relations* which were defined by Jacob starting from their representations as in Corollary 4.29 are relations of finite co-image; they were popularised by a number of authors inspired by Jacob's work (cf. Chap. 7, Sect. 4).

Proposition 4.32. *Let M and N be two graded monoids. An S-relation $\theta\colon M \to N$ with finite co-image is continuous.*

4.3.4 Morphic Image of Recognisable and Rational Series

An S-relation $\theta\colon M \to N$ is *multiplicative* if its restriction to M is a morphism to $S\langle\langle N\rangle\rangle$, viewed as a multiplicative monoid. The definition of S-relations implies in fact that θ is a morphism from M to $S_{\mathsf{c}}\langle\langle N\rangle\rangle$. In particular, the characteristic relation $\underline{\theta}$ of a morphism θ from M to N is a multiplicative S-relation.

We begin with a weighted generalisation of a theorem on recognisable sets.

Proposition 4.33. *Let $\theta\colon M \to N$ be a morphism of monoids and u an S-recognisable series on N. Then $u\underline{\theta}^{-1}$ is an S-recognisable series on M.*

Proof. By assumption, there exists (λ, μ, ν), an S-representation of N, such that for all n in N, $(u, n) = (\lambda \cdot n\mu \cdot \nu)$. Whence, for all m in M,

$$\left(u\underline{\theta}^{-1}, m\right) = (u, m\theta) = \lambda \cdot (m\theta)\mu \cdot \nu.$$

Thus, the S-representation of M $(\lambda, \theta\mu, \nu)$ recognises the series $u\underline{\theta}^{-1}$. □

The hypothesis that the coefficients of an S-relation are taken in S_{c} allows us to establish the following.

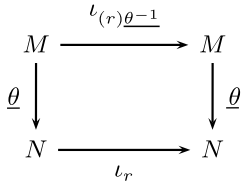


Fig. 12. Lifting of S-intersection with r

Proposition 4.34. *If $\theta\colon M \to N$ is a multiplicative S-relation, then θ is a morphism of S-semialgebras, from $S\langle M\rangle$ to $S\langle\langle N\rangle\rangle$.*

Let M and N be two graded monoids. Let $\theta\colon M \to N$ be a multiplicative S-relation; if for all m in M, $m\theta$ is a *proper series* of $S\langle\langle N\rangle\rangle$, then θ is of finite co-image, hence is defined on all of $S\langle\langle M\rangle\rangle$ and is continuous. In particular, a monoid morphism $\theta\colon M \to N$ is *continuous* if $m\theta \neq 1_N$ for all m in M and then the S-relation $\underline{\theta}$ is a continuous morphism of S-semi-algebras from $S\langle\langle M\rangle\rangle$ to $S\langle\langle N\rangle\rangle$. It follows that if r is in $S\langle\langle M\rangle\rangle$, r^* is defined if and only if $(r\underline{\theta})^*$ is defined and we have $(r^*)\underline{\theta} = (r\underline{\theta})^*$. And the following theorem then holds.

Theorem 4.35. *Let M and N be graded monoids and $\theta\colon M \to N$ a* continuous *morphism of monoids.*

(i) If $r \in S\mathrm{Rat}\, M$, then $r\underline{\theta} \in S\mathrm{Rat}\, N$.
(ii) If θ is surjective and $u \in S\mathrm{Rat}\, N$, then there exists $r \in S\mathrm{Rat}\, M$ such that $r\underline{\theta} = u$.

Example 4.36. Let $\alpha\colon \Sigma^* \to M$ be a surjective morphism; a set R of Σ^* is a *cross-section* of Σ^* for α if α is injective over R and $R\alpha = M$, that is, if $\underline{M} = (\underline{R})\underline{\alpha}$. A monoid M is *rationally enumerable* if such an R exists that is a rational subset of Σ^*.

It easily comes that M is rationally enumerable if and only if it is an unambiguous rational subset of itself: $M \in \mathrm{URat}\, M$, that is, $\underline{M} \in \mathbb{N}\mathrm{Rat}\, M$ and then $\underline{M} \in S\mathrm{Rat}\, M$ for any S.

We prove a last lemma before the result we are aiming at.

Lemma 4.37. *Let $\theta\colon M \to N$ be a function and r a S-series on N. We have (cf. diagram in Fig. 12)*

$$\underline{\theta}\iota_r = \iota_{r\underline{\theta}^{-1}}\underline{\theta}.$$

Proof. For every m in M, we have

$$\begin{aligned}(m\iota_{r\underline{\theta}^{-1}})\underline{\theta} &= \big(r\underline{\theta}^{-1} \odot m\big)\underline{\theta} = \big(\big(r\underline{\theta}^{-1}, m\big)m\big)\underline{\theta} \\ &= (r, m\underline{\theta})m\underline{\theta} = r \odot m\underline{\theta} = (m\underline{\theta})\iota_r. \quad \square\end{aligned}$$

Theorem 4.38. *Let U and V be two commutable sub-semirings of S, u in $V\mathrm{Rat}\, N$ and r in $U\mathrm{Rec}\, N$. Then the Hadamard product of u and r is an S-rational series on N.*

Proof. As N is finitely generated there exists a finite alphabet Σ and a surjective continuous morphism $\theta\colon \Sigma^* \to N$. By Theorem 4.35(ii), there exists a series u in $V\mathrm{Rat}\, \Sigma^*$ such that

$$u\underline{\theta} = u.$$

The coefficients of u commute with those of r, and hence with those of $r\underline{\theta^{-1}}$. Lemma 4.37 allows us to rewrite the equality $r \odot u = r \odot u\underline{\theta}$ as

$$r \odot u = \left[r\underline{\theta^{-1}} \odot u\right]\underline{\theta}. \tag{22}$$

Proposition 4.33 ensures that $r\underline{\theta^{-1}}$ is U-recognisable (on Σ^*), Theorem 4.13 that $r\underline{\theta^{-1}} \odot u$ is S-recognisable, hence S-rational, and finally (22) and Theorem 4.35(i) that $r \odot u$ is S-rational on N. □

Corollary 4.39. *If M is rationally enumerable, then $S\mathrm{Rec}\, M \subseteq S\mathrm{Rat}\, M$.*

Proof. By hypothesis (cf. Example 4.36), $\underline{M} \in S\mathrm{Rat}\, M$ for any S. We have $r \odot \underline{M} = r$ for all r in $S\langle\langle M\rangle\rangle$ and we apply Theorem 4.38. □

Corollary 4.39 is the weighted generalisation of a theorem by McKnight [35], Theorem 4.38, the one of a classical result on subsets of a monoid. As for subsets also, the morphic image of a recognisable series is not necessarily recognisable, the inverse morphic image of a rational series is not necessarily rational.

We stop here with the theory of weighted relations, which could, of course, be further developed. In particular, the composition and evaluation theorems hold for weighted rational relations (cf. [24, 45, 43]). But our aim here was just to set the framework in which we could establish Theorems 4.35, 4.38, and Corollary 4.39, and in the next section, Corollary 5.32.

5 Series over a Free Monoid

So far, we have developed the theory of rational series under the assumption that M is *graded* (so that we knew how to define star). In our presentation, the Kleene–Schützenberger theorem and recognisable series appeared as a last touch added to the fundamental theorem of automata in the case of free monoids. We now require M to be a *free monoid* and change our point of view: rational and recognisable series coincide and somehow recognisable series and their representations become the main subject.

The whole thing takes an algebraic turn. We first give another characterisation of recognisable series, and then under the hypothesis that the semiring of weights is a field, we develop the theory of reduction (that is, minimisation)

of representations. In a third subsection, we review a number of applications of this reduction theory—and first of all, the decidability of equivalence—which in many instances, do not apply only to the case of weights in a field, but also in *any sub-semiring* of a field.

5.1 The Representability Theorem

Representations define recognisable series; we first show how, by means of the *quotient operation*, we can recover a representation from a series when it is recognisable. This is an abstract view since a series is an infinite object; we then give an effective implementation of this result, starting from a rational expression that denotes a rational series; this is another proof of one direction of the Kleene–Schützenberger theorem.

5.1.1 Characterisation of Recognisable Series

The (left) quotient of a series is the generalisation to series of the (left) quotient of a subset of a monoid (a free monoid in this case).

The free monoid Σ^* *acts by quotient* on $S\langle\langle\Sigma^*\rangle\rangle$: for all f in Σ^* and all series r in $S\langle\langle\Sigma^*\rangle\rangle$, the series $f^{-1}r$ is defined by

$$f^{-1}r = \sum_{g\in\Sigma^*} (r, fg)g, \qquad \text{that is,} \qquad \forall g \in \Sigma^* \quad \big(f^{-1}r, g\big) = (r, fg),$$

and in particular

$$\forall f \in \Sigma^* \quad \big(f^{-1}r, \varepsilon\big) = (r, f). \tag{23}$$

As the definition says, the quotient is an *action*, that is,

$$\forall f, g \in \Sigma^* \quad (fg)^{-1}r = g^{-1}\big[f^{-1}r\big],$$

and for every given f, the operation $r \mapsto f^{-1}r$ is an *endomorphism* of the S-semi-module $S\langle\langle\Sigma^*\rangle\rangle$: it is *additive*:

$$f^{-1}(r + u) = f^{-1}r + f^{-1}u,$$

and *commutes with the exterior multiplications* of S on $S\langle\langle\Sigma^*\rangle\rangle$:

$$f^{-1}(kr) = k\big(f^{-1}r\big) \quad \text{and} \quad f^{-1}(rk) = \big(f^{-1}r\big)k.$$

Moreover, it is *continuous*. These three properties ensure that the operation of quotient by f is entirely defined on $S\langle\langle\Sigma^*\rangle\rangle$ by its values on Σ^*.

Example 5.1. Let $r_2 = (\underline{a^*})^2 = \sum_{k\in\mathbb{N}}(k+1)a^k$ in $\mathbb{N}\mathrm{Rat}\, a^*$. For every integer n, we have

$$\big(a^n\big)^{-1}r_2 = \sum_{k\in\mathbb{N}}(n + k + 1)a^k = r_2 + n\underline{a^*}.$$

All quotients of r_2 are distinct.

Example 5.1 shows that, in general, and unlike the case for (recognisable) languages, the family of quotients of a rational, and thus recognisable series is not necessarily finite. On the other hand, and despite its simplicity, it exhibits the property that we seek: of course, there are infinitely many quotients, but they can all be expressed as the linear combination of a *finite number* of suitably chosen series.

Definition 5.2. *A subset U of $S\langle\langle\Sigma^*\rangle\rangle$ is called* stable *if it is closed under quotient; that is, for all r in U and all f in Σ^*, $f^{-1}r$ is still in U.*

Theorem 5.3. *A series on Σ^* with coefficients in S is S-recognisable if and only if it is contained in a finitely generated stable subsemimodule of $S\langle\langle\Sigma^*\rangle\rangle$.*

To allow later references to parts of the proof of this result, it is split into more precise properties and definitions. Let us begin with a notation: Lemma 4.7 shows how close automata and representations are. We shall thus denote the latter in the same way as the former by uppercase gothic letters.

Definition 5.4. *With every S-representation $\mathfrak{A} = (\lambda, \mu, \nu)$ of dimension Q we associate a morphism of S-semimodules $\Phi_{\mathfrak{A}} : S^Q \to S\langle\langle\Sigma^*\rangle\rangle$ by*

$$\forall x \in S^Q \quad (x)\Phi_{\mathfrak{A}} = \big|(x, \mu, \nu)\big| = \sum_{f \in \Sigma^*} (x \cdot f\mu \cdot \nu) f.$$

Proposition 5.5. *If r is a series realised by $\mathfrak{A} = (\lambda, \mu, \nu)$, then $\mathsf{Im}\,\Phi_{\mathfrak{A}}$ is a stable (finitely generated) subsemi-module of $S\langle\langle\Sigma^*\rangle\rangle$ that contains r.*

Proof. The subsemimodule $\mathsf{Im}\,\Phi_{\mathfrak{A}}$ is finitely generated since S^Q is, and it is stable since for all f in Σ^* and all x in S^Q we have

$$f^{-1}\left[(x)\Phi_{\mathfrak{A}}\right] = (x \cdot f\mu)\Phi_{\mathfrak{A}},$$

and contains $r = (\lambda)\Phi_{\mathfrak{A}}$. □

Proposition 5.6. *Let U be a stable subsemimodule of $S\langle\langle\Sigma^*\rangle\rangle$ generated by $G = \{g^{(1)}, g^{(2)}, \ldots, g^{(n)}\}$. Then every series in U is an S-recognisable series of $S\langle\langle\Sigma^*\rangle\rangle$, realised by a representation of dimension n.*

Proof. The set G canonically defines a linear map from S^n onto U:

$$x = (x_1, x_2, \ldots, x_n) \longmapsto x \cdot G = x_1 g^{(1)} + x_2 g^{(2)} + \cdots + x_n g^{(n)}.$$

A series u belongs to U means that there exists at least one x in S^n such that $u = x \cdot G$.

If U is stable, for every a in Σ, and every i, $a^{-1}g^{(i)}$ belongs to U and there exists a vector $m^{(i)}$ in S (at least one) such that $a^{-1}g^{(i)} = m^{(i)} \cdot G$. Let $a\mu$ be the $n \times n$-matrix whose ith row is $m^{(i)}$. As the quotient by a is a linear map, for any u in U, $u = x \cdot G$ it holds $a^{-1}u = (x \cdot a\mu) \cdot G$. These matrices $a\mu$, for a

in Σ, define a representation of Σ^* and as the quotient is an action of Σ^*, for every f in Σ^*, it holds $f^{-1}u = (x \cdot f\mu) \cdot G$.

From (23), follows then $(u, f) = (f^{-1}u, \varepsilon) = ((x \cdot f\mu) \cdot G, \varepsilon)$ and u is realised by the representation $(x, \mu, (G, \varepsilon))$ where (G, ε) denotes the (column) vector $((g^{(1)}, \varepsilon), (g^{(2)}, \varepsilon), \ldots, (g^{(n)}, \varepsilon))$. □

Propositions 5.5 and 5.6 together prove Theorem 5.3.

5.1.2 Derivation of Rational S-Expressions

The *derivation* of rational S-expressions is the lifting to the level of expressions of the quotient of series and will enable us to effectively implement Theorem 5.3: the derived terms of an expression denote a set of generators of a stable subsemimodule that contains the series denoted by the expression. It will give us the weighted generalisation of Antimirov's construction on rational expressions [1]; this is another example where taking multiplicities into account yield better understanding of constructions and results on languages.

S-Derivatives

For the rest of this subsection, addition in S is written $\oplus$ to distinguish it from the $+$ operator in expressions. The addition induced on $S\langle\langle \Sigma^* \rangle\rangle$ is also written $\oplus$. The set of left linear combinations of S-expressions with coefficients in S, or polynomials of $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$, is a left S-semi-module on S:

$$k\mathsf{E} \oplus k'\mathsf{E}' \equiv k'\mathsf{E}' \oplus k\mathsf{E} \quad \text{and} \quad k\mathsf{E} \oplus k'\mathsf{E} \equiv [k \oplus k']\,\mathsf{E}. \tag{$\boldsymbol{B_K}$}$$

In the following, $[k\,\mathsf{E}]$ or $k\,\mathsf{E}$ is a monomial in $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$ whereas $(k\,\mathsf{E})$ is an expression in $S\,\mathsf{RatE}\,\Sigma^*$.

As it is the case in general for semi-modules, there is no multiplication defined on $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$. However, an external right multiplication of an element of $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$ by an expression and by a scalar is needed. This operation is first defined on monomials and then extended to polynomials by linearity:

$$([k\,\mathsf{E}] \cdot \mathsf{F}) \equiv k\,(\mathsf{E} \cdot \mathsf{F}), \qquad ([k\,\mathsf{E}]\,k') \equiv k\,(\mathsf{E}\,k'),$$
$$([\mathsf{E} \oplus \mathsf{E}'] \cdot \mathsf{F}) \equiv (\mathsf{E} \cdot \mathsf{F}) \oplus (\mathsf{E}' \cdot \mathsf{F}), \qquad ([\mathsf{E} \oplus \mathsf{E}']\,k) \equiv (\mathsf{E}\,k) \oplus (\mathsf{E}'\,k).$$

This multiplication on $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$ is *not associative*—since the product operator in expression is not—but is consistent with interpretation: the series denoted by the left-hand sides and right-hand sides are equal.

Definition 5.7. *Let* E *be in* $S\,\mathsf{RatE}\,\Sigma^*$ *and let* a *be in* Σ*. The* S-derivative *of* E *with respect to* a*, denoted by* $\frac{\partial}{\partial a}\mathsf{E}$*, is a polynomial of rational expressions with coefficients in* S*, defined inductively by the following formulas.*

$$\frac{\partial}{\partial a}\,0 = \frac{\partial}{\partial a}\,1 = 0, \qquad \frac{\partial}{\partial a}\,b = \begin{cases} 1 & \text{if } b = a, \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial}{\partial a}(k\,\mathsf{E}) = k\,\frac{\partial}{\partial a}\,\mathsf{E}, \qquad \frac{\partial}{\partial a}(\mathsf{E}\,k) = \left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\,k\right),$$

$$\frac{\partial}{\partial a}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial a}\,\mathsf{E} \oplus \frac{\partial}{\partial a}\,\mathsf{F}, \tag{24}$$

$$\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot\mathsf{F}\right) \oplus \mathsf{c}(\mathsf{E})\frac{\partial}{\partial a}\,\mathsf{F}, \tag{25}$$

$$\frac{\partial}{\partial a}(\mathsf{E}^*) = \mathsf{c}(\mathsf{E})^*\left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot(\mathsf{E}^*)\right). \tag{26}$$

The derivative of a polynomial of expressions is defined by linearity:

$$\frac{\partial}{\partial a}\left(\bigoplus_{i\in I} k_i\,\mathsf{E}_i\right) = \bigoplus_{i\in I} k_i\,\frac{\partial}{\partial a}\,\mathsf{E}_i. \tag{27}$$

Implicitly, the (polynomials of) expressions are reduced with trivial identities, for instance,

$$\frac{\partial}{\partial a}\,\mathsf{E} = 1 \quad\Longrightarrow\quad \frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \mathsf{F} \oplus \mathsf{c}(E)\frac{\partial}{\partial a}\,\mathsf{F}.$$

Notice that (26) is defined only if (E^*) is a valid expression. The S-derivative of an expression with respect to a *word* f is defined by induction on the length of f:

$$\forall f \in \Sigma^+,\ \forall a \in \Sigma \quad \frac{\partial}{\partial fa}\,\mathsf{E} = \frac{\partial}{\partial a}\left(\frac{\partial}{\partial f}\,\mathsf{E}\right). \tag{28}$$

The definition of S-derivatives of S-expressions is consistent with that of quotient of series, as expressed by the following.

Proposition 5.8. $\forall \mathsf{E} \in S\,\mathsf{RatE}\,\Sigma^*,\ \forall f \in \Sigma^+$

$$\left|\frac{\partial}{\partial f}(\mathsf{E})\right| = f^{-1}|\mathsf{E}|.$$

The Derived Term Automaton

Definition 5.9. *The set* $\mathrm{TD}(\mathsf{E})$ *of* true derived terms *of an expression* E *in* $S\,\mathsf{RatE}\,\Sigma^*$ *is inductively defined by the following rules:*

$$\mathrm{TD}(0) = \mathrm{TD}(1) = \emptyset, \qquad \forall a \in \Sigma \quad \mathrm{TD}(a) = \{1\},$$

$$\forall k \in S, \quad \mathrm{TD}(k\,\mathsf{E}) = \mathrm{TD}(\mathsf{E}), \quad \mathrm{TD}(\mathsf{E}\,k) = \bigcup_{\mathsf{K}\in\mathrm{TD}(\mathsf{E})} (\mathsf{K}\,k),$$

$$\mathrm{TD}(\mathsf{E}+\mathsf{F}) = \mathrm{TD}(\mathsf{E}) \cup \mathrm{TD}(\mathsf{F}),$$

$$\mathrm{TD}(\mathsf{E}\cdot\mathsf{F}) = \left[\bigcup_{\mathsf{K}\in\mathrm{TD}(\mathsf{E})} (\mathsf{K}\cdot\mathsf{F})\right] \cup \mathrm{TD}(\mathsf{F}),$$

$$\mathrm{TD}(\mathsf{E}^*) = \bigcup_{\mathsf{K}\in\mathrm{TD}(\mathsf{E})} \big(\mathsf{K}\cdot(\mathsf{E}^*)\big).$$

It follows from the definition that $\mathrm{TD}(\mathsf{E})$ is a *finite* set of monomials of $S\langle S\,\mathsf{RatE}\,\Sigma^*\rangle$, whose cardinal is smaller than or equal to $\ell(\mathsf{E})$. The reason for the two distinct definitions (Definitions 5.7 and 5.9), which may look redundant will be explained below.

The expression E itself does not belong necessarily to $\mathrm{TD}(\mathsf{E})$ and we define the set of *derived terms* of E to be: $\mathrm{D}(\mathsf{E}) = \mathrm{TD}(\mathsf{E}) \cup \{\mathsf{E}\}$. A mechanical induction on the depth of the expressions establishes then the following.

Theorem 5.10. *Let* $\mathrm{D}(\mathsf{E}) = \{\mathsf{K}_1, \dots, \mathsf{K}_n\}$ *be the set of derived terms of an expression* E *in* $S\,\mathsf{RatE}\,\Sigma^*$. *For every letter* a *in* Σ, *there exists an* $n \times n$*-matrix* $a\mu$ *with entries in* S *such that*

$$\forall i \in [n] \quad \frac{\partial}{\partial a}\mathsf{K}_i = \bigoplus_{j\in[n]} a\mu_{i,j}\mathsf{K}_j.$$

From (28), it then follows, by induction on the length of words.

Corollary 5.11. *For every word* f *in* Σ^*, *the* S*-derivative of any expression* E *in* $S\,\mathsf{RatE}\,\Sigma^*$ *with respect to* f *is a linear combination of derived terms of* E.

The statement of Theorem 5.10 is in itself the definition of an S-representation $\mathfrak{A}_{\mathsf{E}} = (\lambda, \mu, \nu)$ of dimension $\mathrm{D}(\mathsf{E})$ if we add

$$\lambda_i = \begin{cases} 1_S & \text{if } \mathsf{K}_i = \mathsf{E}, \\ 0_S & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \nu_j = \mathsf{c}(\mathsf{K}_j).$$

We also write $\mathfrak{A}_{\mathsf{E}}$ for the S-automaton $\langle \lambda, X, \nu\rangle$ where $X = \bigoplus_{a\in\Sigma} a\mu a$ and call it *the derived term automaton* of E.

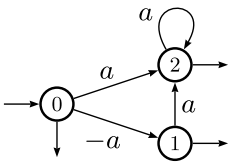
Proposition 5.12. *Let* E *be in* $S\,\mathsf{RatE}\,\Sigma^*$. *Then* $|\mathfrak{A}_{\mathsf{E}}| = |\mathsf{E}|$.

Derivation is thus another means to build an automaton from an expression, different from the one we have seen in the course of the proof of Theorem 3.20 which yielded the *standard automaton* of the expression. The two constructions are related by the following, which is the weighted generalisation of a theorem by Champarnaud and Ziadi [10].

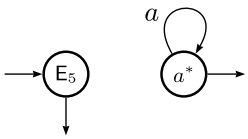
Theorem 5.13 ([33]). *Let* E *be in* $S\,\mathsf{RatE}\,\Sigma^*$. *Then* $\mathfrak{S}_{\mathsf{E}}$ *is an* S*-covering of* $\mathfrak{A}_{\mathsf{E}}$.

Remark 5.14. Definitions 5.7 and 5.9 are both based on an induction on the depth of the expression and then reunited by Theorem 5.10 and Corollary 5.11. It seems that it could be possible, and more natural, to define the derived terms of E as the monomials that appear in the S-derivatives of E.

The problem is that this is not always true if S is not a positive semiring: some derived terms may never appear in an S-derivative—as it can be observed for instance with the $\mathbb{Z}$-expression $\mathsf{E}_5 = (1 - a)a^*$ (cf. Fig. 13). And with such a definition of derived terms, more utilitarian than structural, Theorem 5.13 would not hold anymore.



a The standard automaton: $\mathfrak{S}_{\mathsf{E}_5}$



b The derived term automaton: $\mathfrak{A}_{\mathsf{E}_5}$

Fig. 13. Two $\mathbb{Z}$-automata for E_5

5.2 Reduced Representations

We now suppose that S is a *field*, not necessarily commutative, hence a *skew field*, or *division ring*. The preceding considerations about quotients of series will take on, we might say, a new dimension since the ring of series $S\langle\langle\Sigma^*\rangle\rangle$ is not only an S-algebra, but a left and right S-*vector space*, and the notion of *dimension* of subspaces will give us a new invariant.

5.2.1 Rank of a Series

Definition 5.15. *Let S be a division ring. The* rank of *a series r of $S\langle\langle\Sigma^*\rangle\rangle$ is the dimension of the subspace of $S\langle\langle\Sigma^*\rangle\rangle$ generated by the (left) quotients of r.*

In this setting, and with no further ado, Theorem 5.3 becomes the following theorem.

Theorem 5.16. *A series r over Σ^* with coefficients in a division ring is recognisable if and only its rank is finite.*

From Definition 5.4 and Proposition 5.5, it follows that if r is a series realised by an S-representation $\mathfrak{A} = (\lambda, \mu, \nu)$ of dimension n, the rank of r is smaller than or equal to $\dim(\mathrm{Im}\,\Phi_{\mathfrak{A}})$ which is smaller than or equal to n, that is, the rank of a recognisable series r is smaller than, or equal to, the dimension of any representation that realises it.

Definition 5.17. *A representation of a recognisable series r is* reduced *if its dimension is minimal, equal to the rank of r.*

From Proposition 5.6, it follows that with every base of the subspace generated by the quotients of r is associated a reduced representation. The reduced representations will be characterised by means of the following definition. With every S-representation $\mathfrak{A} = (\lambda, \mu, \nu)$ of dimension Q, we associate the morphism of S-semi-modules $\Psi_{\mathfrak{A}} : S\langle\Sigma^*\rangle \to S^Q$ defined by

$$\forall f \in \Sigma^* \quad (f)\Psi_{\mathfrak{A}} = \lambda \cdot f\mu.$$

Theorem 5.18. *An S-representation $\mathfrak{A} = (\lambda, \mu, \nu)$ is reduced if and only if $\Psi_{\mathfrak{A}}$ is surjective and $\Phi_{\mathfrak{A}}$ injective.*

Proof. Let r be the series realised by $\mathfrak{A}$. The morphism

$$\Psi_{\mathfrak{A}} \circ \Phi_{\mathfrak{A}} : S\langle\langle \Sigma^* \rangle\rangle \to S\langle\langle \Sigma^* \rangle\rangle \quad \text{is such that} \quad (f)[\Psi_{\mathfrak{A}} \circ \Phi_{\mathfrak{A}}] = f^{-1}r$$

for every f in Σ^* and $\mathsf{Im}\,\Psi_{\mathfrak{A}} \circ \Phi_{\mathfrak{A}}$ is the subspace generated by the quotients of r. For the dimension of $\mathsf{Im}\,\Psi_{\mathfrak{A}} \circ \Phi_{\mathfrak{A}}$ be equal to n, the dimension of $\mathfrak{A}$, it is necessary and sufficient that the dimension of both $\mathsf{Im}\,\Psi_{\mathfrak{A}}$ and $\mathsf{Im}\,\Phi_{\mathfrak{A}}$ be equal to n. The second equality holds if and only if the dimension of $\mathsf{Ker}\,\Phi_{\mathfrak{A}}$ is zero. □

Remark 5.19. The significance of the map $\Psi_{\mathfrak{A}}$ goes beyond the case of weights taken in a field. Without linearisation, $(\Sigma^*)\Psi_{\mathfrak{A}}$ is the *reachability set* of $\mathfrak{A}$. If $S = \mathbb{B}$, $(\Sigma^*)\Psi_{\mathfrak{A}}$ is a set of subsets of states of $\mathfrak{A}$, namely the set of states of the determinisation of $\mathfrak{A}$ (by the so-called *subset construction*).

5.2.2 The Reduction Algorithm

It is not enough to know that reduced representations exist and to characterise them. We want to be able to effectively compute them and establish the following.

Theorem 5.20. *A reduced representation of a recognisable series r is effectively computable from any representation that realises r with a procedure whose complexity is cubic in the dimension of the representation.*

For the rest of this section, let $\mathfrak{A} = (\lambda, \mu, \nu)$ be a S-representation of Σ^* of dimension n (that realises the series $r = |\mathfrak{A}|$).

Word Base

The effective computation from $\mathfrak{A}$ of a reduced representation of r is based on one definition and two propositions that are related but whose scope and aim are nevertheless rather different.

Definition 5.21. *We call* word base *for $\mathfrak{A}$ a prefix-closed subset P of Σ^* such that the set $(P)\Psi_{\mathfrak{A}} = \{\lambda \cdot p\mu \mid p \in P\}$ is a base of $\mathsf{Im}\,\Psi_{\mathfrak{A}}$.*

Proposition 5.22. *Word bases for $\mathfrak{A}$ do exist.*

Proof. If $\lambda = 0$, $\mathsf{Im}\,\Psi_{\mathfrak{A}}$ is the null vector space of dimension 0 and the empty set (which is prefix-closed!) is a word base. Assuming that λ is non-zero, the family of prefix-closed subsets P of Σ^* such that $\{\lambda \cdot p\mu \mid p \in P\}$ is a free subset of S^n is not empty since it contains at least the singleton $\{\varepsilon\}$. Every such subset contains at most $k = \dim(\mathsf{Im}\,\Psi_{\mathfrak{A}})$ elements and there exist thus maximal elements (for the inclusion order) in that family.

It remains to show that such a maximal element P is a word base, that is, $(P)\Psi_{\mathfrak{A}}$ generates $\mathsf{Im}\,\Psi_{\mathfrak{A}}$. By way of contradiction, let f in Σ^* such that $\lambda \cdot f\mu$ does not belong to $\langle\!(P)\Psi_{\mathfrak{A}}\rangle\!\rangle$; the word f factorises in $f = pg$, with p in P, and we choose f in such a way that g is of minimal length. The word g is not empty: $g = ah$, with a in Σ, and $\lambda \cdot f\mu = \lambda \cdot (pa)\mu \cdot h\mu$. As P is maximal, $\lambda \cdot (pa)\mu$ belongs to $\langle\!(P)\Psi_{\mathfrak{A}}\rangle\!\rangle$, that is, $\lambda \cdot (pa)\mu = \sum_{p_i \in P} x_i(\lambda \cdot p_i\mu)$. It then follows

$$\lambda \cdot f\mu = \Big(\sum_{p_i \in P} x_i(\lambda \cdot p_i\mu) \Big) \cdot h\mu = \sum_{p_i \in P} x_i\big(\lambda \cdot (p_ih)\mu\big).$$

By the minimality of g, every $\lambda \cdot (p_ih)\mu$ belongs to $\langle\!(P)\Psi_{\mathfrak{A}}\rangle\!\rangle$: contradiction. □

In the sequel, we do not consider the trivial case $\lambda = 0$ anymore.

Proposition 5.23. *With every word base P for $\mathfrak{A}$ of cardinal m is associated a representation $\mathfrak{A}' = (\lambda', \mu', \nu')$ of dimension m—effectively computable from P and $\mathfrak{A}$—which is conjugate to $\mathfrak{A}$ and with the property that $\Psi_{\mathfrak{A}'}$ is surjective. Moreover, if $\Phi_{\mathfrak{A}}$ is injective, then so is $\Phi_{\mathfrak{A}'}$.*

Proof. Let $P = \{p_1 = \varepsilon, p_2, \ldots, p_m\}$ be a word base for $\mathfrak{A}$ and X the $m \times n$-matrix (with entries in S) whose i-th row is $\lambda \cdot (p_i)\mu$. Let us denote $\nu' = X \cdot \nu$ and by λ' the (row) m-vector whose entries are all 0 but the first one which is 1—thus $\lambda' \cdot X = \lambda$.

For every a in Σ, let $a\mu'$ be the $m \times m$-matrix (with entries in S) whose ith row is the vector of coordinates of $\lambda \cdot (p_ia)\mu$ in the base $\lambda \cdot (P)\mu$, that is,

$$\lambda \cdot (p_ia)\mu = \sum_{j=1}^{j=m} (a\mu')_{i,j}(\lambda \cdot p_j\mu). \tag{29}$$

Since $\lambda \cdot (p_ia)\mu = (\lambda \cdot p_i\mu) \cdot a\mu$, the set of equations (29) for all i may be rewritten in a more compact way as

$$a\mu' \cdot X = X \cdot a\mu$$

and $\mathfrak{A}'$ is conjugated to $\mathfrak{A}$ by X.

If P is not a word base for $\mathfrak{A}'$, there exist m coefficients α_i such that $\sum_{i=1}^{i=m} \alpha_i(\lambda' \cdot p_i\mu') = 0$, but multiplying this equality on the right by X yields $\sum_{i=1}^{i=m} \alpha_i(\lambda \cdot p_i\mu) = 0$, a contradiction (with the fact that P is a word base for $\mathfrak{A}$).

If $\Phi_{\mathfrak{A}'}$ is not injective, there exists a non-zero vector y in S^m such that $y \cdot f\mu' \cdot \nu' = 0$, and thus $(y \cdot X) \cdot f\mu \cdot \nu = 0$ for every f in Σ^*. If $\Phi_{\mathfrak{A}}$ is injective, then $y \cdot X = 0$, and thus $y = 0$ for the same reason as above, a contradiction. □

Remark 5.24 (Remark 5.19 continued). Let $\mathfrak{D}$ be the determinisation of a classical automaton $\mathfrak{A}$ (that is, an automaton with weight in $\mathbb{B}$) of dimension Q by the subset construction. If we form the (Boolean) matrix X whose rows are the states of $\mathfrak{D}$ (Boolean vectors of dimension Q), then $\mathfrak{D}$ is conjugate to $\mathfrak{A}$ by X.

Demonstration of the Reduction Theorem (Theorem 5.20)

We first observe that Proposition 5.23 has obviously a dual formulation, which we rather state on the transpose of the representation $\mathfrak{A}$, ${}^{\mathrm{t}}\mathfrak{A} = ({}^{\mathrm{t}}\nu, {}^{\mathrm{t}}\mu, {}^{\mathrm{t}}\lambda)$ where $a{}^{\mathrm{t}}\mu = {}^{\mathrm{t}}(a\mu)$ for every a in Σ and it comes $f{}^{\mathrm{t}}\mu = {}^{\mathrm{t}}({}^{\mathrm{t}}f\mu)$ for every f in Σ^*. We then have the following connection between $\mathfrak{A}$ and ${}^{\mathrm{t}}\mathfrak{A}$.

Lemma 5.25. *If $\Psi_{{}^{\mathrm{t}}\mathfrak{A}}$ is surjective, then $\Phi_{\mathfrak{A}}$ is injective.*

Proof. If $x\Phi_{\mathfrak{A}} = 0$ then $x \cdot f\mu \cdot \nu = 0$ for every f in Σ^* and x belongs to the orthogonal of the subspace generated by the vectors $\{f\mu \cdot \nu \mid f \in \Sigma^*\}$ which is of dimension n by hypothesis: thus $x = 0$. □

Starting from a representation $\mathfrak{A}$, we first compute a word base for ${}^{\mathrm{t}}\mathfrak{A}$ which determines a representation ${}^{\mathrm{t}}\mathfrak{A}'$ such that $\Psi_{{}^{\mathrm{t}}\mathfrak{A}'}$ is surjective, and thus by Lemma 5.25, $\Phi_{\mathfrak{A}'}$ is injective. We then compute a word base for $\mathfrak{A}'$ which determines a representation $\mathfrak{A}''$ such that $\Psi_{\mathfrak{A}''}$ is surjective and $\Phi_{\mathfrak{A}''}$ is injective: $\mathfrak{A}''$ is reduced. The proof of Theorem 5.20 will be complete when we have proved that word bases are effectively computable (with the ascribed complexity).

The foregoing proofs all correspond to effective computations, assuming of course that the operations in S (addition, multiplication, taking the inverse) are effective. All the complexities that follow are calculated assuming that each operation in S has a fixed constant cost, independent of its operands. Computations in S^n are based on the *Gaussian elimination* procedure.

Definition 5.26. *A sequence of k vectors $(v^1, v^2, \ldots, v^k)$ of S^n is an* echelon system *if, for all i in $[k]$:*

(i) ${v^i}_i = 1_S$.
(ii) $\forall j < i \ {v^i}_j = 0_S$.

An echelon system is free, and hence $k \leqslant n$.The following proposition is classic, at least for commutative fields, and its proof is not really different for division rings.

Proposition 5.27 (Gaussian elimination). *Let S be a skew field and let us view S^n as a left vector space over S. Let $U = (v^1, v^2, \ldots, v^k)$ be an echelon system and let w be a vector in S^n.*

(i) We can decide whether w is in $\langle U \rangle$, the subspace generated by U, and in this case, compute effectively the coordinates of w in U.

(ii) If w is not in $\langle U \rangle$, we can compute effectively w' such that $U' = U \cup \{w'\}$ is echelon and generates the same subspace as $U \cup \{w\}$.

The complexity of these operations (deciding whether w is in $\langle U \rangle$ and computing the coordinates of either w or w') is $O(kn)$.

From this proposition, we deduce the effective nature of the assertions, constructions, and specifications used in the proofs of this section. More precisely, the corollary follows.

Corollary 5.28. *Let U be a finite set of vectors of S^n and let w be in S^n.*

(i) We can decide whether w belongs to $\langle U \rangle$.
(ii) We can extract effectively from U a basis V of $\langle U \rangle$.
(iii) We can compute effectively the coordinates in V of an (explicitly given) vector of $\langle U \rangle$.

The following proposition and its proof exhibit the computation underlying Proposition 5.23 (remember, we have defined the *border* of a prefix-closed subset at Sect. 2.2.2).

Proposition 5.29. *Word bases for $\mathfrak{A}$ are effectively computable, with complexity $O(dn^3)$, where d is the cardinal of Σ.*

Proof. We set $P_0 = \{\varepsilon\}$ and $C_0 = \emptyset$. The algorithm to compute a word base P can be written in the following manner.

If $E_k = (P_k\Sigma \backslash P_k) \backslash C_k$ is non-empty, choose an arbitrary f in E_k and decide whether $\lambda \cdot f\mu$ belongs to $\langle \lambda \cdot P_k\mu \rangle$.

(i) If not, then $P_{k+1} = P_k \cup \{f\}$ and $C_{k+1} = C_k$.
(ii) If so, then $P_{k+1} = P_k$ and $C_{k+1} = C_k \cup \{f\}$.

Set $k = k + 1$ and start again.

The algorithm terminates when E_k is empty and at that moment $C_k = P_k\Sigma \backslash P_k$ is the border of P_k. The algorithm must terminate since P_k has at most n elements, so $P_k \cup C_k$ has at most $\|\Sigma\|n+1$ elements and this set grows by 1 at each step of the algorithm.

By construction, P_k is prefix-closed, and each element f of C_k is such that $\lambda \cdot f\mu$ belongs to $\langle \lambda \cdot P_k\mu \rangle$: when E_k is empty, P_k is maximal. □

5.3 Applications of the Reduction of Representations

We consider here three applications: the decidability of equivalence of S-automata (for certain S), the generalisation of the recurrence relation on the coefficients of a rational series over non-commuting variables, and a *structural interpretation* of equivalence of S-automata in terms of conjugacy and covering (again for certain S).

5.3.1 Equivalence Decidability

Even if a series has not a unique reduced representation (they are all *similar*), the existence of reduced representations implies the decidability of equivalence for automata with weights in a field.

Theorem 5.30. *The equivalence of recognisable series over Σ^* with coefficients in a (sub-semiring of a) skew field—and thus of rational series—is decidable, with a procedure which is cubic in the dimension of the representation of the series.*

Proof. Let S be a sub-semiring of a skew field $\mathbb{F}$. Two series r_1 and r_2 of $S\mathrm{Rec}\,\Sigma^*$ are also in $\mathbb{F}\mathrm{Rec}\,\Sigma^*$ and $r_1 = r_2$ holds if, and only if, $(r_1 - r_2)$ is a series of $\mathbb{F}\mathrm{Rec}\,\Sigma^*$ of rank 0, and the rank of $(r_1 - r_2)$ can be computed effectively. □

This result, together with the well-known decidability of equivalence of classical Boolean automata, should not let us think that this is the universal status. For instance, the following holds.

Theorem 5.31 ([28]). *The equivalence of recognisable series over Σ^* with coefficients in the semiring $\mathbb{M} = \langle \mathbb{N}^\infty, \mathsf{min}, + \rangle$ is undecidable.*

Theorem 5.30 has however far reaching and to some extent 'unexpected' consequences, as the following one, discovered by T. Harju and J. Karhumäki.

Corollary 5.32 ([22]). *The equivalence of rational series over $\Sigma_1^* \times \Sigma_2^* \times \cdots \times \Sigma_k^*$ with coefficients in $\mathbb{N}$ is decidable.*

Proof. By Proposition 4.25, a series in $\mathbb{N}\mathrm{Rat}\,\Sigma_1^* \times \Sigma_2^* \times \cdots \times \Sigma_k^*$ is a series in $[\mathbb{N}\mathrm{Rat}\,\Sigma_2^* \times \cdots \times \Sigma_k^*]\mathrm{Rat}\,\Sigma_1^*$. By Corollary 4.29, the latter family is isomorphic to $[\mathbb{N}\mathrm{Rat}\,\Sigma_2^* \times \cdots \times \Sigma_k^*]\mathrm{Rec}\,\Sigma_1^*$ and the decidability of equivalence follows from Theorem 5.33. □

Theorem 5.33. $\mathbb{N}\mathrm{Rat}\,\Sigma_2^* \times \cdots \times \Sigma_k^*$ *is a sub-semiring of a skew field.*

This result is the direct consequence of a series of classical results in mathematics which we shall not prove here (cf. for instance [11]) but simply state.

Definition 5.34 (Hahn–Malcev–Neumann). *Let S be a semiring and G an ordered group. We write $S_{\mathsf{wo}}\langle\langle G \rangle\rangle$ to denote the set of series on G with coefficients in S whose support is a well-ordered subset of G.*

Theorem 5.35 (Birkhoff–Tarski–Neumann–Iwazawa[17]). *A finite direct product of free groups is ordered.*

Theorem 5.36 (Malcev–Neumann). *If S is a skew field and G an ordered group, then $S_{\mathsf{wo}}\langle\langle G \rangle\rangle$ is a skew field.*

[17] And possibly others.

Theorems 5.35 and 5.36 imply that $S_{\mathsf{wo}}\langle\langle F(\Sigma_2) \times \cdots \times F(\Sigma_k)\rangle\rangle$ is a skew field (here $F(\Sigma)$ is the free group generated by Σ). To deduce Theorem 5.33, we must also ensure that $S\mathrm{Rat}\,\Sigma^*$—in fact $S\langle\langle\Sigma^*\rangle\rangle$—is included in $S_{\mathsf{wo}}\langle\langle F(\Sigma)\rangle\rangle$, respectively that

$$S\langle\langle\Sigma_2^* \times \cdots \times \Sigma_k^*\rangle\rangle \subseteq S_{\mathsf{wo}}\langle\langle F(\Sigma_2) \times \cdots \times F(\Sigma_k)\rangle\rangle;$$

that is, to be more precise, that we can order $F(\Sigma_2) \times \cdots \times F(\Sigma_k)$ in such a way that the above inclusion is true and this is not difficult either.

Now, by straightforward computations, 1-way k-tape Turing machines are faithfully modelised by automata over $\Sigma_1^* \times \Sigma_2^* \times \cdots \times \Sigma_k^*$ and two *deterministic* such machines are equivalent if and only if the corresponding automata are equivalent as automata over $\Sigma_1^* \times \Sigma_2^* \times \cdots \times \Sigma_k^*$ *with multiplicity in* $\mathbb{N}$.

Corollary 5.37 ([22]). *The equivalence of* 1*-way* k*-tape deterministic Turing machines is decidable.*

5.3.2 Recurrence Relations

Another consequence of Theorem 5.16 is the generalisation to series over non-commuting variables of the characterisation by linear recurrences of coefficients of rational series over one variable (recall also Lemma 2.27).

Theorem 5.38 ([46]). *A series* r *of* $S\langle\langle\Sigma^*\rangle\rangle$ *is recognisable if and only if there exists a finite prefix-closed subset* P *and its border* $C = P\Sigma\backslash P$*, such that for each pair* (c, p) *in* $C \times P$*, there exists a coefficient* $s_{c,p}$ *in* S *such that*

$$\forall g \in \Sigma^*,\ \forall c \in C \quad (r, cg) = \sum_{p\in P} s_{c,p}(r, pg). \tag{30}$$

Proof. Let P be a word base for an S-representation $\mathfrak{A} = (\lambda, \mu, \nu)$ that recognises r and (λ', μ', ν') the S-representation computed as in Proposition 5.23. For each $c = pa$ in C and all q in P, we set $s_{c,q} = (a\mu')_{p,q}$. From (29) follows that, for all g in Σ^*, it holds:

$$(r, cg) = \lambda \cdot p\mu \cdot a\mu \cdot g\mu \cdot \nu = \sum_{q\in P} a\mu'_{p,q}\lambda \cdot q\mu \cdot g\mu \cdot \nu = \sum_{q\in P} s_{c,q}(r, qg).$$

Conversely, (30) implies that every quotient $f^{-1}r$ belongs to the subspace T generated by $p^{-1}r$ for p in P. This last property is trivially verified if f is in P and (30) can be rewritten as

$$\forall c \in C \quad c^{-1}r = \sum_{p\in P} s_{c,p}p^{-1}r;$$

that is, the property is verified for f in C. *A contrario*, suppose that $f^{-1}r$ does not belong to T; by Lemma 2.27, we have $f = cg$ and choose f such that g is of minimal length. By (30), we have, for all h in Σ^*,

$$(r, cgh) = \sum_{p \in P} s_{c,p}(r, pgh) \qquad \text{that is,} \qquad f^{-1}r = \sum_{p \in P} s_{c,p}(pg)^{-1}r.$$

For each p in P, either pg is in P, or $pg = c'g'$ with c' in C, then $|c'| > |p|$; hence, $|g'| < |g|$ and $(pg)^{-1}r$ is in T by the assumption of minimality of g. Hence, $f^{-1}r$ belongs to T, which is a contradiction. Also, r is recognisable by Theorem 5.16. □

Remark 5.39. If $\Sigma = \{a\}$, every prefix-closed subset of Σ^* has the form $P = \{\varepsilon, a, \ldots, a^{j-1}\}$ for some integer j, and C is a singleton: $C = \{a^j\}$. Equation (30) becomes

$$\forall n \in \mathbb{N} \quad \big(r, a^{n+j}\big) = s_{j-1}\big(r, a^{n+j-1}\big) + s_{j-2}\big(r, a^{n+j-2}\big) + \cdots + s_0\big(r, a^n\big);$$

that is, a linear recurrence in its standard form.

Another way to exploit Proposition 5.23, is by 'computing' the coefficients of a *reduced representation* of a recognisable series as a function of the coefficients of the series itself. Going from the series back to the representation does not so much correspond to an effective procedure like those described in Proposition 5.23 and Theorem 5.38, as it expresses a fundamental property of recognisable series on a field (see an application with Theorem 6.4).

Proposition 5.40 ([46]). *Let S be a skew field, r an S-recognisable series of rank n, and (λ, μ, ν) a reduced representation of r. There exist two sets of n words: $P = \{p_1, p_2, \ldots, p_n\}$ and $Q = \{q_1, q_2, \ldots, q_n\}$ (which we can choose to be respectively prefix-closed and suffix-closed) and two $n \times n$ matrices α_P and β_Q such that*

$$\forall f \in \Sigma^* \quad f\mu = \alpha_P \cdot ((r, p_i f q_j)) \cdot \beta_Q,$$

where $((r, p_i f q_j))$ denote the $n \times n$ matrix whose entry (i, j) is $(r, p_i f q_j)$.

5.3.3 From Equivalence to Conjugacy

At Section 3.3, we have seen that it directly follows from the definition that two conjugate automata are equivalent (Proposition 3.23). For certain semirings S, this statement can be given a kind of converse, which reads as follows.

Theorem 5.41 ([4]). *Let S be $\mathbb{B}$, $\mathbb{N}$, $\mathbb{Z}$, or any (skew) field. Two S-automata are equivalent if and only if there exists a third S-automaton that is conjugate to both of them.*

The proof of Theorem 5.41 relies on the idea of *joint reduction* which is defined by means of the notion of *representation*. Let $\mathfrak{A} = \langle \lambda, \mu, \nu \rangle$ be an S-representation of dimension Q and the associated map $\Psi_{\mathfrak{A}} \colon \Sigma^* \to S^Q$. We have already seen (Proposition 5.23 and Remark 5.24) that, in the two contrasting cases of the Boolean semiring and of a field, we can choose a word base P such that:

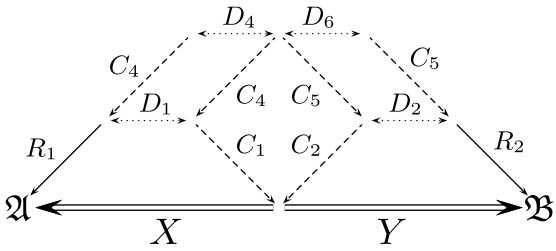


Fig. 14. Structural decomposition of the equivalence of two S-automata

(i) $\{\lambda \cdot p\mu \mid p \in P\}$ is a set of vectors, which is equal to $(\Sigma^*)\Psi_{\mathfrak{A}}$ in the Boolean case, which generates the same S-vector space in the field case.
(ii) There exists an automaton $\mathfrak{R}$ which is conjugate to $\mathfrak{A}$ by the transfer matrix X whose rows are the vectors $\{\lambda \cdot p\mu \mid p \in P\}$.

Let now $\mathfrak{A} = \langle \lambda, \mu, \nu \rangle$ and $\mathfrak{B} = \langle \eta, \kappa, \chi \rangle$ be two S-representations of dimension Q and R, respectively, and let $\mathfrak{C}$ be the *sum* of $\mathfrak{A}$ and $\mathfrak{B}$: $\mathfrak{C} = \langle \zeta, \pi, \omega \rangle$ is an S-representation of dimension $Q \cup R$, $\zeta = [\lambda\ \eta]$ is the horizontal concatenation of λ and η, $\omega = \left[\begin{smallmatrix} \nu \\ \chi \end{smallmatrix}\right]$ the vertical concatenation of ν and χ, and $\pi = \left[\begin{smallmatrix} \mu & 0 \\ 0 & \kappa \end{smallmatrix}\right]$ is the representation whose diagonal blocs are μ and κ. We perform the same construction as before on $\mathfrak{C}$; we consider the set of vectors $(\Sigma^*)\Psi_{\mathfrak{C}} = \{[\lambda \cdot f\mu\ \eta \cdot f\kappa] \mid f \in \Sigma^*\}$ and look for a *finite* set V of vectors $[x\ y]$ which, roughly speaking, generates the same S-semi-module as $(\Sigma^*)\Psi_{\mathfrak{C}}$.

The computation of V provides indeed at the same time an automaton $\mathfrak{Z}$ which is conjugate to $\mathfrak{C}$ by the transfer matrix Z whose rows are the vectors in V. *If* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are equivalent*, then $\mathfrak{Z}$, or a slight modification of it (depending on which semiring S the computations are currently done), is conjugate to both $\mathfrak{A}$ and $\mathfrak{B}$ by the transfer matrices X and Y, respectively, where X and Y are respectively the 'left' and 'right' parts of the matrix Z. In every case listed in Theorem 5.41, the finite set V is effectively computable, a proof that has to be done separately for each case (cf. [4]).

Together with the result of decomposition of conjugacy by means of a sequence of co-covering, circulation, and covering (Theorem 3.33), and Proposition 3.36 that allows us to build diagrams upward; this result yields a structural decomposition of the equivalence of two S-automata as shown in Fig. 14. In the case $S = \mathbb{N}$, this decomposition takes the following form.

Corollary 5.42. *Two equivalent* $\mathbb{N}$*-automata can be transformed, one into the other, by a chain of two state-splittings (in- and out-) and two state-mergings (out- and in-).*

6 Support of Rational Series

It follows directly from Proposition 4.5 that for any (graded) monoid M, we have the following corollary.

Corollary 6.1. *If S is a positive semiring, the support of an S-recognisable series over M is a recognisable subset of M.*

The assumption on S is necessary, even in the case where M is a free monoid Σ^*, as shown by the following example.

Example 6.2 (Example 4.2 continued). We have seen that $u_1 = \sum_{f\in\Sigma^*} |f|_b f$ is a $\mathbb{Z}$-rational series, and thus so is $r_1 = \sum_{f\in\Sigma^*} |f|_a f$. The series $z_1 = r_1 - u_1 = \sum_{f\in\Sigma^*}(|f|_a - |f|_b)f$ is a $\mathbb{Z}$-rational series. The complement of $\mathsf{supp}\ z_1 = \{f \in \Sigma^* \mid |f|_a \neq |f|_b\}$ is the language $Z_1 = \{f \in \Sigma^* \mid |f|_a = |f|_b\}$, which we know is not rational.

In this short section, we study certain conditions which ensure the rationality of the support of a series, and some closure properties of the family of languages thus defined. We end with several undecidable properties for $\mathbb{Z}$-rational series, somewhat surprising in this context where properties seem to be all decidable and effective.

Recall that a series r of $S\langle\langle\Sigma^*\rangle\rangle$ is fundamentally a map from Σ^* to S. It is therefore natural to write, for every subset U of S, Ur^{-1} for the set of words of Σ^* whose coefficient in r belongs to U:

$$Ur^{-1} = \{f \in \Sigma^* \mid (r, f) \in U\}.$$

The first result concerns *locally finite* semirings (defined in Chap. 1).

Proposition 6.3. *Let S be a* locally finite *semiring and let r be an S-rational series over Σ^*. For all subsets U of S, Ur^{-1} is rational.*

Proof. Since r is also recognisable, r is recognised by a S-representation (λ, μ, ν), of finite dimension Q, that is, $\mu\colon \Sigma^* \to S^{Q\times Q}$ is a morphism. Since S is locally finite, the image $(\Sigma^*)\mu = M$ is a *finite submonoid* of $S^{Q\times Q}$. The language Ur^{-1} is recognised by the morphism $\mu\colon \Sigma^* \to M$, a well-known characterisation of rational (or recognisable) languages of Σ^*. □

Another way to state (and to prove indeed) Proposition 6.3 is to remark that if S is locally finite, then the *reachability set* $(\Sigma^*)\Psi_{\mathfrak{A}}$ of any S-representation $\mathfrak{A}$ is finite—opening the way to the immediate construction for equivalent deterministic or minimal automata, a basic fact that seems to have been often overlooked, and thus often rediscovered (cf. also Chap. 12). To express it in another way again: *Counting in a (locally) finite semiring is not counting.*

Proposition 6.3 generalises in a remarkable way if S is a field. But it is not a trivial remark anymore; it follows from the whole algebraic theory we have built in this case.

Theorem 6.4 ([46]). *Let S be a (skew) field. If r is an S-rational series over Σ^* with a finite image, then kr^{-1} is rational for all k in S.*

Proof. Let (λ, μ, ν) be a reduced representation that recognises r. By Proposition 5.40, the image $(\Sigma^*)\mu$ is a *finite sub-monoid* of $S^{Q\times Q}$ if r has a finite image and the conclusion follows as in Proposition 6.3. □

Since the family of supports of S-rational series over Σ^* strictly contains $\operatorname{Rat}\Sigma^*$, a natural question is to ask under which operations this family is closed. The answer certainly depends on S; a fairly complete one can be given for sub-semirings of $\mathbb{R}$.

Proposition 6.5 ([46]). *Let S be a sub-semiring of $\mathbb{R}$. The set of supports of S-rational series on Σ^* contains $\operatorname{Rat}\Sigma^*$ and is closed under union, product, star, and intersection.*

Proof. The first assertion is a restatement of Proposition 3.12. Since $S\operatorname{Rat}\Sigma^*$ is closed under the Hadamard product, we deduce first the closure by intersection, then because r and $r \odot r$ have the same support, it follows that every support of an S-rational series is the support of an S-rational series with *non-negative coefficients*. Then for such series, we clearly have

$$\operatorname{supp}(r + r') = \operatorname{supp} r \cup \operatorname{supp} r', \qquad \operatorname{supp}(rr') = \operatorname{supp} r \operatorname{supp} r' \qquad \text{and}$$
$$\operatorname{supp}(r^*) = (\operatorname{supp} r)^*. \quad \square$$

The closure under morphisms and inverse morphisms is somewhat more difficult to establish.

Proposition 6.6 (Fliess [17]). *Let S be a sub-semiring of $\mathbb{R}$. The set of supports of S-rational series over Σ^* is closed under morphisms and inverse morphisms.*

The set $\operatorname{Rat}\Sigma^*$ is also closed under complement, but if S is not positive, the set of supports of S-rational series can strictly contain $\operatorname{Rat}\Sigma^*$. The closure under complement is precisely characteristic of membership of $\operatorname{Rat}\Sigma^*$ as stated in the following result. Besides the reduction theory, its proof is based upon the strongest version of the iteration theorem (or pumping lemma) for rational languages, due to A. Ehrenfeucht, R. Parikh, and G. Rozenberg [13], and itself is based on Ramsey's theorem.

Theorem 6.7 (Restivo–Reutenauer [38]). *Let S be a (sub-semiring of a) skew field. If a language and its complement are each the support of an S-rational series over Σ^*, then this language is rational.*

We then construct, with this simple model of finite weighted automata, some series for which we cannot answer some elementary questions, as soon as the semiring of coefficients contains $\mathbb{Z}$.

Theorem 6.8. *It is undecidable if the support of a $\mathbb{Z}$-rational series over Σ^* is all of Σ^*.*

Proof. Let $\Delta = \{x, y\}$; the morphism $\alpha\colon \Delta^* \to \mathbb{N}^{2\times 2}$ defined by

$$x\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad y\alpha = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

is *injective* (cf. the automaton $\mathfrak{C}_1$ at Example 3.4).

Then let $\theta\colon \Sigma^* \to \Delta^*$ and $\theta\colon \Sigma^* \to \Delta^*$ be two morphisms. For i and j equal to 1 or 2, the series $r_{i,j}$ defined by

$$\forall f \in \Sigma^* \quad (r_{i,j}, f) = \big((f\theta)\alpha\big)_{i,j} - \big((f\mu)\alpha\big)_{i,j}$$

are $\mathbb{Z}$-rational, hence so are the series $u_{i,j} = r_{i,j} \odot r_{i,j}$, and the series

$$u = \sum_{i,j} u_{i,j}.$$

The support of u is not all of Σ^* if and only if there exists f such that $(u, f) = 0$; that is, since α is *injective*, if and only if $f\theta = f\mu$, which we know to be undecidable (Post Correspondence Problem). □

With the same construction, we easily obtain the following corollary.

Corollary 6.9. *Let r be a $\mathbb{Z}$-rational series over Σ^*. It is undecidable whether:*

(i) r has infinitely many coefficients equal to zero.
(ii) r has at least one positive coefficient.
(iii) r has infinitely many positive coefficients.

Corollary 6.10. *It is undecidable whether the supports of two $\mathbb{Z}$-rational series over Σ^* are equal.*

7 Notes

I am grateful to M. Droste, Ch. Reutenauer, and W. Kuich who pointed out some interesting references to me.

7.1 General Sources

As already mentioned in the Introduction, this chapter is essentially an epitome of Chap. III and of a part of Chap. IV of [43] where more details, proofs, and examples are to be found. More precise references to some of them are given below.

A classical, and above all pioneering, reference on the subject is the treatise by S. Eilenberg [14] whose influence is willingly acknowledged. Each of the references quoted in the Introduction [45, 29] or in Chaps. 1 and 3 develop a particular point of view that is worth interest. But the most advanced one is the book of Jean Berstel and Christophe Reutenauer [5], a new revised edition of which is now available and anyone really interested in weighted automata should certainly not miss this work.

7.2 Notes to Sect. 2: Rational Series

One can say that it is equation (2) that justifies the choice of Eilenberg, op. cit., to call *rational* what was called *regular* in the foregoing literature. Schützenberger and his school, to which I acknowledge membership, followed him but one must recognise it has not been a universal move. If the terminology is still disputable for languages, and expressions, I do not think the question may even be asked when it comes to series. On the other hand and in the same work, Eilenberg calls a monoid with the property that every element is finitely decomposable a *locally finite monoid*. This terminology inconveniently conflicts with another accepted meaning of the phrase: a monoid such that every finitely generated submonoid is finite (cf. [49]).

I was led to define *strong semirings*, a terminology suggested to me by J. van der Hoeven, to be able to prove the equivalence between the existence of the star of an arbitrary series and that of the star of its constant term.

7.3 Notes to Sect. 3: Weighted Automata

The construction of $\mathfrak{S}_{\mathsf{E}}$ is the version given in [33] of the generalisation to weighted automata of the construction of the *Glushkov automaton* or *position automaton* first given by Caron and Flouret [9].

In a sense the Fundamental Theorem is what Kleene showed for automata over Σ [26], or its usual weighted generalisation (often called the Kleene–Schützenberger theorem). However, because these results apply to automata over free monoids, their standard form—cf. Theorem 2.12, Chap. 3—states the identity between rational and recognisable languages or series, which *no longer holds* for automata (weighted or otherwise) over arbitrary monoids. Kleene's theorem was therefore split in two, as it were: one part which holds for automata over arbitrary monoids and which, considering what the proof involves, concentrates the substance of the theorem; and one part which holds only for automata over free monoids and which is nearly a formality; this distinction seems to appear for the first time in [42].

Proposition 3.21 can be credited to Conway [12] and Krob [27]; an elementary proof is given in [43, 44].

The matter of Sect. 3.3 is taken from [43] and [3]. Conjugacy of $\mathfrak{A}$ to $\mathfrak{B}$ by X is called *simulation* from $\mathfrak{A}$ into $\mathfrak{B}$ in [7]. In a different setting, this kind of mapping was called *morphism of 'modules sériels'* by Fliess in [18]. The definition of S-covering as conjugacy by an amalgamation matrix is a hint for similarity between S-coverings and *state amalgamation* in *symbolic dynamical systems* [31, Sect. 2.4]. If $\mathfrak{B}$ is obtained from $\mathfrak{A}$ by an In-amalgamation, then $\mathfrak{A}$ is an $\mathbb{N}$-covering of $\mathfrak{B}$. But the converse is not true. Roughly speaking, and with the notations of Proposition 3.28, $\mathfrak{A} = \langle I, E, T\rangle$ is an S-covering of $\mathfrak{B}$ if the rows with 'equivalent' indices of the matrix $E{\cdot}H_{\varphi}$ are equal while $\mathfrak{B}$ is obtained by amalgamation from $\mathfrak{A}$ if the rows with 'equivalent' indices of the matrix E are equal. The notion of 'equisubtractivity' used in Sect. 3.3.3 in order to

express conjugacy in terms of coverings and co-coverings is very similar to a property introduced by Tarski in [51] where an extension of Lemma 3.35 to infinite sums is established.

The presentation of the minimal S-quotient is taken from [43], whereas the notion itself probably exists in many other works; for instance, two S-automata are *in bisimulation* if and only if their minimal S-quotients are isomorphic.

7.4 Notes to Sect. 4: Recognisable Series

The definition of representations in the form (λ, μ, ν) is due to Fliess [18]. Lemma 4.10 is a classic statement in matrix theory and can be found already in Gröbner [21] (cf. also [29, Theorem 4.33]). Theorem 4.12 is due to Schützenberger [48], including the more general formulation of Theorem 4.13. Theorem 4.18 is also due to Fliess [19]; the proof given here is that of [43].

The 'shuffle identity' (20) is an unpublished result of M. Petitot and was indicated to me by M. Waldschmidt (personal communication); the proof I gave for it in [43] was the starting point of [3].

The matter of Sect. 4.3, and especially the definition of weighted relations, is taken from Chap. IV of [43]. Another theory of weighted relations, slightly different from what I have very briefly sketched here, is that of Jacob [24, 25]. It consists of defining with *regulated rational transductions* the largest possible family of relations which satisfy the evaluation and composition theorems and which correspond to total maps (and hence maps whose composition is also always defined), and to do that *independently of the semiring of coefficients*. This point of view was adopted in related works [45, 29] which popularised the work of Jacob.

7.5 Notes to Sect. 5: Series over a Free Monoid

Some authors speak of the *translation* of a series instead of quotient; I have preferred to use the same term as for languages.

The original work is due to Schützenberger [46, 47]. The characterisation of recognisable series (Theorem 5.3) is a generalisation, due to Jacob [24], of the property stated by Fliess for the case of series on a field [18].

The derivation of weighted expressions is a generalisation of V. Antimirov's work [1] (where derived terms were called *partial derivatives*). We note once more that the introduction of weights clarifies and structures a result on languages, even if having to take into account that not necessarily positive semirings adds a certain complexity. This presentation is taken from [32]. With somewhat different techniques, Rutten [40, 41] also proved Theorem 5.10 and Proposition 5.12.

The original work for reduction of representations is again from Schützenberger [46]. The presentation here follows roughly [5] but as in a background

and owes much to my discussions with S. Lombardy. It keeps the Hankel matrix of a series—which could be given the central role as M. Fliess did in [18]—as a subliminal object. It is important for the sequel that the theory is generalised to non-commutative fields. In [20], it was also observed that Schützenberger's reduction algorithm applies to the case of series on a skew field, but with a reference to a previous theory of non-commutative determinants [39]. The cubic complexity of the reduction algorithm was already established in [8].

The problem of the decidability of equivalence of *deterministic* k-tape automata was posed in [37] and was solved for $k = 2$ by M. Bird [6] by an *ad hoc* method, then by L. Valiant [52] as a corollary of the decidability of the equivalence of 'finite-turn' deterministic pushdown automata. The problem remained open for $k \geqslant 3$ until the solution in [22]. The material for Theorems 5.35 and 5.36 is standard if not elementary algebra, and is explained in sufficiently comprehensive treatises such as [11]. A self-contained presentation and proof of this is given in [43, IV.7]. The original proof of Theorem 5.36 by Neumann [36] has been greatly simplified by Higman [23] where he proved what is often known as 'Higman's lemma'. The Russian version of the same result was proved in [34].

Section 5.3.2 is adapted from [5] and Sect. 5.3.3 from [4]. A result analogous to Theorem 5.41 holds for functional transducers as well, but this, its proof, and its consequences somewhat fall out of the scope of this chapter (cf. [4]).

7.6 Notes to Sect. 6: Support of Rational Series

The subject is hardly touched there and the reader is referred once again to [45] or to [5]. Theorem 6.4 has been generalised to commutative rings by Sontag [50]. The proof of Theorem 6.8 is taken from [14].

References

1. V. Antimirov. Partial derivatives of regular expressions and finite automaton constructions. *Theoretical Computer Science*, 155:291–319, 1996.
2. A. Arnold. *Systèmes de transitions finis et sémantique des processus communiquants*. Masson, Paris, 1992. Translation: *Finite Transitions Systems*. Prentice–Hall, New York, 1994.
3. M.-P. Béal, S. Lombardy, and J. Sakarovitch. On the equivalence of $\mathbb{Z}$-automata. In *ICALP 2005*, volume 3580 of *Lecture Notes in Computer Science*, pages 397–409. Springer, Berlin, 2005.
4. M.-P. Béal, S. Lombardy, and J. Sakarovitch. Conjugacy and equivalence of weighted automata and functional transducers. In *CSR 2006*, volume 3967 of *Lecture Notes in Computer Science*, pages 58–69. Springer, Berlin, 2006.

5. J. Berstel and C. Reutenauer. *Les séries rationnelles et leurs langages*. Masson, Paris, 1984. Translation: *Rational Series and Their Languages*. Springer, Berlin, 1988. New revised English edition available from http://www-igm.univ-mlv.fr/~berstel/.
6. M. Bird. The equivalence problem for deterministic two-tape automata. *Journal of Computer and System Sciences*, 7:218–236, 1973.
7. S.L. Bloom and Z. Ésik. *Iteration Theories*. Springer, Berlin, 1993.
8. A. Cardon and M. Crochemore. Détermination de la représentation standard d'une série reconnaissable. *Theoretical Informatics and Applications, RAIRO*, 14:371–379, 1980.
9. P. Caron and M. Flouret. Glushkov construction for multiplicities. In A. Paun and S. Yu, editors, *CIAA 2000*, volume 2088 of *Lecture Notes in Computer Science*, pages 67–79. Springer, Berlin, 2001.
10. J.-M. Champarnaud and D. Ziadi. Canonical derivatives, partial derivatives and finite automaton constructions. *Theoretical Computer Science*, 289:137–163, 2002.
11. P.M. Cohn. *Algebra*. Wiley, New York, 1974. 2nd edition: volume I, 1982; volume II, 1989; volume III, 1991.
12. J.H. Conway. *Regular Algebra and Finite Machines*. Chapman & Hall, London, 1971.
13. A. Ehrenfeucht, R. Parikh, and G. Rozenberg. Pumping lemmas for regular sets. *SIAM Journal on Computing*, 10:536–541, 1981.
14. S. Eilenberg. *Automata, Languages and Machines, volume A*. Academic Press, San Diego, 1974.
15. S. Eilenberg and M.P. Schützenberger. Rational sets in commutative monoids. *Journal of Algebra*, 13:173–191, 1969.
16. C.C. Elgot and J.E. Mezei. On relations defined by generalized finite automata. *IBM Journal of Research and Development*, 9:47–68, 1965.
17. M. Fliess. Formal languages and formal power series. In *Séminaire Logique et Automates, IRIA, 1971*, pages 77–85.
18. M. Fliess. Matrices de Hankel. *Journal de Mathématiques Pures et Appliquées*, 53:197–222, 1974. Erratum in: *Journal de Mathématiques Pures et Appliquées*, 54, 1975.
19. M. Fliess. Sur divers produit de séries formelles. *Bulletin de la Société Mathématique de France*, 102:181–191, 1974.
20. M. Flouret and E. Laugerotte. Noncommutative minimization algorithms. *Information Processing Letters*, 64:123–126, 1997.
21. W. Gröbner. *Matrizenrechnung*. Oldenburg, München, 1956.
22. T. Harju and J. Karhumäki. The equivalence problem of multitape finite automata. *Theoretical Computer Science*, 78:347–355, 1991.
23. G. Higman. Ordering by divisibility in abstract algebra. *Proceedings of the London Mathematical Society. Second Series*, 2:326–336, 1952.
24. G. Jacob. Représentations et substitutions matricielles dans la théorie algébrique des transductions. Thèse Sci. Math. Univ. Paris VII, 1975.
25. G. Jacob. Sur un théorème de Shamir. *Information and Control*, 27:218–261, 1975.

26. S.C. Kleene. Representation of events in nerve nets and finite automata. In C. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, Princeton, 1956.
27. D. Krob. Complete systems of B-rational identities. *Theoretical Computer Science*, 89:207–343, 1991.
28. D. Krob. The equality problem for rational series with multiplicities in the tropical semiring is undecidable. In W. Kuich, editor, *ICALP'92*, volume 623 of *Lecture Notes in Computer Science*, pages 101–112. Springer, Berlin, 1992.
29. W. Kuich and A. Salomaa. *Semirings, Automata, Languages*. Springer, Berlin, 1986.
30. F.W. Levi. On semigroups. *Bulletin of the Calcutta Mathematical Society*, 36:141–146, 1944 and 38:123–124, 1946.
31. D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.
32. S. Lombardy and J. Sakarovitch. Derivation of rational expressions with multiplicity. In *MFCS'02*, volume 2420 of *Lecture Notes in Computer Science*, pages 471–482. Springer, Berlin, 2002.
33. S. Lombardy and J. Sakarovitch. Derivation of rational expressions with multiplicity. *Theoretical Computer Science*, 332:141–177, 2005.
34. A.I. Malcev. On the embedding of group algebras in division algebras. *Doklady Akademii Nauk SSSR (N.S.)*, 60:1409–1501, 1948 (in Russian).
35. J. McKnight. Kleene's quotient theorems. *Pacific Journal of Mathematics*, 14:43–52, 1964.
36. B.H. Neumann. On ordered division ring. *Transactions of the American Mathematical Society*, 66:202–252, 1949.
37. M.O. Rabin and D. Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3:125–144, 1959. Reprinted in: E. Moore, editor, *Sequential Machines: Selected Papers*, Addison–Wesley, Reading, 1965.
38. A. Restivo and C. Reutenauer. On cancellation properties of languages which are support of rational series. *Journal of Computer and System Sciences*, 29:153–159, 1984.
39. A.R. Richardson. Simultaneous linear equations over a division algebra. *Proceedings of the London Mathematical Society*, 28:395–420, 1928.
40. J.M. Rutten. Automata, power series, and coinduction: Taking input derivatives seriously. In J. Wiedermann, P. van Emde Boas, and M. Nielsen, editors, *ICALP'99*, volume 1644 of *Lecture Notes in Computer Science*, pages 645–654. Springer, Berlin, 1999.
41. J.M. Rutten. Behavioural differential equations: A coinductive calculus of streams, automata, and power series. *Theoretical Computer Science*, 308:1–53, 2003.
42. J. Sakarovitch. Kleene's theorem revisited. In A. Kelemenova and K. Kelemen, editors, *Trends, Techniques and Problems in Theoretical Computer*

Science, volume 281 of *Lecture Notes in Computer Science*, pages 39–50. Springer, Berlin, 1987.
43. J. Sakarovitch. *Éléments de théorie des automates*. Vuibert, Paris, 2003. Corrected English edition: *Elements of Automata Theory*, Cambridge University Press, Cambridge, 2009.
44. J. Sakarovitch. The language, the expression and the (small) automaton. In *CIAA 2005*, volume 3845 of *Lecture Notes in Computer Science*, pages 15–30. Springer, Berlin, 2005.
45. A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series*. Springer, Berlin, 1977.
46. M.P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.
47. M.P. Schützenberger. Certain elementary families of automata. In *Symposium on Mathematical Theory of Automata, 1962*, pages 139–153.
48. M.P. Schützenberger. On a theorem of R. Jungen. *Proceedings of the American Mathematical Society*, 13:885–889, 1962.
49. I. Simon. Limited subsets of a free monoid. In *FOCS'78, 1978*, pages 143–150.
50. E.D. Sontag. On some questions of rationality and decidability. *Journal of Computer and System Sciences*, 11:375–385, 1975.
51. A. Tarski. *Cardinal Algebras*. Oxford University Press, London, 1949.
52. L.G. Valiant. The equivalence problem for deterministic finite-turn pushdown automata. *Information and Control*, 25:123–133, 1974.