Derivatives of rational expressions with multiplicity

Sylvain Lombardy[1] and Jacques Sakarovitch[2]

[1] Université Paris 7
Laboratoire d'Informatique Algoritmique: Fondements et Applications
2, place Jussieu 75251 Paris Cedex 05, France
lombardy@liafa.jussieu.fr

[2] CNRS - Laboratoire Traitement et Communication de l'Information (UMR 5141)
Ecole Nationale Supérieure des Télécommunications
46, rue Barrault 75634 Paris Cedex 13, France
sakarovitch@enst.fr

The second author is pleased to dedicate this article
to Rick and Harriet Meyer.
During summer 2000, he was able to develop the first ideas
that gave birth to that paper in the best possible environment
thanks to their wonderful hospitality.

Abstract. This paper addresses the problem of turning a rational (*ie* regular) expression into a finite automaton. We formalize and generalize the idea of "*partial derivatives*" introduced in 1995 by V. Antimirov, in order to obtain a construction of an automaton with multiplicity from a rational expression describing a formal power series with coefficients in a semiring.

We first define precisely what is such a rational expression with multiplicity and which hypothesis should be put on the semiring of coefficients in order to keep the usual identities.

We then define the *derivative* of such a rational expression as a *linear combination* of expressions called *derived terms* and we show that all derivatives of a given expression are generated by a finite set of derived terms, that yields a finite automaton with multiplicity whose behaviour is the series denoted by the expression. We also prove that this automaton is a *quotient* of the standard (or Glushkov) automaton of the expression. Finally, we propose and discuss some possible modifications to our definition of derivation.

Introduction

The problem of turning a rational (*ie* regular) expression into a finite automaton has attracted much attention since the beginning of the theory of automata and has been since then an area of active research. Lexical analysis, pattern matching, and many other problems involving sequences of symbols give strong motivation for a dedicated study of this subject.

The story begins with McNaughton and Yamada's paper ([15]) which contains two algorithms: one turning an automaton into an expression (usually known as McNaughton–Yamada algorithm) and one turning an expression into a *deterministic* automaton, thus giving an effective proof of Kleene's Theorem. In the numerous subsequent works on the subject, let us pick three papers that are milestones for us.

In [11], V. Glushkov has given an algorithm that associates to every rational expression of *litteral length* n a (non deterministic) automaton with $n+1$ states — which is often called *the Glushkov*, or *the position, automaton* of the expression. We shall call it here the standard automaton of the expression. (Indeed, the subset construction applied to the position automaton gives back the one defined by McNaughton and Yamada.)

In [3], J. Brzozowski defined the *derivatives* of a rational expression. He showed that, *modulo* the axioms of associativity, commutativity, and idempotency of the addition (on the set of languages) — the ACI properties — the set of derivatives of a given expression is finite. It is well-known that a language is rational iff it has a finite number of (left) quotients (or left residuals). The virtue of Brzozowski's result is not only, not so much, that it yields both a new proof for (one direction of) Kleene's Theorem and an algorithm turning an expression into a deterministic finite automaton but that it lifts this property of languages at the symbolic level of expressions.

In 1995, V. Antimirov made another fundamental contribution by defining the so-called "*partial derivatives*" of an expression ([1]). Using his own words, "this construction allows us to take into account the ACI-properties of only *some* of the occurence of "+" in a regular term r, namely of those which appear at the very upper level of r".

Roughly speaking, the derivation proposed by Antimirov has two effects. First, it performs the "Brzozowski" derivation and, second, it breaks the derivatives into "parts" — hence the name *partial derivatives* — such that the result is the sum of the parts. This construction has several outcomes: the number m of partial derivatives is not only finite but "small" : smaller than or equal to the *litteral length* of the expression; they are easier to compute than the Brzozowski derivatives and they yield a *non-deterministic* finite automaton with (at most) $m+1$ states; finally, the subset construction applied to that automaton gives back the deterministic one computed by Brzozowski's algorithm. (The computation of Antimirov automaton has been shown to be of quadratic complexity by Champarnaud and Ziadi [7].)

The main purpose of this paper is at the same time a *formalization* and a *generalization* of Antimirov's construction from rational expressions and languages to rational expressions *with multiplicity* and *formal power series*, yielding what we call *derived terms* which correspond to partial derivatives and an automaton with multiplicity which is the counterpart of Antimirov automaton. As it turns out, this is another example where taking multiplicities into account and considering series rather than languages yield simplification and better understanding of constructions and results on languages. The paper is organized into three parts: the first one (Sec. 1 and 2) sets the framework, the second and main part (Sec. 3 to 5) describes the derivation of an expression with multiplicity and the construction of the automaton of derived terms and the third one (Sec. 6 to 8) provides comments, complements and variations to the main construction.

When one deals with (formal) power series, one makes use — implicitely or explicitely — of a topology on the semiring $\mathbb{K}$ of coefficients. Unlike many authors, and along the line we develop for the theory of automata with multiplicity (*cf.* [18]), we do not want to restrict ourselves to the discrete topology on $\mathbb{K}$ and want to be able to take the star of series which are not necessarily proper. The price we pay for that generalization and in order to recover the classical identities is to restrict ourselves to a class of semirings, that we call *strong*, and that contains all useful semirings. We then carefully define rational expressions with multiplicity in a semiring $\mathbb{K}$, or $\mathbb{K}$-expressions.

The core of the paper begins with the definition of the derivation of a $\mathbb{K}$-expression with respect to a letter and then to a word. The main feature of our definition, that indeed realizes Antimirov's main idea, is that the result of the derivation of a $\mathbb{K}$-expression is not a $\mathbb{K}$-expression anymore but *a linear combination*, with coefficients in $\mathbb{K}$, of $\mathbb{K}$-expressions.

The generalization of Antimirov's results is then straightforward and gives our main result (Theorem 1 and Theorem 3). For any $\mathbb{K}$-expression E, there exists a finite set of $\mathbb{K}$-expressions, called the *derived terms* of E, such that the derivation of E with respect to any word is a linear combination with coefficients in $\mathbb{K}$ of derived terms. The number of derived terms is smaller than or equal to the *litteral length* of E. When $\mathbb{K} = \mathbb{B}$, the Boolean semiring, the derived terms of E are exactly the partial derivatives of E defined by Antimirov.

A direct consequence, or corollary, of these results is that a rational series belongs to a finitely generated $\mathbb{K}$-module which is *stable* by derivatives (or residuals); but it should be stressed that this new proof of a well-known (and fundamental) result is *not* the main account, or purpose of this paper. Rather, what we show here is that, when a rational series s is defined by an expression E, a *symbolic computation* performed on E *effectively* gives —through their expressions— a finite set of generators of the stable module that contains this series s. An automaton with multiplicity, and which recognizes the series denoted by E, is then built, the states of which are the derived terms.

Let us note that Brzozowski's point of view certainly could not be adopted in this context as the addition of series is not idempotent anymore. It is precisely when we tried to extend Brzozowski's derivatives to expressions with multiplicity that we were led to a definition that we understood to be equivalent to the one of Antimirov in the case of multiplicities taken in the Boolean semiring. After a first version of this work was completed, we learned that J.J. Rutten had established the equivalent of Theorem 3 ([16, 17]). The setting is different and the result is obtained as a byproduct of his theory of coinduction on series. It uses explicitely the same set of polynomials of expressions and yields as well the construction of a $\mathbb{K}$-automaton whose states are $\mathbb{K}$-expressions.

As a matter of fact, let us quote that Champarnaud and Duchamp [9] have also proposed to extend Brzozowski's derivatives to expressions with multiplicity. They define a set of equivalences on $\mathbb{K}$-expressions that could play a role similar to the ACI-equivalences on Boolean expressions. This does not yield a finite $\mathbb{K}$-automaton in general, since “deterministic” finite $\mathbb{K}$-automata do not always exist for $\mathbb{K}$-rational series. This approach is also orthogonal to Antimirov's one (and thus to ours), since the idea of *partial derivatives* was to avoid the use of any equivalence or (identities) on expressions in order to obtain a practical method to build an automata from an expression.

The third part deals with three different questions. The first one draws the attention on a a phenomenon that can arise only if $\mathbb{K}$ is not a positive semiring that is contains elements whose sum may be equal to $0_{\mathbb{K}}$. In this case, the derivation may generate some derived terms which are “useless” in the sense they appear with a null coefficient in the derivation of the expression with respect to any word. We call them *shadow terms*. The corresponding states may then be suppressed from the automaton of derived terms; such a possibility could not have been considered of course by Antimirov.

We next establish the generalization of a property that was recently shown by J.-M. Champarnaud and D. Ziadi ([8]): there exists a morphism from the position automaton of a rational expression onto its Antimirov automaton. We have generalized Antimirov automata to derived term $\mathbb{K}$-automata. In [4], Caron and Flouret have generalized the Glushkov's construction to expressions with multiplicity, yielding a $\mathbb{K}$-automaton which we call the *standard* $\mathbb{K}$-*automaton* of the expression (we recall their construction here). We present then, with the notion of $\mathbb{K}$-*covering*, the adequate generalization of morphisms of automata and are thus able to prove that there exists a $\mathbb{K}$-covering from the standard $\mathbb{K}$-automaton of an expression onto its $\mathbb{K}$-automaton of derived terms.

The true difficulty or meaning of this statement does not lay so much in its proof, that remains somewhat mechanical, but in the definition of derived term that it requires in order to hold. This definition that we give in Section 4 is indeed independent from the one of derivatives and more formal than the one we would otherwise have chosen, and actually that we choosed in a first version of that work ([13]).

Quite opposite, the last subject tackled in this third part is the study of the modifications that can be brought to derivation in order to get more compact or adequate derived term automata. We first describe two variants of the derivation: the *unitary derivation* that amounts to grouping successive scalar multiplications and the *breaking derivation* that consists in speeding up the splitting of the derived terms. Both are described via the definition of an appropriate *derivation with respect to the empty word*. And finally, we mention a "mild breaking derivation" that we have been using in a recent work of ours, where we have investigate the possibility of reversing McNaughton–Yamada algorithm.

A preliminary version of this work has been presented at the 2002 edition of the MFCS conference ([13]). The present version is significantly different and enriched.

Acknowledgements The authors are pleased to thank J.J. Rutten and several anonymous referees for their corrections and their comments that have helped them to improve the content and the presentation of this paper.

1 Rational series

Let us first recall, in order to fix the terminology and notation, the basic and standard definitions for rational series. We shall then set up a more general framework for rationality.

1.1 Basic definitions

For (rational) power series, their definitions, their notations and the related results, we refer to Berstel and Reutenauer's book [2] which we basically follow. Let A be a finite alphabet and $\mathbb{K}$ a semiring. The *associative* and *commutative* addition of $\mathbb{K}$ is denoted by $\oplus$, its multiplication simply by concatenation.

The semiring of formal power series over A^* with multiplicity in $\mathbb{K}$ is denoted by $\mathbb{K}\langle\langle A^*\rangle\rangle$. The inherited *commutative* addition in $\mathbb{K}\langle\langle A^*\rangle\rangle$ is denoted by $\oplus$, the (Cauchy) product by concatenation. The product of two series is always defined since each of its coefficients is obtained from a finite sum.

The coefficient of a word f of A^* in a series s of $\mathbb{K}\langle\langle A^*\rangle\rangle$ is denoted by $\langle s, f\rangle$. The *constant term* of s is the coefficient of the empty word 1_{A^*} in s and is denoted by $\mathsf{c}(s)$. A series is *proper* if $\mathsf{c}(s) = 0_{\mathbb{K}}$. The *proper part* of s, denoted by s_p, is the proper series which coincide with s on all words f different from 1_{A^*} and it holds: $s = \mathsf{c}(s)\, 1_{A^*} \oplus s_p$.

The *star* of a series is an infinite sum:

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

Giving a meaning to this infinite sum requires that $\mathbb{K}\langle\langle A^*\rangle\rangle$ be endowed with a topology and that the family $\{s^n\}_{n\in\mathbb{N}}$ be summable for that topology.

The standard topology chosen on $\mathbb{K}\langle\langle A^*\rangle\rangle$ is defined by setting the distance between two series s and t as the inverse of the length of a shortest word whose coefficients in s and t are different. This implies that the star of a proper series is always defined as the family $\{s^n\}_{n\in\mathbb{N}}$ is locally finite for a proper series s.

These definitions are sufficient to set up the theory of rational expressions with multiplicity and the subsequent developments we present here. However, it is possible to define the star of a series in a more general setting and this is what we shall be doing in the next subsection.

1.2 Strong semirings

The standard topology defined above on $\mathbb{K}\langle\langle A^*\rangle\rangle$ corresponds to the product topology — or weak convergence topology — derived from the *discrete* topology on $\mathbb{K}$. But this is an unnecessary, and even unpleasant, assumption. For instance, it holds:

$$(1/2)^* = 1 + 1/2 + (1/2)^2 + \cdots = 2$$

in the field $\mathbb{Q}$ equipped with the usual topology, whereas $(1/2)^*$ is not defined if $\mathbb{Q}$ is equipped with the discrete topology. Of course we would like the former equality be part of a theory of rationality.

In what follows, we suppose that $\mathbb{K}$ is a topological semiring ($\mathbb{K}$ is equipped with a topology and both addition and multiplication are continuous operations) and that $\mathbb{K}\langle\langle A^*\rangle\rangle$ is equipped with the product topology derived from the topology on $\mathbb{K}$. In other words, a family of series is summable if, and only if, the family of the coefficients of every word is summable.

The star of a series is still defined as above, but we need a further assumption on $\mathbb{K}$ in order to preserve the usual properties of that operation.

Definition 1. *We shall say that a topological semiring is* strong *if the product of two summable families is summable.*

This definition is an extension of the property of finite distributivity. For instance, every sub-semiring of $\mathbb{C}$, every positive semiring (boolean semiring, $(\max, +)$-semirings, subsets of a monoid) or every semiring with discrete topology is strong. The following is also easily verified:

Proposition 1. *If $\mathbb{K}$ is a strong semiring, so is $\mathbb{K}\langle\langle A^*\rangle\rangle$.*

The definition of a strong semiring is used in the proof of the following statement that is classical in the usual framework where rational series are considered (*cf.* [2, Exer. I.4.1]).

Lemma 1. *Let $\mathbb{K}$ be a strong topological semiring. Let s be a series in $\mathbb{K}\langle\langle A^*\rangle\rangle$, s_0 its constant term and s_p its proper part. Then s^* is defined iff s_0^* is defined and it holds:*

$$s^* = (s_0^*\, s_p)^* s_0^* = s_0^*(s_p\, s_0^*)^*. \tag{1}$$

Proof. The condition is necessary since $\langle s^n, 1_{A^*}\rangle = {s_0}^n$ and, if s^* is defined, the coefficients of 1_{A^*} in $\{s^n\}_{n\in\mathbb{N}}$ are a summable family.

Let assume, now, that $\{s_0^n\}_{n\in\mathbb{N}}$ is summable, and its sum is equal to s_0^*. For every pair of integers k and l, let:

$$P_{k,l} = \sum_{\substack{i_0,i_1,\ldots,i_k\in\mathbb{N}\\ i_0+i_1+\cdots+i_k=l}} s_0^{i_0}\, s_p\, s_0^{i_1}\, s_p \ldots s_0^{i_{k-1}}\, s_p\, s_0^{i_k}. \tag{2}$$

By convention, we set $P_{0,l} = s_0^l$ et $P_{k,0} = s_p^k$. We easily check that, for every n:

$$s^n = (s_0 + s_p)^n = \sum_{l=0}^{l=n} P_{n-l,l}. \tag{3}$$

By induction on k, we prove that the family:

$$F_k = \{s_0^{i_0}\, s_p\, s_0^{i_1}\, s_p \ldots s_0^{i_{k-1}}\, s_p\, s_0^{i_k} \mid i_0, i_1, \ldots, i_k \in \mathbb{N}\}$$

is summable in $\mathbb{K}\langle\langle A^*\rangle\rangle$, and its sum is:

$$Q_k = (s_0^*\, s_p)^k\, s_0^* = s_0^*\, (s_p\, s_0^*)^k.$$

Indeed, the hypothesis on s_0 guarantees the property for $k = 0$ and, moreover, that the family $G = \{s_0^n\, s_p \mid n \in \mathbb{N}\}$ is summable in $\mathbb{K}\langle\langle A^*\rangle\rangle$, with a sum equal to $s_0^* s_p$. The family F_{k+1} is the product of families G and F_k and the strongness of $\mathbb{K}$, and thus the strongness of $\mathbb{K}\langle\langle A^*\rangle\rangle$ leads to conclusion. (Figure 1 is given below with the hope it helps to follow the proof.)

Therefore, the family $\{P_{k,l} \mid l \in \mathbb{N}\}$ is summable for every k and its sum is equal to Q_k. The family $\{Q_k \mid k \in \mathbb{N}\}$ is locally finite, thus summable, and its sum is:

$$t = \sum_{k=0}^{\infty} Q_k = (s_0^*\, s_p)^*\, s_0^* = s_0^*\, (s_p\, s_0^*)^*. \tag{4}$$

We can then show by classical methods that the family $\{P_{k,l} \mid k, l \in \mathbb{N}\}$ is summable and its sum is equal to t. By Equation 3, we obtain that the family $\{s^n \mid n \in \mathbb{N}\}$ is summable with a sum equal to t. □

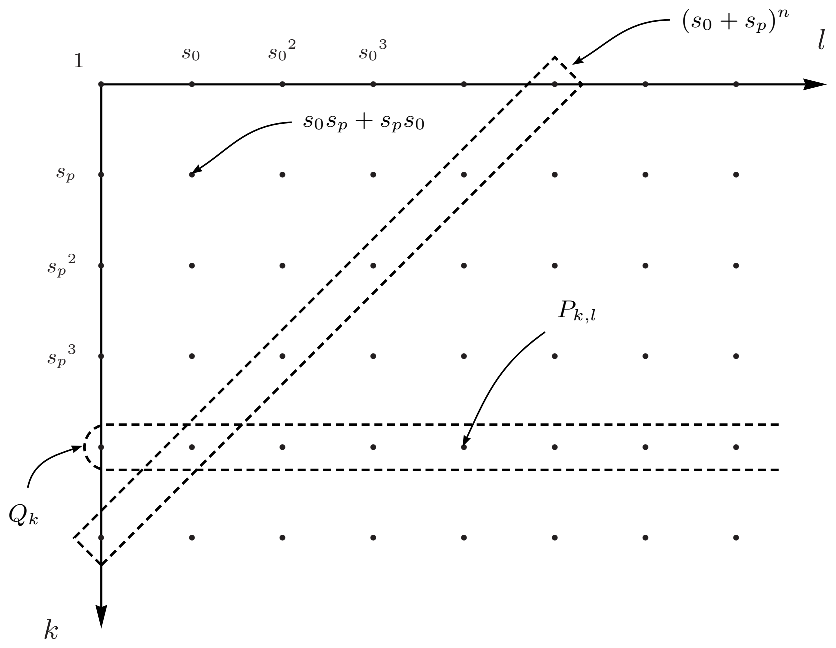


Fig. 1. The decomposition of the star of a series

2 Rational expressions with multiplicity

The definition of *rational expressions over A^* with multiplicity in* $\mathbb{K}$ goes as the one of classical rational expressions: it amounts to the construction of a set of well-formed formulae. Such an expression is called a $\mathbb{K}$-expression.

There are indeed two ways to deal with scalars in these formulae. A first option is to consider that they are atomic formulae, like letters. In this case, there are two binary operation, $+$ and $\cdot$ and one unary operation, $*$. Rational expressions are then given by the following grammar:

$$\mathsf{E} \to a \in A \mid k \in \mathbb{K} \mid (\mathsf{E} + \mathsf{E}) \mid (\mathsf{E} \cdot \mathsf{E}) \mid (\mathsf{E}^*) \tag{5}$$

This is for instance the way they are defined by J.J. Rutten in [17]. This method gives to scalars the same status as to letters. It is convenient as it gives a compact definition of expressions (only one kind of multiplication, the series zero is simply described by the scalar 0,...).

But this symmetry is not very convenient for our purpose, since letters and scalars do not play the same role in our algorithm. We prefer considering that every scalar can be applied on the left or the rigth side of an expression. Thus, to be able to represent the series zero or any constant series, we need two symbols which are constant. In this framework, atomic formulae are the letters and both symbols $\mathsf{0}$ and $\mathsf{1}$, the binary operations are still $+$ and $\cdot$, but the unary operations are $*$ and for every k in $\mathbb{K}$, the multiplication by k to the left or the right of an expression. Rational expressions are then given by the following grammar:

$$\mathsf{E} \to a \in A \mid \mathsf{0} \mid \mathsf{1} \mid (\mathsf{E} + \mathsf{E}) \mid (\mathsf{E} \cdot \mathsf{E}) \mid (\mathsf{E}^*) \mid (k\mathsf{E})\ k \in \mathbb{K} \mid (\mathsf{E}k)\ k \in \mathbb{K} \tag{6}$$

In the sequel, we use the latter definition to emphasize the special status of scalar in derivatives. We denote by $\mathbb{K}\,\mathsf{RatE}\,A^*$ the set of rational expressions over A with multiplicity in $\mathbb{K}$.

The complexity of a rational expression can be mesured by different parameters. The *litteral length*, denoted by $\ell(\mathsf{E})$, is the number of atomic formulae in E that are letters. Remark that $\ell((\mathsf{E}+\mathsf{F})) = \ell((\mathsf{E}\cdot\mathsf{F})) = \ell(\mathsf{E}) + \ell(\mathsf{F})$ and that $\ell((\mathsf{E}^*)) = \ell((k\,\mathsf{E})) = \ell((\mathsf{E})\,k) = \ell(\mathsf{E})$. This is the parameter on which the automaton we build in section 5 depends. Meanwhile, many of our proofs are actually by induction on the depth of a rational expression. The *depth*[1] $\mathsf{d}(\mathsf{E})$ of an expression E is inductively defined by:

$$\begin{gathered}
\mathsf{d}(\mathsf{0}) = \mathsf{d}(\mathsf{1}) = 0, \qquad\qquad \forall a \in A \qquad \mathsf{d}(a) = 0,\\
\mathsf{d}((\mathsf{E}^*)) = \mathsf{d}((k\,\mathsf{E})) = \mathsf{d}((\mathsf{E}\,k)) = \mathsf{d}(\mathsf{E}) + 1,\\
\mathsf{d}((\mathsf{E}\cdot\mathsf{F})) = \mathsf{d}((\mathsf{E}+\mathsf{F})) = \max(\mathsf{d}(\mathsf{E}), \mathsf{d}(\mathsf{F})) + 1.
\end{gathered}$$

As we said, a rational $\mathbb{K}$-expression is a formula, and the purpose of such a formula is to describe series, which is done *via* the following definition. The series *denoted* by an expression E, which we note by $|\mathsf{E}|$, is defined by induction on the depth $\mathsf{d}(\mathsf{E})$ of E.

$$\begin{gathered}
|\mathsf{0}| = 0_{\mathbb{K}}\,, \quad |\mathsf{1}| = 1_{A^*}\,, \quad |a| = a\,, \quad \text{for every } a \text{ in } A\,,\\
|(k\,\mathsf{E})| = k\,|\mathsf{E}|\,, \quad |(\mathsf{E}\,k)| = |\mathsf{E}|\,k\,,\\
|(\mathsf{E}+\mathsf{F})| = |\mathsf{E}| \oplus |\mathsf{F}|\,, \quad |(\mathsf{E}\cdot\mathsf{F})| = |\mathsf{E}|\,|\mathsf{F}|\,, \quad \text{and} \quad |(\mathsf{E}^*)| = |\mathsf{E}|^*.
\end{gathered}$$

The star operations involved in this computation may or may not be defined. If they are all defined, E is said to be valid and $|\mathsf{E}|$ is well-defined. Whether a $\mathbb{K}$-expression E is valid or not can be computed on the formulae E itself provided one can compute in $\mathbb{K}$, *via* the definition of the *constant term* of an expression E, which we denote by $\mathsf{c}(\mathsf{E})$.

$$\begin{gathered}
\mathsf{c}(\mathsf{0}) = 0_{\mathbb{K}}\,, \quad \mathsf{c}(\mathsf{1}) = 1_{\mathbb{K}}\,, \quad \forall a \in A \qquad \mathsf{c}(a) = 0_{\mathbb{K}}\,,\\
\mathsf{c}((k\,\mathsf{E})) = k\,\mathsf{c}(\mathsf{E})\,, \quad \mathsf{c}((\mathsf{E}\,k)) = \mathsf{c}(\mathsf{E})\,k\,,\\
\mathsf{c}((\mathsf{E}+\mathsf{F})) = \mathsf{c}(\mathsf{E}) \oplus \mathsf{c}(\mathsf{F})\,, \quad \mathsf{c}((\mathsf{E}\cdot\mathsf{F})) = \mathsf{c}(\mathsf{E})\,\mathsf{c}(\mathsf{F})\,,\\
\text{and} \quad \mathsf{c}((\mathsf{E}^*)) = \mathsf{c}(\mathsf{E})^* \quad \text{iff the latter is defined in } \mathbb{K}.
\end{gathered}$$

[1] We choose “depth ” rather than “height ” in order to avoid any confusion with the “star height ” of a rational expression (which we are dealing with in other papers of ours).

The definitions of the constant term of a series and of a $\mathbb{K}$-expression are consistent as the following statement holds:

Lemma 2. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$*, then* $\mathsf{c}(\mathsf{E})$ *is defined if, and only if,* E *is valid and in this case we have* $\mathsf{c}(\mathsf{E}) = \mathsf{c}(|\mathsf{E}|)$.

Proof. By induction on the depth of E, as shown by the following equations:

$$\mathsf{c}(0) = 0_{\mathbb{K}} = \mathsf{c}(0_{\mathbb{K}}) \qquad \mathsf{c}(1) = 1_{\mathbb{K}} = \mathsf{c}(1_{A^*}) \qquad \mathsf{c}(a) = 0_{\mathbb{K}} = \mathsf{c}(|a|)$$

$$\begin{aligned}
\mathsf{c}((k\,\mathsf{E})) &= k\,\mathsf{c}(\mathsf{E}) = k\,\mathsf{c}(|\mathsf{E}|) = \mathsf{c}(k\,|\mathsf{E}|) = \mathsf{c}(|(k\,\mathsf{E})|)\\
\mathsf{c}((\mathsf{E}\,k)) &= \mathsf{c}(\mathsf{E})\,k = \mathsf{c}(|\mathsf{E}|)\,k = \mathsf{c}(|\mathsf{E}|\,k) = \mathsf{c}(|(\mathsf{E}\,k)|)\\
\mathsf{c}((\mathsf{E}+\mathsf{F})) &= \mathsf{c}(\mathsf{E}) \oplus \mathsf{c}(\mathsf{F}) = \mathsf{c}(|\mathsf{E}|) \oplus \mathsf{c}(|\mathsf{F}|) = \mathsf{c}(|\mathsf{E}| \oplus |\mathsf{F}|) = \mathsf{c}(|(\mathsf{E}+\mathsf{F})|)\\
\mathsf{c}((\mathsf{E}\cdot\mathsf{F})) &= \mathsf{c}(\mathsf{E})\,\mathsf{c}(\mathsf{F}) = \mathsf{c}(|\mathsf{E}|)\,\mathsf{c}(|\mathsf{F}|) = \mathsf{c}(|\mathsf{E}|\,|\mathsf{F}|) = \mathsf{c}(|(\mathsf{E}\cdot\mathsf{F})|)\\
\mathsf{c}((\mathsf{E}^*)) &= \mathsf{c}(\mathsf{E})^* = \mathsf{c}(|\mathsf{E}|)^* = \mathsf{c}(|\mathsf{E}|^*) = \mathsf{c}(|(\mathsf{E}^*)|)
\end{aligned}$$

The last equation is made consistent by Lemma 1: $|(\mathsf{E}^*)|$ is defined iff $\mathsf{c}(\mathsf{E})^*$ is defined thus iff (E^*) is valid. □

Lemma 1 allows thus to decide whether the star of a series denoted by an expression E is defined by considering only the constant term of the expression E, and this will be used in the proof of Proposition 4.

Example 1. In this example, $A = \{a, b\}$ and $\mathbb{K} = \mathbb{Q}$.
$\mathsf{E}_1 = (((\frac{1}{6}\,a^*)+(\frac{1}{3}\,b^*))^*)$. Let $\mathsf{F}_1 = ((\frac{1}{6}\,a^*)+(\frac{1}{3}\,b^*))$. It comes $\mathsf{c}(\mathsf{F}_1) = \frac{1}{2}$, hence $\mathsf{c}(\mathsf{F}_1)^* = 2$. Thus, although $|\mathsf{F}_1|$ is not proper, the series denoted by E_1 is well-defined.

$\mathbb{K}$-rational series and $\mathbb{K}$-rational expressions are related by the well-known proposition:

Proposition 2. *A series of* $\mathbb{K}\langle\langle A^*\rangle\rangle$ *is* $\mathbb{K}$*-rational iff it is denoted by a valid rational* $\mathbb{K}$*-expression.* □

Trivial identities By definition, two $\mathbb{K}$-expressions are equivalent if and only if the series that they denote are equal, and two equivalent $\mathbb{K}$-expressions give rise to a $\mathbb{K}$-identity. The description of all $\mathbb{K}$-identities is known to be one of the most difficult problem in the theory and it will not be touched here.

We just select a set of such $\mathbb{K}$-identities that are used in the sequel in order to prove the correctness of our propositions and algorithms; they are very simple indeed and we thus call them *trivial identities*:

$$(k\,0) \equiv (0\,k) \equiv 0\,, \quad (0_{\mathbb{K}}\,\mathsf{E}) \equiv (\mathsf{E}\,0_{\mathbb{K}}) \equiv 0\,, \quad (0\cdot\mathsf{E}) \equiv (\mathsf{E}\cdot 0) \equiv 0\,, \tag{7}$$

$$0+\mathsf{E} \equiv \mathsf{E}+0 \equiv \mathsf{E}\,, \quad (1_{\mathbb{K}}\,\mathsf{E}) \equiv (\mathsf{E}\,1_{\mathbb{K}}) \equiv \mathsf{E}\,, \quad (1\cdot\mathsf{E}) \equiv (\mathsf{E}\cdot 1) \equiv \mathsf{E}\,, \tag{8}$$

$$(1\,k) \equiv (k\,1)\,, \qquad ((k\,1)\cdot\mathsf{E}) \equiv (k\,\mathsf{E})\,, \quad (\mathsf{E}\cdot(k\,1)) \equiv (\mathsf{E}\,k)\,, \tag{8'}$$

These identities are obviously consistent with the interpretation of rational expressions. Identity (7) states that zero is an absorbing element. Identity (8) and (8') reflects that zero and one are identities for addition and multiplication respectively.

One can consider the trivial identities as rewriting rules (that consist in replacing every sub-expression that have the same form as a left term of these equalities with the corresponding right term), and it should be clear that this leads to a "reduced form" for every rational expression which is unique and which can be computed in a time proportional to the (total) length of the expression.

Remark 1. Let us stress that we shall *not consider* in the sequel any other $\mathbb{K}$-identities, like those deriving from the associativity of $+$ and "$\cdot$", the distibutivity of "$\cdot$" over $+$ or the commutativity of $+$. For instance, $(a + (b + c))$, $((a + b) + c)$ and $(a + (c + b))$ are three different expressions and will be treated as such.

Remark 2. If $\mathbb{K}$ is a commutative semiring, the grammar describing rational expressions takes the simplified form:

$$\mathsf{E} \to a \in A \mid \mathsf{0} \mid \mathsf{1} \mid (\mathsf{E} + \mathsf{E}) \mid (\mathsf{E} \cdot \mathsf{E}) \mid (\mathsf{E}^*) \mid (k\mathsf{E}) \; k \in \mathbb{K}$$

and the equations corresponding to the right multiplication become useless.

3 Derivatives

We now introduce polynomials of expressions and their derivatives. We denote by:

$$\mathbb{K}\langle \mathbb{K} \,\mathsf{RatE}\, A^* \rangle \,,$$

the set of left linear combinations of rational expressions, or *polynomials of expressions*. It is a left $\mathbb{K}$-module; the addition is commutative and the multiplication by an element of $\mathbb{K}$ is distributive:

$$k\,\mathsf{E} \oplus k'\,\mathsf{F} = k'\,\mathsf{F} \oplus k\,\mathsf{E}, \qquad k\,\mathsf{E} \oplus k'\,\mathsf{E} = [k \oplus k']\mathsf{E}. \tag{9}$$

In the following, $[k\,\mathsf{E}]$ or $k\,\mathsf{E}$ is a monomial whereas $(k\,\mathsf{E})$ is an expression. The series denoted by a polynomial of rational expressions is obtained by extending by linearity the interpretation defined on rational expressions.

As it is the case in general for modules, there is no multiplication defined on $\mathbb{K}\langle \mathbb{K} \,\mathsf{RatE}\, A^* \rangle$. We need however to define an external right multiplication of an element of $\mathbb{K}\langle \mathbb{K} \,\mathsf{RatE}\, A^* \rangle$ by an expression and by a scalar. This operation is first defined on monomials and then extended to polynomials by linearity.

$$([k\,\mathsf{E}] \cdot \mathsf{F}) \equiv k\,(\mathsf{E} \cdot \mathsf{F}) \qquad ([k\,\mathsf{E}]\,k') \equiv k\,(\mathsf{E}\,k') \tag{10}$$

$$([\mathsf{E} \oplus \mathsf{E}'] \cdot \mathsf{F}) \equiv (\mathsf{E} \cdot \mathsf{F}) \oplus (\mathsf{E}' \cdot \mathsf{F}) \qquad ([\mathsf{E} \oplus \mathsf{E}']\; k) \equiv (\mathsf{E}\,k) \oplus (\mathsf{E}'\,k) \tag{11}$$

Remark 3. The set of polynomials of rational expressions is not a semialgebra. We do not try to defined such a multiplication because, first, we do not need it in the sequel and, second, this multiplication would not be associative anyway (because the multiplication of expressions is not).

Definition 2. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$ *and let* a *be in* A*. The* derivative *of* E *with respect to* a*, denoted by* $\dfrac{\partial}{\partial a}\,\mathsf{E}$*, is a polynomial of rational expressions with coefficients in* $\mathbb{K}$*, defined inductively by the following formulae.*

$$\frac{\partial}{\partial a}\,0 = \frac{\partial}{\partial a}\,1 = 0\,, \qquad \frac{\partial}{\partial a}\,b = \begin{cases} 1 & \textit{if} \quad b = a \\ 0 & \textit{otherwise} \end{cases} \tag{12}$$

$$\frac{\partial}{\partial a}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial a}\,\mathsf{E} \oplus \frac{\partial}{\partial a}\,\mathsf{F} \tag{13}$$

$$\frac{\partial}{\partial a}(k\,\mathsf{E}) = k\,\frac{\partial}{\partial a}\,\mathsf{E} \tag{14}$$

$$\frac{\partial}{\partial a}(\mathsf{E}\,k) = \left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\,k\right) \tag{15}$$

$$\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot\mathsf{F}\right) \oplus \mathsf{c}(\mathsf{E})\,\frac{\partial}{\partial a}\,\mathsf{F} \tag{16}$$

$$\frac{\partial}{\partial a}(\mathsf{E}^*) = \mathsf{c}(\mathsf{E})^*\,\left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) \tag{17}$$

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

$$\frac{\partial}{\partial a}\left(\bigoplus_{i\in I} k_i\,\mathsf{E}_i\right) = \bigoplus_{i\in I} k_i\,\frac{\partial}{\partial a}\,\mathsf{E}_i \tag{18}$$

Implicitely, the (polynomials of) expressions are reduced by trivial identities:

$$\frac{\partial}{\partial a}\,\mathsf{E} = 1 \Longrightarrow \frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \mathsf{F} \oplus \mathsf{c}(E)\,\frac{\partial}{\partial a}\,\mathsf{F}. \tag{19}$$

The difference between the equations (13) to (17) and the equations stated by Brzozowski is that "+" is replaced by $\oplus$ in the right handside of (13) and (16), and this realizes the generalization of the idea of Antimirov. Equations (14), (15), (17) and (18) are the natural ones that are necessary for the generalization to expressions with multiplicity. Notice that Equation (17) is defined only if (E^*) is a valid expression.

In contrast to the Boolean case, the number of polynomials obtained by iterating the derivation process can be infinite. Theorem 3 will state that all these different polynomials are linear combinations of a fixed finite number of expressions.

The derivative of an expression with respect to a *word* f is defined by induction on the length of f (by convention, the derivation with respect to the empty word is the identity):

$$\forall f \in A^*\,,\ \forall a \in A \qquad \frac{\partial}{\partial fa}\,\mathsf{E} = \frac{\partial}{\partial a}\left(\frac{\partial}{\partial f}\,\mathsf{E}\right)\,. \tag{20}$$

The following composition lemma then holds.

Lemma 3. $\forall \mathsf{E} \in \mathbb{K}\,\mathsf{RatE}\,A^*\,,\ \forall f, g \in A^* \qquad \dfrac{\partial}{\partial fg}\,\mathsf{E} = \dfrac{\partial}{\partial g}\left(\dfrac{\partial}{\partial f}\,\mathsf{E}\right)\,.$

Proof. The proof is by induction on the length of g. If g is a letter, the equality is the same as (20). If the equality holds for a words g in A^*, then, for every letter a in A:

$$\frac{\partial}{\partial fga}\,\mathsf{E} = \frac{\partial}{\partial a}\left(\frac{\partial}{\partial fg}\,\mathsf{E}\right) = \frac{\partial}{\partial a}\left(\frac{\partial}{\partial g}\left(\frac{\partial}{\partial f}\,\mathsf{E}\right)\right) = \frac{\partial}{\partial ga}\left(\frac{\partial}{\partial f}\,\mathsf{E}\right)\,. \qquad \square$$

There are explicit formulae that give the derivatives with respect to a word:

Proposition 3. $\forall \mathsf{E}, \mathsf{F} \in \mathbb{K}\,\mathsf{RatE}\,A^*,\ \forall f, g \in A^+,$

i) $\dfrac{\partial}{\partial f}(k\,\mathsf{E}) = k\ \dfrac{\partial}{\partial f}\,\mathsf{E}, \qquad \dfrac{\partial}{\partial f}(\mathsf{E}\,k) = \left(\left[\dfrac{\partial}{\partial f}\,\mathsf{E}\right]\ k\right),$

ii) $\dfrac{\partial}{\partial f}(\mathsf{E}+\mathsf{F}) = \dfrac{\partial}{\partial f}\,\mathsf{E} \oplus \dfrac{\partial}{\partial f}\,\mathsf{F}$

iii) $\dfrac{\partial}{\partial f}(\mathsf{E}\cdot\mathsf{F}) = \left[\dfrac{\partial}{\partial f}\,\mathsf{E}\right]\cdot\mathsf{F} \oplus \left[\ \bigoplus_{\substack{f=g\,h \\ g\in A^*,\,h\in A^+}} \mathsf{c}(\dfrac{\partial}{\partial g}\,\mathsf{E})\ \dfrac{\partial}{\partial h}\,\mathsf{F}\right]$

iv) $\dfrac{\partial}{\partial f}(\mathsf{E}^*) = \bigoplus_{\substack{f=g_1g_2\cdots g_n \\ g_1,g_2,\ldots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\ \prod_{i=1}^{n-1}\left[\mathsf{c}(\dfrac{\partial}{\partial g_i}\,\mathsf{E})\,\mathsf{c}(\mathsf{E})^*\right]\ \left[\dfrac{\partial}{\partial g_n}\,\mathsf{E}\right]\cdot(\mathsf{E}^*)$

Proof. The proof is by induction on the length of f. We denote:

$$\frac{\partial}{\partial a}\,\mathsf{E} = \bigoplus k_i\,\mathsf{E}_i \qquad \text{and} \qquad \frac{\partial}{\partial a}\,\mathsf{F} = \bigoplus k_j\,\mathsf{F}_j.$$

The equations of the proposition hold for letters. If these equations hold for a word f, then, for every letter a in A:

$$\frac{\partial}{\partial a}\,\mathsf{E} = \bigoplus k_i\,\mathsf{E}_i \qquad \text{and} \qquad \frac{\partial}{\partial a}\,\mathsf{F} = \bigoplus k_j\,\mathsf{F}_j.$$

And it then holds:

i)

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

$$\begin{aligned}\frac{\partial}{\partial af}(\mathsf{E}\,k) = \frac{\partial}{\partial f}\,\frac{\partial}{\partial a}(\mathsf{E}\,k) = \frac{\partial}{\partial f}[\bigoplus k_i\,(\mathsf{E}_i\,k)] = \bigoplus k_i\ \frac{\partial}{\partial f}(\mathsf{E}_i\,k) &= \bigoplus k_i([\frac{\partial}{\partial f}\,\mathsf{E}_i]\,k) \\ = \left([\bigoplus k_i\ \frac{\partial}{\partial f}\,\mathsf{E}_i]\,k\right) &= \left(\frac{\partial}{\partial f}[\bigoplus k_i\mathsf{E}_i]\,k\right) \\ = \left(\left[\frac{\partial}{\partial f}[\frac{\partial}{\partial a}\,\mathsf{E}]\right]\ k\right) &= \left([\frac{\partial}{\partial af}\,\mathsf{E}]\,k\right).\end{aligned}$$

ii)

$$\frac{\partial}{\partial af}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial f}[\frac{\partial}{\partial a}\,\mathsf{E} \oplus \frac{\partial}{\partial a}\,\mathsf{F}] = \frac{\partial}{\partial f}[\frac{\partial}{\partial a}\,\mathsf{E}] \oplus \frac{\partial}{\partial f}[\frac{\partial}{\partial a}\,\mathsf{F}] = \frac{\partial}{\partial af}\,\mathsf{E} \oplus \frac{\partial}{\partial af}\,\mathsf{F}.$$

iii)

$$\begin{aligned}\frac{\partial}{\partial af}(\mathsf{E}\cdot\mathsf{F}) &= \frac{\partial}{\partial f}\left[\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F})\right] = \frac{\partial}{\partial f}\left[\left([\frac{\partial}{\partial a}\,\mathsf{E}]\cdot\mathsf{F}\right) \oplus \mathsf{c}(\mathsf{E})\ \frac{\partial}{\partial a}\,\mathsf{F}\right] \\ &= \left(\frac{\partial}{\partial f}[\frac{\partial}{\partial a}\,\mathsf{E}]\cdot\mathsf{F}\right) \oplus \left[\ \bigoplus_{\substack{f=g\,h \\ g\in A^*,\,h\in A^+}} \mathsf{c}(\frac{\partial}{\partial g}[\frac{\partial}{\partial a}\,\mathsf{E}])\ \frac{\partial}{\partial h}\,\mathsf{F}\right] \oplus \mathsf{c}(\mathsf{E})\ \frac{\partial}{\partial af}\,\mathsf{F} \\ &= \left([\frac{\partial}{\partial af}\,\mathsf{E}]\cdot\mathsf{F}\right) \oplus \left[\ \bigoplus_{\substack{af=g\,h \\ g\in A^*,\,h\in A^+}} \mathsf{c}(\frac{\partial}{\partial g}\,\mathsf{E})\ \frac{\partial}{\partial h}\,\mathsf{F}\right]\end{aligned}$$

iv)

$$\begin{aligned}
\frac{\partial}{\partial af}(\mathsf{E}^*) &= \mathsf{c}(\mathsf{E})^* \ \frac{\partial}{\partial f}\left([\frac{\partial}{\partial a}\,\mathsf{E}]\cdot(\mathsf{E}^*)\right)\\
&= \mathsf{c}(\mathsf{E})^* \left[\left([\frac{\partial}{\partial af}\,\mathsf{E}]\cdot(\mathsf{E}^*)\right) \oplus \bigoplus_{\substack{af=g\,h\\ g\in A^*,\,h\in A^+}} [\mathsf{c}(\frac{\partial}{\partial ag}\,\mathsf{E})\ \frac{\partial}{\partial h}(\mathsf{E}^*)]\right]\\
&= \mathsf{c}(\mathsf{E})^* \left([\frac{\partial}{\partial af}\,\mathsf{E}]\cdot(\mathsf{E}^*)\right) \oplus\\
&\quad \bigoplus_{\substack{af=g_0\,h\\ g_0\in A^*,\,h\in A^+}} \ \bigoplus_{\substack{h=g_1\ldots g_n\\ g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\mathsf{c}(\frac{\partial}{\partial ag_0}\,\mathsf{E})\,\mathsf{c}(\mathsf{E})^*\mathsf{c}(\frac{\partial}{\partial g_1}\,\mathsf{E})\ldots\mathsf{c}(\mathsf{E})^*\left([\frac{\partial}{\partial g_n}\,\mathsf{E}]\cdot(\mathsf{E}^*)\right)\\
&= \bigoplus_{\substack{af=g_1\ldots g_n\\ g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_1}\,\mathsf{E})\ldots\mathsf{c}(\frac{\partial}{\partial g_{n\text{-}1}}\,\mathsf{E})\,\mathsf{c}(\mathsf{E})^*\left([\frac{\partial}{\partial g_n}\,\mathsf{E}]\cdot(\mathsf{E}^*)\right) \qquad \square
\end{aligned}$$

Example 2. (Ex. 1 continued)

$$\begin{aligned}
\frac{\partial}{\partial a}\,\mathsf{E}_1 &= \frac{\partial}{\partial a}\,(\mathsf{F_1}^*) = 2\,\frac{\partial}{\partial a}\left(\tfrac{1}{6}\,a^*\right)\cdot\mathsf{F_1}^* \oplus 2\,\frac{\partial}{\partial a}\left(\tfrac{1}{3}\,b^*\right)\cdot\mathsf{F_1}^* = \tfrac{1}{3}\,(a^*\cdot\mathsf{F_1}^*)\\
\frac{\partial}{\partial b}\,\mathsf{E}_1 &= 2\,\frac{\partial}{\partial b}\left(\tfrac{1}{3}\,b^*\right)\cdot\mathsf{F_1}^* = \tfrac{2}{3}\,(b^*\cdot\mathsf{F_1}^*)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial a}\,(a^*\cdot\mathsf{F_1}^*) &= \left(\frac{\partial}{\partial a}\,a^*\right)\cdot\mathsf{F_1}^* \oplus \mathsf{c}(a^*)\ \frac{\partial}{\partial a}\,(\mathsf{F_1}^*)\\
&= (a^*\cdot\mathsf{F_1}^*) \oplus \tfrac{1}{3}\,(a^*\cdot\mathsf{F_1}^*) = \tfrac{4}{3}\,(a^*\cdot\mathsf{F_1}^*)\\
\frac{\partial}{\partial b}\,(a^*\cdot\mathsf{F_1}^*) &= \left(\frac{\partial}{\partial b}\,a^*\right)\cdot\mathsf{F_1}^* \oplus \mathsf{c}(a^*)\ \frac{\partial}{\partial b}\,(\mathsf{F_1}^*) = \tfrac{2}{3}\,(b^*\cdot\mathsf{F_1}^*)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial a}\,(b^*\cdot\mathsf{F_1}^*) &= \left(\frac{\partial}{\partial a}\,b^*\right)\cdot\mathsf{F_1}^* \oplus \mathsf{c}(b^*)\ \frac{\partial}{\partial a}\,(\mathsf{F_1}^*) = \tfrac{1}{3}\,(a^*\cdot\mathsf{F_1}^*)\\
\frac{\partial}{\partial b}\,(b^*\cdot\mathsf{F_1}^*) &= \left(\frac{\partial}{\partial b}\,b^*\right)\cdot\mathsf{F_1}^* \oplus \mathsf{c}(b^*)\ \frac{\partial}{\partial b}\,(\mathsf{F_1}^*)\\
&= (b^*\cdot\mathsf{F_1}^*) \oplus \tfrac{2}{3}\,(b^*\cdot\mathsf{F_1}^*) = \tfrac{5}{3}\,(b^*\cdot\mathsf{F_1}^*)
\end{aligned}$$

Example 3. In this example, the multiplicity semiring is $\big(\mathcal{P}(\{\mathbf{x},\mathbf{y}\}^*),.,\cup\big)$.
Let $\mathsf{E}_2 = \mathbf{1} + a\cdot\big(\mathbf{x}a + b\cdot(\mathbf{xy}a + \mathbf{yx}b)\big)^*\cdot(\mathbf{x1} + \mathbf{xy}b)$.

$$\begin{aligned}
&\frac{\partial}{\partial b}\,\mathsf{E}_2 = \emptyset &&\frac{\partial}{\partial a}\,\mathsf{E}_2 = \big(\mathbf{x}\,a + b\cdot(\mathbf{xy}\,a + \mathbf{yx}\,b)\big)^*\cdot(\mathbf{x}\,\mathbf{1} + \mathbf{xy}\,b) = \mathsf{F}_2\\
&\frac{\partial}{\partial a}\,\mathsf{F}_2 = \mathbf{x}\,\mathsf{F}_2 &&\frac{\partial}{\partial b}\,\mathsf{F}_2 = (\mathbf{xy}\,a + \mathbf{yx}\,b)\cdot\mathsf{F}_2 \cup \mathbf{xy}\,\mathbf{1} = \mathsf{G}_2 \cup \mathbf{xy}\,\mathbf{1}\\
&\frac{\partial}{\partial a}\,\mathsf{G}_2 = \mathbf{xy}\,\mathsf{F}_2 &&\frac{\partial}{\partial b}\,\mathsf{G}_2 = \mathbf{yx}\,\mathsf{F}_2
\end{aligned}$$

$$\mathsf{c}(\mathsf{E}_2) = 1_{\{\mathbf{x},\mathbf{y}\}^*} \qquad \mathsf{c}(\mathsf{F}_2) = \mathbf{x}$$
$$\mathsf{c}(\mathsf{G}_2) = \emptyset \qquad \mathsf{c}(1) = 1_{\{\mathbf{x},\mathbf{y}\}^*}$$

Derivatives and quotients. The next thing we have to do is to establish the relationship between derivation and *(left) quotient*[2] Let us recall that if s is a series in $\mathbb{K}\langle\langle A^*\rangle\rangle$, the (left) quotient of s by a word f in A^* is the series $f^{-1}s$ defined by $\langle f^{-1}s, g\rangle = \langle s, f\,g\rangle$, for every g in A^*. The link between derivative and quotient is expressed in the following theorem.

Theorem 1. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$*. For every* f *in* A^+*, it holds:* $|\frac{\partial}{\partial f}\mathsf{E}| = f^{-1}|\mathsf{E}|$.

This theorem is a direct consequence of Lemma 3 and Proposition 4:

Proposition 4. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$*. For every* f *in* A^+*, it holds:* $\langle|\mathsf{E}|, f\rangle = \mathsf{c}(\frac{\partial}{\partial f}\mathsf{E})$.

Proof. The proof goes by induction on the depth of the expression and makes use of Proposition 3.

The result is true for $\mathsf{0}$ and $\mathsf{1}$: the derivation with respect to any word f in A^+ is null and the coefficient of f in the series 0 and in the characteristic series of 1_{A^*} is actually null.

The proof for $(k\,\mathsf{E})$, $(\mathsf{E}\,k)$, $(\mathsf{E}+\mathsf{F})$ and $(\mathsf{E}\cdot\mathsf{F})$ directly follows from the definition of $\mathsf{c}(\mathsf{E})$. In the proof for (E^*), we use Lemma 1 before applying the same arguments, in order to avoid an infinite sum:

$$\begin{aligned}
\langle|(\mathsf{E}^*)|, f\rangle &= \langle(\mathsf{c}(\mathsf{E})^*\,|\mathsf{E}|_p)^*\,\mathsf{c}(\mathsf{E})^*, f\rangle = \langle(\mathsf{c}(\mathsf{E})^*\,|\mathsf{E}|_p)^*, f\rangle\,\mathsf{c}(\mathsf{E})^* \\
&= \bigoplus_{\substack{f=g_1g_2\cdots g_n \\ g_1,g_2,\ldots,g_n\in A^+}} \langle\mathsf{c}(\mathsf{E})^*\,|\mathsf{E}|_p, g_1\rangle\ldots\langle\mathsf{c}(\mathsf{E})^*\,|\mathsf{E}|_p, g_n\rangle\,\mathsf{c}(\mathsf{E})^* \\
&= \bigoplus_{\substack{f=g_1g_2\cdots g_n \\ g_1,g_2,\ldots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\,\langle|\mathsf{E}|_p, g_1\rangle\ldots\mathsf{c}(\mathsf{E})^*\,\langle|\mathsf{E}|_p, g_n\rangle\,\mathsf{c}(\mathsf{E})^*
\end{aligned}$$

As the g_i are all *different* from 1_{A^*}, $\langle|\mathsf{E}|_p, g_i\rangle = \langle|\mathsf{E}|, g_i\rangle$.

$$\begin{aligned}
\langle|(\mathsf{E}^*)|, f\rangle &= \bigoplus_{\substack{f=g_1g_2\cdots g_n \\ g_1,g_2,\ldots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_1}\mathsf{E})\ldots\mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_n}\mathsf{E})\,\mathsf{c}(\mathsf{E})^* \\
&= \mathsf{c}\left(\bigoplus_{\substack{f=g_1g_2\cdots g_n \\ g_1,g_2,\ldots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_1}\mathsf{E})\ldots\mathsf{c}(\mathsf{E})^*\,\left[\frac{\partial}{\partial g_n}\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) = \mathsf{c}(\frac{\partial}{\partial f}\mathsf{E}^*) \qquad \square
\end{aligned}$$

Proof. (of Theorem 1) For every pair of words f and g in A^*,

$$\langle f^{-1}|\mathsf{E}|, g\rangle = \langle|\mathsf{E}|, fg\rangle = \mathsf{c}(\frac{\partial}{\partial fg}\mathsf{E}) = \mathsf{c}\left(\frac{\partial}{\partial g}\left[\frac{\partial}{\partial f}\mathsf{E}\right]\right) = \langle|\frac{\partial}{\partial f}\mathsf{E}|, g\rangle \qquad \square$$

Remark 4. If $\mathbb{K}$ is a commutative semiring, the definition of derivatives becomes simpler: Equation (15) and everything that concerns the right multiplication by a scalar becomes useless.

[2] Sometimes called *residual*.

4 Derived terms

We shall define the derived terms of a $\mathbb{K}$-expression E in a purely formal, and inductive, way. This could seem a bit heavy and clumsy for the computation of the successive derivatives of E would lead to the same result — at first sight. There is still a slight difference, which will be explained and used later in Section 7, and, therefore, we proceed as announced.

Definition 3. *The set* $D(\mathsf{E})$ *of* derived terms *of an expression* E *in* $\mathbb{K}\,\mathsf{RatE}\,A^*$ *is inductively defined by the following rules:*

$$D(\mathsf{0}) = D(\mathsf{1}) = \emptyset$$

$$\forall a \in A, \qquad D(a) = \{\mathsf{1}\}$$

$$\forall k \in \mathbb{K}, \forall \mathsf{E} \in \mathbb{K}\,\mathsf{RatE}\,A^*, \qquad D(k\,\mathsf{E}) = D(\mathsf{E}) \qquad D(\mathsf{E}\,k) = \bigcup_{\mathsf{K}\in D(\mathsf{E})} (\mathsf{K}\,k)$$

$$\forall \mathsf{E}, \mathsf{F} \in \mathbb{K}\,\mathsf{RatE}\,A^*, \qquad D(\mathsf{E}+\mathsf{F}) = D(\mathsf{E}) \cup D(\mathsf{F})$$

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

$$D(\mathsf{E}^*) = \bigcup_{\mathsf{K}\in D(\mathsf{E})} (\mathsf{K}\cdot(\mathsf{E}^*))$$

It can be noted that E itself does not belong necessarily to $D(\mathsf{E})$. We state now our main theorem, from which the generalization of Antimirov's result ensues.

Theorem 2. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$ *and let* $D(\mathsf{E})$ *be the set of derived terms of* E*. Then:*

(a) $D(\mathsf{E}) = \{\mathsf{K}_1, ...\mathsf{K}_n\}$ *is finite and* $n \leqslant \ell(\mathsf{E})$*;*

(b) there exist n *coefficients,* $k_1^{(a)}, k_2^{(a)}, \ldots, k_n^{(a)}$*, and* n^2 *coefficients* $\{z_{i,j}^{(a)}\}_{i,j\in[n]}$ *in* $\mathbb{K}$ *such that:*

$$\frac{\partial}{\partial a}\mathsf{E} = \bigoplus_{i\in[n]} k_i^{(a)}\,\mathsf{K}_i; \qquad \text{and} \qquad \forall i \in [n] \qquad \frac{\partial}{\partial a}\mathsf{K}_i = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\,\mathsf{K}_j. \tag{21}$$

In the case where $\mathbb{K} = \mathbb{B}$, these K_i are exactly what Antimirov called "*partial derivatives*" of E, with the explanation that they are "*parts*" of the (Brzozowski) *derivatives* of E (*cf.* [1]).

Example 4. (Ex. 1 continued) The derived terms of E_1 are $(a^*\cdot{\mathsf{F}_1}^*)$ and $(b^*\cdot{\mathsf{F}_1}^*)$.

Proof. Both Statements (a) and (b) are established by induction on the depth of the expression E (not on its *litteral* length). They obviously hold for $\mathsf{0}$ and $\mathsf{1}$ and for $\mathsf{E} = a$, $a \in A$. Statement (a) is easily deduced from the definition of the derived terms and from the definition of the litteral length. Statement (b) needs to be verified for each rational operation. Let thus E and F be two expressions for which (a) and (b) hold [we write $\frac{\partial}{\partial a}\mathsf{F} = \bigoplus_{p\in[s]} l_p^{(a)}\mathsf{L}_p$, for every a in A].

(i) Consider $(k\,\mathsf{E})$, $k \in \mathbb{K}$: the derived terms of $(k\,\mathsf{E})$ are the same as those of E.

(ii) Consider $(\mathsf{E}\,k)$, $k \in \mathbb{K}$. the derived terms of $(\mathsf{E}\,k)$ are the $(\mathsf{K}_i\,k)$, where K_i are the derived terms of E. This set is indeed closed under derivation.

(iii) Consider $(\mathsf{E}+\mathsf{F})$. It holds:

$$\frac{\partial}{\partial a}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial a}\mathsf{E} \oplus \frac{\partial}{\partial a}\mathsf{F} = \bigoplus_{i\in[n]} k_i^{(a)}\,\mathsf{K}_i \oplus \bigoplus_{p\in[s]} l_p^{(a)}\,\mathsf{L}_p.$$

The set of derived terms of $(\mathsf{E}+\mathsf{F})$ is the union of those of E and F and this set clearly satisfies the proposition.

(iv) Consider $(\mathsf{E}\cdot\mathsf{F})$. The set of derived terms of $(\mathsf{E}\cdot\mathsf{F})$ is the union of the sets $\{(\mathsf{K}_i\cdot\mathsf{F})\}_{i\in[n]}$ and $\{\mathsf{L}_p\}_{p\in[s]}$. It holds:

$$\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot\mathsf{F}\right)\oplus \mathsf{c}(\mathsf{E})\;\frac{\partial}{\partial a}\,\mathsf{F} = \bigoplus_{i\in[n]} k_i^{(a)}\,(\mathsf{K}_i\cdot\mathsf{F})\oplus\bigoplus_{p\in[s]}\left(\mathsf{c}(\mathsf{E})\,l_p^{(a)}\right)\,\mathsf{L}_p,$$

and , for every i in $[n]$ and every a in A,

$$\frac{\partial}{\partial a}(\mathsf{K}_i\cdot\mathsf{F}) = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\,(\mathsf{K}_j\cdot\mathsf{F})\oplus\bigoplus_{p\in[s]}\left(\mathsf{c}(\mathsf{K}_i)\,l_p^{(a)}\right)\,\mathsf{L}_p.$$

Hence, Statement (b) holds for $D(\mathsf{E}\cdot\mathsf{F})$.

(v) Consider (E^*) . The set of derived terms of (E^*) is $\{(\mathsf{K}_i\cdot\mathsf{E}^*)\}_{i\in[n]}$. It holds

$$\frac{\partial}{\partial a}(\mathsf{E}^*) = \mathsf{c}(\mathsf{E})^*\;\left(\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) = \bigoplus_{i\in[n]}\left(\mathsf{c}(\mathsf{E})^*\,k_i^{(a)}\right)\,(\mathsf{K}_i\cdot\mathsf{E}^*),$$

and, for every i in $[n]$ and every a in A,

$$\frac{\partial}{\partial a}(\mathsf{K}_i\cdot\mathsf{E}^*) = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\,(\mathsf{K}_j\cdot\mathsf{E}^*)\oplus\bigoplus_{j\in[n]}\left(\mathsf{c}(\mathsf{K}_i)\,\mathsf{c}(\mathsf{E})^*\,k_j^{(a)}\right)\,(\mathsf{K}_j\cdot\mathsf{E}^*).$$

Hence, Statement (b) holds for $D(\mathsf{E}^*)$. □

From Theorem 2 directly follows Theorem 3, that states that derivatives of a $\mathbb{K}$-expression belong to a finitely generated $\mathbb{K}$-submodule of $\mathbb{K}\langle\mathbb{K}\,\mathsf{RatE}\,A^*\rangle$.

Theorem 3. *Let E be in $\mathbb{K}\,\mathsf{RatE}\,A^*$. There exists an integer m, $m\leqslant\ell(\mathsf{E})$, and m rational expressions K_1, K_2, ..., K_m such that for every word f in A^+, there exist m coefficients in $\mathbb{K}$, $k_1^{(f)}$, $k_2^{(f)}$, ..., $k_m^{(f)}$, such that:*

$$\frac{\partial}{\partial f}\,\mathsf{E} = \bigoplus_{i\in[m]} k_i^{(f)}\,\mathsf{K}_i.$$

Proof. The proof is by induction on the length of words. If $|f|=1$, Theorem 2 i) gives the result. If $|f|>1$, we prove that the following equality holds between coefficients described in Theorem 2 and Theorem 3 :

$$\forall f\in A^+, \forall a\in A,\qquad k_j^{(fa)} = \bigoplus_{i\in[m]} k_i^{(f)}\,z_{i,j}^{(a)}\,. \tag{22}$$

Actually,

$$\begin{aligned}\frac{\partial}{\partial fa}\,\mathsf{E} &= \frac{\partial}{\partial a}\left[\frac{\partial}{\partial f}\,\mathsf{E}\right] = \frac{\partial}{\partial a}\left[\bigoplus_{i\in[m]} k_i^{(f)}\,\mathsf{K}_i\right] = \bigoplus_{i\in[m]} k_i^{(f)}\,\frac{\partial}{\partial a}\,\mathsf{K}_i \\ &= \bigoplus_{i\in[m]} k_i^{(f)}\left(\bigoplus_{j\in[m]} z_{i,j}^{(a)}\,\mathsf{K}_j\right) = \bigoplus_{j\in[m]}\left[\bigoplus_{i\in[m]} k_i^{(f)}\,z_{i,j}^{(a)}\right]\mathsf{K}_j\end{aligned}$$

□

Remark 5. The derivation and the left quotient are right actions of A^* on the set $\mathbb{K}\langle\mathbb{K}\,\mathsf{RatE}\,A^*\rangle$ of polynomials of rational expressions and on the set $\mathbb{K}\mathsf{Rat}A^*$ of rational series respectively. Theorem 3 says that the orbit of a rational expression with multiplicity under the action of A^* belongs to a finitely generated $\mathbb{K}$-module. The function which maps a polynomial of expressions P onto the rational power series $|P|$ is a morphism of actions. Therefore, Theorem 3 implies that the orbit of a rational series under the action of A^* belongs to a finitely generated $\mathbb{K}$-module as well, and provides a new proof for this classical result [2]. Derived terms give an explicit representation of a set of generators of this $\mathbb{K}$-module. An anonymous referee has drawn our attention to the fact that derived terms are a generalization of Mirkin's prebases (B.G. Mirkin: On dual automata, Kibernetika 2(1) (1966), 7–10).

5 The automaton of derived terms

In this paper, we do not consider the most general automata over A^* with multiplicity in $\mathbb{K}$ (as we do in [18]) but only a simple class of them. To emphasize the difference with classical (Boolean) automata, we still call them $\mathbb{K}$-automata.

Definition 4. *A $\mathbb{K}$-automaton over A is a 5-tuple $\mathcal{A} = \langle Q, A, E, I, T\rangle$, where Q is the finite set of states, I (*resp. *T) is a row (*resp. *column) vector of dimension Q with entries in $\mathbb{K}$ and E is a square matrix of size Q whose entries are linear combinations of letters in A with coefficients in $\mathbb{K}$.*

The behaviour of $\mathcal{A}$ is the series $|\mathcal{A}|$ such that, for every word f:

$$<|\mathcal{A}|, f> = <I \cdot E^* \cdot T, f> = <I \cdot E^{|f|} \cdot T, f> .$$

With the notations of this definition, one can define the $\mathbb{K}$-*representation* (of $|\mathcal{A}|$) as the triple (I, ζ, T), where ζ is a morphism from A^* to $\mathbb{K}^{Q\times Q}$ such that $E = \sum_{a\in A}(a)\zeta a$. This is a part of the proof of the so-called Kleene-Schützenberger Theorem (*cf.* [2]). It is clear that it holds:

$$|\mathcal{A}| = \bigoplus_{f\in A^*} (I \cdot (f)\zeta \cdot T)\, f.$$

To any rational $\mathbb{K}$-expression E in $\mathbb{K}\,\mathsf{RatE}\,A^*$, we associate a $\mathbb{K}$-automaton in the following way.

Definition 5. *Let $P_{\mathsf{E}} = D(\mathsf{E}) \cup \{\mathsf{E}\}$. The expression E may already belong to $D(\mathsf{E})$; if it does not, we set $\mathsf{K}_0 = \mathsf{E}$. The* automaton of derived terms of E *is the $\mathbb{K}$-automaton $\mathcal{A}_{\mathsf{E}} = \langle P_{\mathsf{E}}, A, Z, I, T\rangle$ defined by:*

$$I_{\mathsf{K}_i} = \begin{cases} 1_{\mathbb{K}} & \text{if } \ \mathsf{K}_i = \mathsf{E}\,, \\ 0_{\mathbb{K}} & \text{otherwise}\,, \end{cases} \qquad Z_{\mathsf{K}_i,\mathsf{K}_j} = \bigoplus_{a\in A} z_{i,j}^{(a)}\, a\,, \qquad T_{\mathsf{K}_j} = \mathsf{c}(\mathsf{K}_j),$$

where the $z_{i,j}^{(a)}$ have been defined at Theorem 2 and if E does not belong to $D(\mathsf{E})$, for every i, $z_{i,0}^{(a)} = 0_{\mathbb{K}}$, and, for every $j > 0$, $z_{0,j}^{(a)} = k_j^{(a)}$.

The vectors I and T are vectors of dimension P_{E} with coefficients in $\mathbb{K}$; this is the usual generalisation of initial and final states for $\mathbb{K}$-automata.

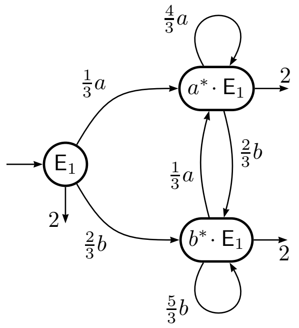
Theorem 4. *Let E be in $\mathbb{K}\,\mathsf{RatE}\,A^*$. The series realized by the $\mathbb{K}$-automaton of derived terms of E is equal to the series denoted by E: $|\mathcal{A}_{\mathsf{E}}| = |\mathsf{E}|$.*

Proof. Let $n = \mathsf{Card}(P_{\mathsf{E}})$ and (I, ζ, T) the $\mathbb{K}$-representation corresponding to $\mathcal{A}_{\mathsf{E}}$: $(a)\zeta_{i,j} = z_{i,j}^{(a)}$, for every a in A and every pair of integers i and j in $[n]$. Equation (22) directly shows, by induction on the length of f, that:

$$\forall f \in A^+ \,,\ \forall i \in [n] \qquad (I \cdot (f)\zeta)_i = k_i^{(f)}, \tag{23}$$

and then, by Proposition 4, $\langle|\mathcal{A}_{\mathsf{E}}|, f\rangle = \bigoplus_{i \in [n]} k_i^{(f)}\, \mathsf{c}(\mathsf{K}_i) = \langle|\mathsf{E}|, f\rangle$. for every f in A^+. The coefficient of the empty word is $\langle|\mathcal{A}_{\mathsf{E}}|, 1_{A^*}\rangle = \mathsf{c}(\mathsf{K}_i)$ for $\mathsf{K}_i = \mathsf{E}$, thus $\langle|\mathcal{A}_{\mathsf{E}}|, 1_{A^*}\rangle = \mathsf{c}(\mathsf{E})$ and the theorem is verified. □

Example 5. (Ex. 1 continued)
The opposite figure shows $\mathcal{A}_{\mathsf{E}_1}$, the automaton of derived terms of E_1.



Example 6. (Ex. 3 continued) We build $\mathcal{A}_{\mathsf{E}_2}$, the automaton of derived terms of E_2, with multiplicity in $\big(\mathcal{P}(\{\mathbf{x}, \mathbf{y}\}^*), ., \cup\big)$; this is nothing else than a transducer and thus we write the coefficients as outputs.[3]

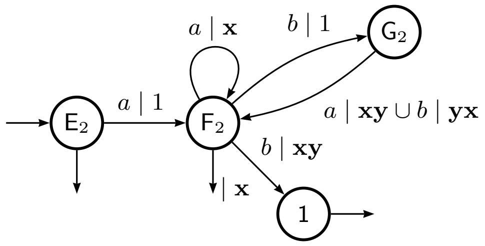


Fig. 2. The automaton of derived terms of E_2

Definition 5 and Theorem 4 are the aim of the definition of derivation for $\mathbb{K}$-expressions: giving an algorithm to turn a rational $\mathbb{K}$-expression into a $\mathbb{K}$-automaton.

The following sections will give some insight on this $\mathbb{K}$-automaton: the notion of shadow terms, a property (Theorem 6) and possible variations of the definition of derivation.

6 Shadow terms

Every term that appears in derivatives of a $\mathbb{K}$-rational expression with respect to a word is a derived term. (*i.e.* the rational expressions of Theorem 3 are derived terms). But it is not true that every derived term necessary appears in some derivative of the expression. Such a derived term is called a *shadow* derived term.

[3] This transducer realizes the replacement of a's by $\mathbf{x}$'s, of b's by $\mathbf{y}$'s and the "cautious" rewriting of factors $a\,b\,b$ by $\mathbf{y\,x\,x}$ in the reading from left to right of the words (*cf.* [18]).

Definition 6. *A* shadow *derived term of a $\mathbb{K}$-rational expression* E *is a derived term* K *such that, for every f in A^*,*

$$<\frac{\partial}{\partial f}\,\mathsf{E},\mathsf{K}> = 0_{\mathbb{K}}\,.$$

Proposition 5. *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$ *and let* $\mathcal{A}_{\mathsf{E}}$ *be its automaton of derived terms. The* $\mathbb{K}$*-automaton* $\mathcal{A}'_{\mathsf{E}}$ *obtained from* $\mathcal{A}_{\mathsf{E}}$ *by erasing the states that correspond to shadow derived terms and by trimming the result realizes the same series:* $|\mathcal{A}'_{\mathsf{E}}| = |\mathsf{E}|$.

Proof. If K_i is a shadow derived term, then, by (23), $k_i^{(f)} = (I \cdot (f)\zeta)_i = 0_{\mathbb{K}}$, for every f in A^*. This state K_i can be erased in the representation (I,ζ,T) without changing the realized series. Further, trimming the automaton does not change the realized series either. □

Example 7. Let $\mathsf{E}_3 = \big(a\,b\,a + (a\,(a - b\,a))\big)$. The derivatives of E_3 are:

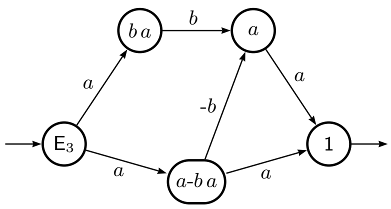
$$\frac{\partial}{\partial a}\,\mathsf{E}_3 = b\,a \oplus (a - b\,a) \qquad\qquad \frac{\partial}{\partial b}\,\mathsf{E}_3 = 0$$

$$\frac{\partial}{\partial b}\,ba = a \qquad\qquad \frac{\partial}{\partial b}(a - ba) = (-1_{\mathbb{K}})\,a$$

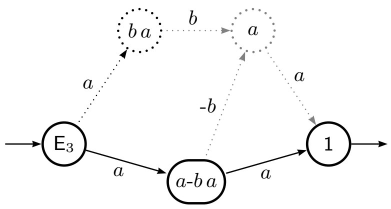
$$\frac{\partial}{\partial aa}\,\mathsf{E}_3 = \frac{\partial}{\partial a}\,b\,a \oplus \frac{\partial}{\partial a}(a - b\,a) = 1$$

$$\frac{\partial}{\partial ab}\,\mathsf{E}_3 = \frac{\partial}{\partial b}\,b\,a \oplus \frac{\partial}{\partial b}(a - b\,a) = a \oplus (-1_{\mathbb{K}})\,a = 0$$

The derived term a does not appear in a derivative $\dfrac{\partial}{\partial f}\,\mathsf{E}_3$ for any f in A^*: it is a shadow term. Figure 3 (a) below shows the automaton of derived terms of E_3; Figure 3 (b) shows the effect of, first, the suppression of the state a and then the trimming that erases the state $b\,a$.



(a) The automaton of derived terms of E_3



(b) The automaton of non-shadow derived terms of E_3

Fig. 3. Deletion of shadow terms

Deciding whether a derived term is shadow or not depends on the semiring of coefficients. In many cases —and in particular in all classical cases— this computation does not bring any problem. On the one hand, if the semiring is positive, there is no shadow term. For instance, Boolean semiring, subsemirings of $\mathbb{R}_+$, $(\max,+)$-semirings, $(\mathcal{P}(A^*),\cup,.)$, *etc.* are positive. On the other hand, K_i is a shadow term if and only if the series realized by the automaton $\mathcal{A}_i = \langle P_{\mathsf{E}}, A, Z, I, \{\mathsf{K}_i\}\rangle$ is equal to zero. This can be easily decided if the semiring of coefficients is a subsemiring of a field (*cf.* [10, 2], Equality theorem). Examples of such semirings are $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Z}[X]$, *etc.* If $\mathbb{K}$ is not of that kind, the complexity of the computation of shadow terms may be more difficult and, in particular, does not depend only on the number of states of the automaton but also on the coefficients.

Anyway, this computation is not really necessary as the automaton of derived terms is already a *finite* automaton that realizes the series denoted by an expression E and that can be effectively computed from E. This notion of shadow term is necessary to explain the discrepancy that may occur between the derived terms that appear in the derivatives of an expression and the ones that are given by the formulae of Definition 3 applied to the same expression.

7 The automaton of derived terms and the standard automaton

Champarnaud and Ziadi have shown that, in the Boolean case, the automaton of partial derivatives of Antimirov is obtained from the standard automaton as the result of some merging of states, that is:

Theorem 5. *[8] There is a morphism from the standard automaton of an expression onto the automaton of partial derivatives of this expression.*

In order to extend this result to $\mathbb{K}$-automata, we have to generalize both the definitions of morphisms — that we shall call $\mathbb{K}$-*coverings* — and the statement of this result.

Theorem 6. *There is a $\mathbb{K}$-covering from the standard automaton of a $\mathbb{K}$-expression onto the automaton of derived terms of this expression.*

This statement holds because we have defined the derived terms in the way we did at Section 5. We shall explain this phenomenon with more detail at Remark 7.

7.1 The standard $\mathbb{K}$-automaton of a $\mathbb{K}$-expression

A standard automaton is an automaton with a particular form. To every rational expression, one can associate a standard automaton. This automaton, often called the "Glushkov automaton" of the expression, is canonical. It can be built by considering that every state, except the initial state, corresponds to an occurrence of a letter in the expression. This is the reason why it is sometimes called the *position automaton*. Caron et Flouret [4] have generalized the definition of Glushkov automaton, and its construction by the position algorithm, to automata with multiplicity. Their construction is reproduced here, for sake of completeness and in a way which is more convenient to our purpose.

Definition 7. *A $\mathbb{K}$-automaton is* standard *if it has one and only one initial state with initial multiplicity equal to $1_{\mathbb{K}}$ and such that this state is not the end of any transition. Moreover, every state of the automaton is accessible from this initial state.*

Every automaton can easily be turned into a standard automaton:

Definition 8. *Let $\mathcal{A} = \langle Q, A, E, I, T\rangle$ be a $\mathbb{K}$-automaton. Let i be a state that do not belong to Q. The* standardized *automaton $\mathsf{s}(\mathcal{A})$ of $\mathcal{A}$ is the accessible part of the $\mathbb{K}$-automaton $\langle Q \cup \{i\}, A, E', I', T'\rangle$, where:*

$$I'_p = \begin{cases} 1_{\mathbb{K}} & \textit{if} \quad p = i \\ 0_{\mathbb{K}} & \textit{otherwise} \end{cases} \qquad\qquad T'_p = \begin{cases} \langle I, T\rangle & \textit{if}^4 \quad p = i \\ T_p & \textit{otherwise} \end{cases}$$

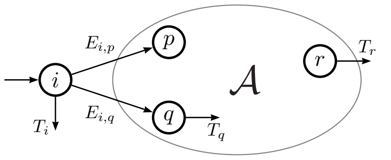
$$E'_{p,q} = \begin{cases} 0_{\mathbb{K}} & \textit{if} \quad q = i \\ \bigoplus_{r \in Q} I_r E_{r,q} & \textit{if} \quad p = i \\ E_{p,q} & \textit{otherwise} \end{cases}$$

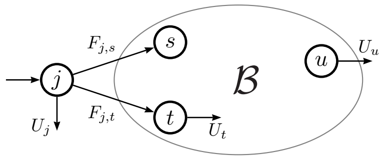
¿From these definitions it is easily seen that the standardized of a standard automaton is equal (isomorphic) to that automaton.

Remark 6. The constant term of $|\mathsf{s}(\mathcal{A})| = |\mathcal{A}|$ is precisely the final multiplicity of the initial state of $\mathsf{s}(\mathcal{A})$.

Rational operations on standard automata A classical way to prove the easy part of the Kleene theorem is to prove that rational operations on languages can be translated onto automata that accept them, either by using ε-transitions (Thompson construction) or normalized automata. Likewise, such operations can be defined on standard automata and this is what we shall describe now.

Let $\mathcal{A} = \langle Q, A, E, \{i\}, T\rangle$ and $\mathcal{B} = \langle R, A, F, \{j\}, U\rangle$ be two standard automata:

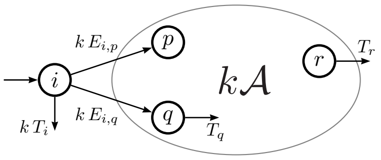




• The standard automata $k\mathcal{A} = \langle Q, A, E^{(k.)}, \{i\}, T^{(k.)}\rangle$ is defined by:

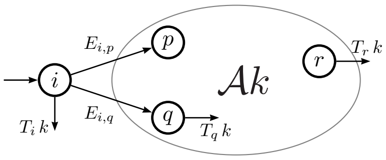
$$\forall p,q \in Q\,, \qquad E^{(k.)}_{p,q} = \begin{cases} k\,E_{p,q} & if \quad p = i \\ E_{p,q} & otherwise \end{cases}$$

$$\forall p \in Q\,, \qquad T^{(k.)}_{p} = \begin{cases} k\,T_{p} & if \quad p = i \\ T_{p} & otherwise \end{cases}$$



• The standard automata $\mathcal{A}k = \langle Q, A, E, \{i\}, T^{(.k)}\rangle$ is defined by:

$$\forall p \in Q\,, \qquad T^{(.k)}_{p} = T_p\,k$$



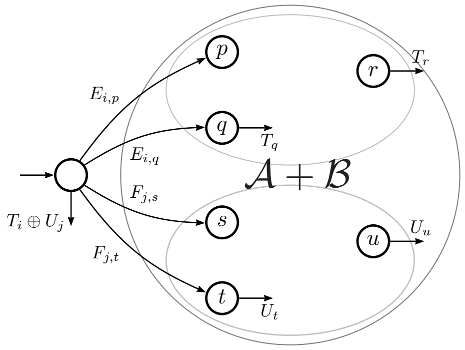
• The standard automaton $\mathcal{A} + \mathcal{B} = \langle Q \cup R \setminus \{j\}, A, G, \{i\}, V\rangle$ is defined as:

$\forall p,q \in Q \cup R \setminus \{j\}\,,$

$$G_{p,q} = \begin{cases} E_{p,q} & if \quad p,q \in Q \\ F_{p,q} & if \quad p,q \in R \\ F_{j,q} & if \quad p = i \text{ and } q \in R \\ 0_{\mathbb{K}} & otherwise \end{cases}$$

$\forall p \in Q \cup R \setminus \{j\}\,,$

$$V_{p} = \begin{cases} T_i \oplus U_j & if \quad p = i \\ T_{p} & if \quad p \in Q \setminus \{i\} \\ U_p & if \quad p \in R \end{cases}$$

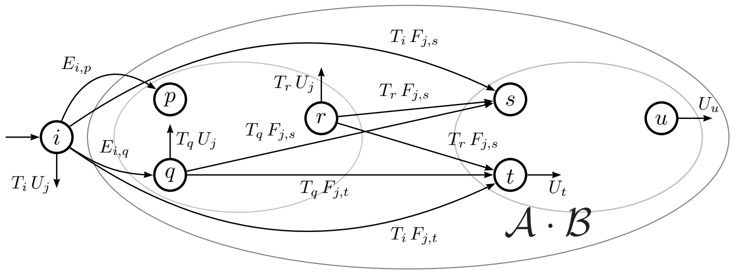


4 $\langle I, T\rangle$ is the scalar product of the vectors I and T.

• The standard automaton $\mathcal{A} \cdot \mathcal{B} = \langle Q \cup R \setminus \{j\}, A, G, \{i\}, V \rangle$ is:

$\forall p, q \in Q \cup R \setminus \{j\}$, $\qquad\qquad$ $\forall p \in Q \cup R \setminus \{j\}$,

$$G_{p,q} = \begin{cases} E_{p,q} & if \quad p, q \in Q \\ F_{p,q} & if \quad p, q \in R \\ T_p F_{j,q} & if \quad p \in Q \text{ and } q \in R \\ 0_{\mathbb{K}} & otherwise \end{cases} \qquad V_p = \begin{cases} U_p & if \quad p \in R \\ T_p\, U_j & if \quad p \in Q \end{cases}$$



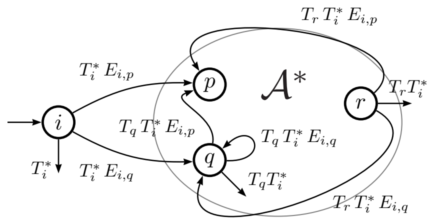
• The standard automaton $\mathcal{A}^* = \langle Q, A, E^{(*)}, \{i\}, T^{(*)} \rangle$ is:

$\forall p, q \in Q$,

$$E^{(*)}_{p,q} = \begin{cases} T_i^* E_{i,q} & if \quad p = i \\ T_p T_i^* E_{i,q} \oplus E_{p,q} & otherwise \end{cases}$$

$\forall p \in Q$,

$$T^{(*)}_p = \begin{cases} T_i^* & if \quad p = i \\ T_p T_i^* & otherwise \end{cases}$$



All these definitions have been taken in order that the following equations—that are easily verified—hold:

$$|(k\,\mathcal{A})| = k\,|\mathcal{A}|\,, \quad |(\mathcal{A}\,k)| = |\mathcal{A}|\,k\,,$$
$$|(\mathcal{A} + \mathcal{B})| = |\mathcal{A}| \oplus |\mathcal{B}|\,, \quad |(\mathcal{A} \cdot \mathcal{B})| = |\mathcal{A}|\,|\mathcal{B}|\,, \quad \text{and} \quad |\mathcal{A}^*| = |\mathcal{A}|^*.$$

Standard automaton of a rational expression

Definition 9. *For every rational $\mathbb{K}$-expression E, there exists a canonical standard $\mathbb{K}$-automaton $\mathcal{S}_{\mathsf{E}}$ that realizes the series denoted by the expression. It is defined by induction on the depth of the expression as follows:*

expression	0	1	$a \in A$	$(k\,\mathsf{E})$	$(\mathsf{E}\,k)$	$(\mathsf{E}+\mathsf{F})$	$(\mathsf{E}\cdot\mathsf{F})$	(E^*)
automaton	$\to\bigcirc$	$\to\bigcirc\to$	$\to\bigcirc\xrightarrow{a}\bigcirc\to$	$k\,\mathcal{S}_{\mathsf{E}}$	$\mathcal{S}_{\mathsf{E}}k$	$\mathcal{S}_{\mathsf{E}} + \mathcal{S}_{\mathsf{F}}$	$\mathcal{S}_{\mathsf{E}} \cdot \mathcal{S}_{\mathsf{F}}$	$\mathcal{S}_{\mathsf{E}}^*$

Theorem 7. *[4] The standard automaton of a rational $\mathbb{K}$-expression E has $\ell(\mathsf{E}) + 1$ states and realizes the series $|E|$.*

In particular, the last of the above equalities, together with Remark 6, is another proof of Lemma 2. It follows also that the number of states of the position automaton is greater than or equal to the number of states of the automaton of derived terms.

Note that not every standard $\mathbb{K}$-automaton is *the* standard (or *Glushkov*) $\mathbb{K}$-automaton of a $\mathbb{K}$-expression. It can be determined whether a given standard $\mathbb{K}$-automaton is the standard $\mathbb{K}$-automaton of a $\mathbb{K}$-expression [5, 6].

7.2 $\mathbb{K}$-coverings

A $\mathbb{K}$-covering between $\mathbb{K}$-automata is a relation that is more precise than the notion of morphisms between Boolean automata. $\mathbb{K}$-coverings have been defined in [18]. Intuitively, a $\mathbb{K}$-covering $\varphi : \mathcal{A} \to \mathcal{B}$ merges two states of $\mathcal{A}$ only if they have the "same" outgoing transitions *modulo* φ. This means that two states have the same image by φ only if the sums of their outgoing transitions to any class of states modulo φ are equal. More formally, $\mathbb{K}$-coverings are defined as follow:

Definition 10. *Let $\mathcal{A} = \langle I, E, T\rangle$ and $\mathcal{B} = \langle J, F, U\rangle$ be two $\mathbb{K}$-automata. A surjective mapping φ from states of $\mathcal{A}$ onto states of $\mathcal{B}$ induces a* $\mathbb{K}$-covering *from $\mathcal{A}$ onto $\mathcal{B}$ if $\mathcal{A}$ is such that:*

$$\forall p, p' \in Q, \qquad p\varphi = p'\varphi \Longrightarrow \begin{cases} i) & T_p = T_{p'} \\ ii)\ \forall q \in Q,\ \text{let } S_q = q\varphi\varphi^{-1} & \bigoplus_{q' \in S_q} E_{p,q'} = \bigoplus_{q' \in S_q} E_{p',q'} \end{cases}$$

and if $\mathcal{B}$ satisfies the following:

$$\begin{array}{llll} iii) & \forall r \in R, & U_r = T_p & \text{for any } p \text{ in } r\varphi^{-1} \\ iv) & \forall s \in R, & J_s = \bigoplus_{q \in s\varphi^{-1}} I_q & \\ v) & \forall (r,s) \in R^2, & F_{r,s} = \bigoplus_{q \in s\varphi^{-1}} E_{p,q} & \text{for any } p \text{ in } r\varphi^{-1} \end{array}$$

This definition is illustrated on Figure 4.

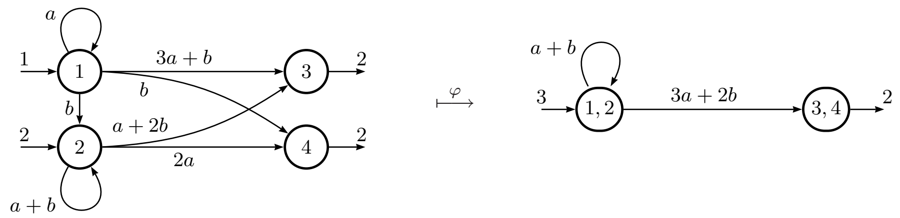


Fig. 4. A $\mathbb{N}$-covering

Proposition 6. *[18] Let $\mathcal{A}$ and $\mathcal{B}$ be two $\mathbb{K}$-automata. If $\varphi : \mathcal{A} \to \mathcal{B}$ is a $\mathbb{K}$-covering, then $|\mathcal{A}| = |\mathcal{B}|$.*

Proof. Let $\mathcal{A} = \langle Q, A, E, I, T\rangle$ and $\mathcal{B} = \langle R, A, F, J, U\rangle$. Let ζ (*resp.* η) be the $\mathbb{K}$-representation induced by the definition of $\mathcal{A}$ (*resp.* $\mathcal{B}$), as in the proof of Theorem 4. We prove, by induction on the length of words, that, for every word u, for every state r of $\mathcal{B}$,

$$(J \cdot (u)\eta)_r = \bigoplus_{p \in r\varphi^{-1}} (I \cdot (u)\zeta)_p \,. \tag{24}$$

If u is the empty word, $(u)\eta$ and $(u)\zeta$ are identities. Thus, Equation (24) amounts to Definition 10,ii). If u is a word such that (24) holds, then for every a in A,

$$\begin{aligned}
\big(J \cdot (ua)\eta\big)_s &= \bigoplus_{r \in R} \big(J \cdot (u)\eta\big)_r (a)\eta_{r,s} \\
&= \bigoplus_{r \in R} \left[\left(\bigoplus_{p \in r\varphi^{-1}} \big(I \cdot (u)\zeta\big)_p \right) (a)\eta_{r,s} \right] && \text{(induction)} \\
&= \bigoplus_{r \in R} \bigoplus_{p \in r\varphi^{-1}} \left[\big(I \cdot (u)\zeta\big)_p (a)\eta_{r,s} \right] \\
&= \bigoplus_{r \in R} \bigoplus_{p \in r\varphi^{-1}} \left[\big(I \cdot (u)\zeta\big)_p \bigoplus_{q \in s\varphi^{-1}} (a)\zeta_{p,q} \right] && \big(\text{Def. 10,v}\big) \\
&= \bigoplus_{p \in Q} \bigoplus_{q \in s\varphi^{-1}} \big(I \cdot (u)\zeta\big)_p (a)\zeta_{p,q} \\
&= \bigoplus_{q \in s\varphi^{-1}} \bigoplus_{p \in Q} \big(I \cdot (u)\zeta\big)_p (a)\zeta_{p,q} = \bigoplus_{q \in s\varphi^{-1}} \big(I \cdot (ua)\zeta\big)_q
\end{aligned}$$

Hence, (24) holds for every word of A^*, and, for every u in A^*,

$$\langle|\mathcal{B}|, u\rangle = \bigoplus_{r \in R} \big(J \cdot (u)\eta\big)_r T_r = \bigoplus_{r \in R} \bigoplus_{p \in r\varphi^{-1}} \big(I \cdot (u)\zeta\big)_p T_p = \bigoplus_{p \in Q} \big(I \cdot (u)\zeta\big)_p T_p = \langle|\mathcal{A}|, u\rangle \qquad \square$$

Remark 7. If $\mathbb{K}$ is the Boolean semiring, $\varphi : \mathcal{A} \to \mathcal{B}$ is a $\mathbb{B}$-covering if and only if it is a surjective morphism such that, for every state p of $\mathcal{A}$, it induces a surjective mapping of outgoing transitions of p onto outgoing transitions of $\varphi(p)$, and if $\varphi(p)$ is terminal, so is p. In this case, we say that $\mathcal{B}$ is a *quotient* of $\mathcal{A}$.

7.3 Proof of Theorem 6

Another example of the definition of $\mathbb{K}$-coverings is the following lemma which will be used in the sequel.

Lemma 4. *Let $\mathcal{A}$ be a $\mathbb{K}$-automaton and $\mathsf{s}(\mathcal{A})$ its standardized automaton. If $\mathcal{A}$ has one and only one initial state with initial multiplicity $1_{\mathbb{K}}$, the canonical mapping $\sigma_{\mathcal{A}} : \mathsf{s}(\mathcal{A}) \to \mathcal{A}$ is a $\mathbb{K}$-covering.*

Proof. Let p be the initial state of $\mathcal{A}$. With the notation of Definition 8, $\sigma_{\mathcal{A}}$ is a one-to-one mapping except for i and p: $p\sigma_{\mathcal{A}}^{-1} = \{p, i\}$. Let us check that σ is a $\mathbb{K}$-covering:

$$\forall q \in Q\,, \quad \begin{cases} I'_p = 0_{\mathbb{K}}, I'_i = 1_{\mathbb{K}} \Rightarrow I_q = 1_{\mathbb{K}} = I'_p \oplus I'_i & \text{if} \quad q = p \quad (q\sigma_{\mathcal{A}}^{-1} = \{p, i\}) \\ I'_q = 0_{\mathbb{K}} = I_q & \text{otherwise} \end{cases}$$

$$\forall q \in Q\,, \quad \begin{cases} T'_p = T_q \quad \text{and} \quad T'_i = I_q\, T_q = T_q & \text{if} \quad q = p \quad (q\sigma_{\mathcal{A}}^{-1} = \{p, i\}) \\ T'_q = T_q & \text{otherwise} \end{cases}$$

$$\forall q, r \in Q\,, \quad \begin{cases} \text{if} \quad q = p \quad \text{and} \quad r = p & \left.\begin{matrix} E'_{pp} = E'_{ip} = E_{qr} \\ E'_{pi} = E'_{ii} = 0_{\mathbb{K}} \end{matrix}\right\} \Rightarrow E_{qr} = E'_{pp} \oplus E'_{pi} = E'_{ip} \oplus E'_{ii} \\ \text{if} \quad q \neq p \quad \text{and} \quad r = p & E'_{qp} = E_{qr} \quad \text{and} \quad E'_{qi} = 0_{\mathbb{K}} \Rightarrow E_{qr} = E'_{qp} \oplus E'_{qi} \\ \text{if} \quad q = p \quad \text{and} \quad r \neq p & E'_{pr} = E_{pr} \quad \text{and} \quad E'_{ir} = E_{pr} \\ \text{otherwise} & E'_{qr} = E_{qr} \end{cases}$$

□

Remark 8. ¿From Definition 5, it follows that the derived term automaton $\mathcal{A}_{\mathsf{E}}$ of a $\mathbb{K}$-expression E meets the hypothesis of Lemma 4.

Lemma 5. *Let E be in $\mathbb{K}\,\mathsf{RatE}\,A^*$. Let $D(\mathsf{E}) = \{\mathsf{K}_1, \mathsf{K}_2, \ldots, \mathsf{K}_n\}$ be the set of the derived terms of E. Let P be the* disjoint *union of $D(\mathsf{E})$ and an element K_0. The automaton $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$ is isomorphic to the automaton $\mathcal{B} = \langle P, A, Z, I, T\rangle$ defined by:*

$$I_{\mathsf{K}_i} = \begin{cases} 1_{\mathbb{K}} & \textit{if} \quad i = 0\,; \\ 0_{\mathbb{K}} & \textit{otherwise}\,. \end{cases} \qquad Z_{\mathsf{K}_i, \mathsf{K}_j} = \begin{cases} 0_{\mathbb{K}} & \textit{if} \quad j = 0\,; \\ \bigoplus_{a \in A} k_j^{(a)}\, a & \textit{if} \quad i = 0\,; \\ \bigoplus_{a \in A} z_{i,j}^{(a)}\, a & \textit{otherwise}\,. \end{cases}$$

$$T_{\mathsf{K}_i} = \begin{cases} \mathsf{c}(\mathsf{E}) & \textit{if} \quad i = 0\,; \\ \mathsf{c}(\mathsf{K}_i) & \textit{otherwise}\,. \end{cases}$$

where the $k_j^{(a)}$ and the $z_{i,j}^{(a)}$ have been defined at Theorem 2.

Proof. If $\mathcal{B}$ is different from $\mathcal{A}_{\mathsf{E}}$, then the expression E is equal to one of its derived terms: K_i. In this case, $\mathsf{c}(\mathsf{E}) = \mathsf{c}(\mathsf{K}_i)$, and, for every a in A, $\frac{\partial}{\partial a}\mathsf{E} = \frac{\partial}{\partial a}\mathsf{K}_i$. Hence, the terminal fonction is the same in E and K_i and the transitions that leave both states in the standard automaton of derived terms are the same. Hence, the $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$ is equal to $\mathcal{B}$.

If $\mathcal{B}$ is equal to $\mathcal{A}_{\mathsf{E}}$, the expression E is not a derived term and $\mathcal{A}_{\mathsf{E}}$ is standard. □

Proposition 7. *Let E be in $\mathbb{K}\,\mathsf{RatE}\,A^*$. There exists a $\mathbb{K}$-covering from $\mathcal{S}_{\mathsf{E}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$.*

Proof. The proof is by induction on the depth of the expression.

If $\mathsf{E} = \mathsf{0}$, $\mathsf{E} = \mathsf{1}$, or $\mathsf{E} = a$, $\mathcal{S}_{\mathsf{E}}$ and $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$ are equal. The identity is obviously a $\mathbb{K}$ covering.

There are two kind of cases whether there is or there is not a bijection between the set of derived terms of the expression E and the set of derived terms of the expressions from which E is built.

i) If $\mathsf{E} = (\mathsf{F}\,k)$, $\mathsf{s}(\mathcal{A}_{\mathsf{E}}) = \mathsf{s}(\mathcal{A}_{\mathsf{F}})\,k$, hence if there is a $\mathbb{K}$-covering φ from $\mathcal{S}_{\mathsf{F}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{F}})$, there is a $\mathbb{K}$-covering φ from $\mathcal{S}_{\mathsf{E}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$.

If $\mathsf{E} = (k\,\mathsf{F})$, $\mathsf{s}(\mathcal{A}_{\mathsf{E}}) = k\,\mathsf{s}(\mathcal{A}_{\mathsf{F}})$, hence if there is a $\mathbb{K}$-covering φ from $\mathcal{S}_{\mathsf{F}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{F}})$, there is a $\mathbb{K}$-covering φ from $\mathcal{S}_{\mathsf{E}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{E}})$.

If $\mathsf{E} = \mathsf{F}^*$, $\mathcal{S}_{\mathsf{E}} = \mathcal{S}_{\mathsf{F}}^*$; let φ_{F} be the $\mathbb{K}$-covering from $\mathcal{S}_{\mathsf{F}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{F}})$ There is a bijection from derived terms of F onto derived terms of E: the derived terms of E are the $(\mathsf{K}_i \cdot \mathsf{E})$, where K_i are the derived terms of F. Hence, $\mathsf{s}(\mathcal{A}_{\mathsf{E}}) = \mathsf{s}(\mathcal{A}_{\mathsf{F}})^*$. φ induces a $\mathbb{K}$-covering from $\mathcal{S}_{\mathsf{F}}^* = \mathcal{S}_{\mathsf{E}}$ onto $\mathsf{s}(\mathcal{A}_{\mathsf{F}})^* = \mathsf{s}(\mathcal{A}_{\mathsf{E}})$.

ii) The two last cases are respectively $\mathsf{E} = \mathsf{F} + \mathsf{G}$ and $\mathsf{E} = \mathsf{F} \cdot \mathsf{G}$. In these cases, one can form the sum (*resp.* the product) $\mathcal{B}$ of the automata $\mathsf{s}(\mathcal{A}_\mathsf{F})$ and $\mathsf{s}(\mathcal{A}_\mathsf{G})$. The problem is that $D(\mathsf{F})$ and $D(\mathsf{G})$ (*resp.* $D(\mathsf{F}).\mathsf{G}$ and $D(\mathsf{G})$) may have some elements in common.

If $\mathsf{E} = \mathsf{F} + \mathsf{G}$, let φ_F (*resp.* φ_G) be the $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{F}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{F})$ (*resp.* from $\mathcal{S}_\mathsf{G}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{G})$). There is a $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{E}$ onto $\mathcal{B} = \mathsf{s}(\mathcal{A}_\mathsf{F}) + \mathsf{s}(\mathcal{A}_\mathsf{G})$. We prove that there is a $\mathbb{K}$-covering from $\mathcal{B}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{E})$. Let φ be the canonical surjection from states of $\mathcal{B}$ onto states of $\mathsf{s}(\mathcal{A}_\mathsf{E})$ (the image of the initial state is the initial state, the image of another state (which is a derived term) is the state that is the same derived term). The mapping φ is not injective when some derived terms of F are derived terms of G. It is not difficult to check that φ is nevertheless a $\mathbb{K}$-covering.

If $\mathsf{E} = \mathsf{F} \cdot \mathsf{G}$, $\mathcal{S}_\mathsf{E} = \mathcal{S}_\mathsf{F} \cdot \mathcal{S}_\mathsf{G}$; let φ_F (*resp.* φ_G) be the $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{F}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{F})$ (*resp.* from $\mathcal{S}_\mathsf{G}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{G})$). The set of derived terms of E is the union of the set of derived terms of G and the $(\mathsf{K}_i \cdot \mathsf{G})$, where K_i are the derived terms of F. We consider $\mathcal{C} = \mathsf{s}(\mathcal{A}_\mathsf{F}) \cdot \mathsf{s}(\mathcal{A}_\mathsf{G})$. This automaton may be different from $\mathsf{s}(\mathcal{A}_\mathsf{E})$ because some derived terms of G may be equal to some $\mathsf{K}_i \cdot \mathsf{G}$.

However, there is a canonical surjection ψ from $\mathcal{C}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{E})$ that maps every pair (p_1, p_2) of states corresponding to the same derived term (the one coming from $(\mathsf{K}_i \cdot \mathsf{G})$, the other one from G) onto the same state p. None of these states is initial (hence $I_p = I_{p_1} \oplus I_{p_2}$). As they correspond to the same derived term K, $\mathsf{c}(\mathsf{K}) = T_p = T_{p_1} = T_{p_2}$.

For every pair (r, s) of states of $\mathsf{s}(\mathcal{A}_\mathsf{E})$, the transitions that comes from r depends only on the corresponding derived term K_r. Therefore, for every state p in $r\psi^{-1}$, $\mathsf{K}_p = \mathsf{K}_r$, and it holds:

$$E_{r,s} = \bigoplus_{a \in A} \langle \frac{\partial}{\partial a} \mathsf{K}_r, \mathsf{K}_s \rangle \, a = \bigoplus_{a \in A} \langle \frac{\partial}{\partial a} \mathsf{K}_p, \mathsf{K}_s \rangle \, a \, .$$

Thus ψ is a $\mathbb{K}$-covering. The following sequence holds:

$$\mathcal{S}_\mathsf{E} = \mathcal{S}_\mathsf{F} \cdot \mathcal{S}_\mathsf{G} \quad \xrightarrow{\varphi} \quad \mathcal{C} = \mathsf{s}(\mathcal{A}_\mathsf{F}) \cdot \mathsf{s}(\mathcal{A}_\mathsf{G}) \quad \xrightarrow{\psi} \quad \mathsf{s}(\mathcal{A}_\mathsf{E}) \, .$$

Therefore, there is a $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{E}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{E})$. $\square$

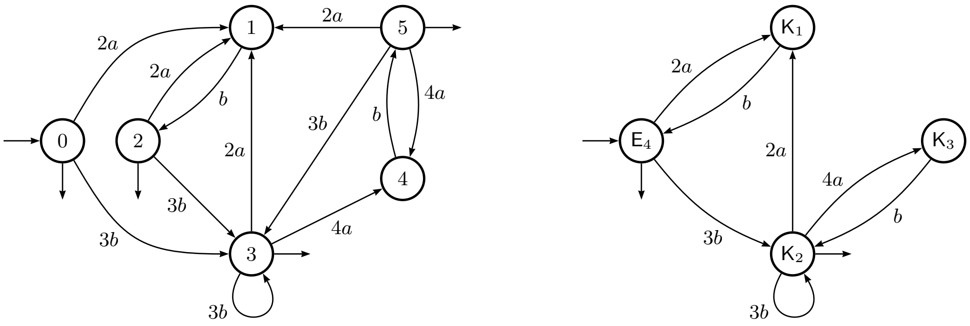
By Lemma 4, there is a $\mathbb{K}$-covering from $\mathsf{s}(\mathcal{A}_\mathsf{E})$ onto $\mathcal{A}_\mathsf{E}$; by Proposition 7,there is a $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{E}$ onto $\mathsf{s}(\mathcal{A}_\mathsf{E})$. Thus, there is a $\mathbb{K}$-covering from $\mathcal{S}_\mathsf{E}$ onto $\mathcal{A}_\mathsf{E}$ and this concludes the proof of Theorem 6. $\square$

Example 8. Let $\mathsf{E}_4 = \big(2ab + (3b) \cdot (4ab)^*\big)^*$. Figure 5 shows the standard automaton and the automaton of derived term of E_4 — we write $\mathsf{K}_1 = b\mathsf{E}_4$, $\mathsf{K}_2 = (4ab)^*\mathsf{E}_4$ and $\mathsf{K}_3 = b\mathsf{K}_2$. There is a $\mathbb{N}$-covering that maps states 0 and 2 onto E_4, and states 3 and 5 onto K_2.

Remark 9. As we said above, Proposition 7 holds because the derived terms of an expression have been defined as in Definition 3. We could have defined[5] the derived terms of an expression E as the smallest family $\mathcal{R}$ of expressions that contains E and such that any derivative of any element of $\mathcal{R}$ is a linear combination of elements of $\mathcal{R}$. The discrepancy between $D(\mathsf{E})$ and $\mathcal{R}$ is the possible existence of shadow terms that are hidden in the derivation with respect to a *letter*. And if these terms are missing in the automaton of derived terms, Theorem 6 does not hold anymore, as shown by the following example:

$$\mathsf{E}_5 = (1 - a)a^*.$$

[5] And this is what we have done in [13] where we had not stated Theorem 6.



(a) The standard automaton of E_4: $\mathcal{S}_{\mathsf{E}_4}$ (b) The automaton of derived terms of E_4: $\mathcal{A}_{\mathsf{E}_4}$

Fig. 5. Two $\mathbb{K}$-automata for E_4

Figure 6 represents $\mathcal{S}_{\mathsf{E}_5}$ and $\mathcal{A}_{\mathsf{E}_5}$. The singleton $\{\mathsf{E}_5\}$ itself is closed under derivation since $\frac{\partial}{\partial a}\,\mathsf{E}_5 = 0$ and it plays the role of the family $\mathcal{R}$. There is no $\mathbb{Z}$-covering from $\mathcal{S}_{\mathsf{E}_5}$ onto the automaton reduced to the single state E_5 of $\mathcal{A}_{\mathsf{E}_5}$.

(a) The standard automaton of E_5: $\mathcal{S}_{\mathsf{E}_5}$ (b) The automaton of derived terms of E_5: $\mathcal{A}_{\mathsf{E}_5}$

Fig. 6. Two $\mathbb{K}$-automata for E_5

8 Variations

The definition of derivation and of derived terms has been chosen such a way they can stand as a perfect generalization of those defined by Antimirov and, in particular, in order to get the same bound on the number of states of the automaton of derived terms and to establish Theorem 6. However, this definition makes the operation of derivation utterly faithful to the structure of the expression. As the goal of the derivation process is not only to state theorems but also to design algorithms that build automata from expressions, it may be efficient to enhance the derivation in such a way it perform some “semantic simplifications” such as grouping successive scalar multiplications or speeding up the splitting of derived terms.

This is what we shall define now with three kinds of derivation. It is remarkable that the three new derivations may be described by solely changing the definition of the derivation with respect to the empty word and its use. Recall that, when we have defined the derivation with respect to a word, we have assumed that the derivation with respect to the empty word is the identity:

$$\frac{\partial}{\partial 1_{A^*}}\,\mathsf{E} = \mathsf{E}.$$

This simple modification will imply deep differences in resulting automata.

8.1 Building the derivatives with unitary expressions

In the automaton of derived terms that we have defined, every state is characterized by an expression. So E and $(k\,\mathsf{E})$, which are of course distinct expressions for k different from $1_{\mathbb{K}}$, label two different states. We can define derivatives such that such a pair of states will finally be merged in the automaton of derived terms.

We define the *unitary derivation* with respect to the empty word, $\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}$, by induction on the depth of the expression:

$$\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{0} = \mathsf{0}, \qquad \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{1} = \mathsf{1}, \qquad \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,a = a,$$

$$\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}(k\,\mathsf{E}) = k\,\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}, \qquad \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}(\mathsf{E}\,k) = \left(\left[\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}\right]\,k\right),$$

$$\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}(\mathsf{E}+\mathsf{F}) = (\mathsf{E}+\mathsf{F}), \qquad \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}\right]\cdot\mathsf{F}\right),$$

$$\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}(\mathsf{E}^*) = (\mathsf{E}^*).$$

This operation is idempotent.

An expression E of $\mathbb{K}\,\mathsf{RatE}\,A^*$ is said to be *unitary* if $\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E} = \mathsf{E}$. Let us note that this definition is more subtle than to say that an expression E is unitary iff it is not of the form $(k\,\mathsf{F})$; for instance, $((2a)b)$ would be unitary then and it is not the definition we have chosen. For any E, $\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}$ is the product of a coefficient in $\mathbb{K}$ by a unitary expression.

We then define the *unitary derivation* with respect to a letter a as:

$$\forall \mathsf{E} \in \mathbb{K}\,\mathsf{RatE}\,A^* \quad \frac{\partial_{\mathsf{u}}}{\partial a}\,\mathsf{E} = \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\left[\frac{\partial}{\partial a}\,\mathsf{E}\right].$$

This unitary derivation is then extended to words by composition. The result of the unitary derivation with respect to any word of any expression is a linear combination of unitary expressions.

Remark 10. It is easy to prove (by induction on the depth of the expression) that, for every $\mathbb{K}$-expression E, it holds:

$$\frac{\partial}{\partial a}\left[\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}\right] = \frac{\partial}{\partial a}\,\mathsf{E}\,.$$

Hence, for every f in A^*, it comes:

$$\frac{\partial_{\mathsf{u}}}{\partial f}\,\mathsf{E} = \frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\left[\frac{\partial}{\partial f}\,\mathsf{E}\right].$$

Remark 11. The unitary derivation of an expression *is not* a process of factorization:

$$\frac{\partial_{\mathsf{u}}}{\partial a}\big((2\,a)+(4\,b)\big)=\big((2\,a)+(4\,b)\big),\qquad \frac{\partial_{\mathsf{u}}}{\partial a}\big(2\,\big(a+(2\,b)\big)\big)=2\,\big(a+(2\,b)\big).$$

As the aim of this new derivation is to build automata as small as possible, we abandon the formalism of Definition 3 and use now a closure property for the definition of derived terms.

Definition 11. *Let* E *be an expression in* $\mathbb{K}\,\mathsf{RatE}\,A^*$. *Let* K'_1 *be the unitary expression such that* $\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}\,\mathsf{E}=k_1\,\mathsf{K}'_1$ *and let* $P'_{\mathsf{E}}=\{\mathsf{K}'_1,\mathsf{K}'_2,\ldots,\mathsf{K}'_n\}$ *be the smallest set of unitary expressions that is closed under unitary derivation, that is such that there exist a set* $\{{z'}^{(a)}_{i,j}\in\mathbb{K}\mid i,j\in[n],a\in A\}$ *such that*

$$\frac{\partial_{\mathsf{u}}}{\partial a}\,\mathsf{K}'_i=\bigoplus_{j\in[n]}{z'}^{(a)}_{i,j}\,\mathsf{K}'_j. \tag{25}$$

The family P'_{E} *is called the set of* unitary derived terms *of* E.

With a proof analoguous to the one of Theorem 4, we have:

Theorem 8. *Let* E *be an expression in* $\mathbb{K}\,\mathsf{RatE}\,A^*$. *The* $\mathbb{K}$*-automaton* $\mathcal{A}'_{\mathsf{E}}=\langle P'_{\mathsf{E}},A,Z',I',T'\rangle$ *defined by:*

$$I'_{\mathsf{K}'_i}=\begin{cases}k_1 & \text{if}\quad i=1\\ 0_{\mathbb{K}} & \text{otherwise}\end{cases},\qquad Z'_{\mathsf{K}'_i,\mathsf{K}'_j}=\bigoplus_{a\in A}{z'}^{(a)}_{i,j}\,a,\qquad T'_{\mathsf{K}'_j}=\mathsf{c}(\mathsf{K}'_j)$$

realizes the series denoted by E: $|\mathcal{A}'_{\mathsf{E}}|=|\mathsf{E}|$. *The automaton* $\mathcal{A}'_{\mathsf{E}}$ *is called the* automaton of unitary derived terms *of* E. ☐

There is a canonical surjection of derived terms onto unitary derived terms (given by $\frac{\partial_{\mathsf{u}}}{\partial 1_{A^*}}$). Therefore, the number of unitary derived terms is smaller than the number of derived terms and, obviously, the number of states of the automaton of unitary derived terms is smaller than the litteral length of the expression.

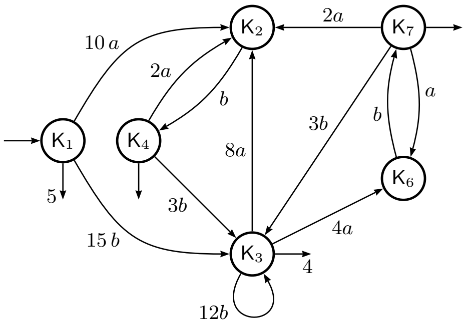
Example 9. Let[6] $\mathsf{F}_6=(2ab+(3b)\cdot(4(ab)^*))^*$; $\mathsf{E}_6=(5\,\mathsf{F}_6)$. We compute the derived terms (left column) and the unitary derived terms (right column):

$$\begin{aligned}
\mathsf{K}_1 &= \mathsf{E}_6 & \mathsf{K}'_1 &= \mathsf{F}_6\quad k_1=5\\
\frac{\partial}{\partial a}\,\mathsf{K}_1 &= 10\,(b\cdot\mathsf{F}_6)=10\,\mathsf{K}_2 & \frac{\partial_{\mathsf{u}}}{\partial a}\,\mathsf{K}'_1 &= 2\,(b\cdot\mathsf{F}_6)=2\,\mathsf{K}'_2\\
\frac{\partial}{\partial b}\,\mathsf{K}_1 &= 15\,((4\,(ab)^*)\cdot\mathsf{F}_6)=15\,\mathsf{K}_3 & \frac{\partial_{\mathsf{u}}}{\partial b}\,\mathsf{K}'_1 &= 12\,((ab)^*\cdot\mathsf{F}_6)=12\,\mathsf{K}'_3\\
\frac{\partial}{\partial b}\,\mathsf{K}_2 &= \mathsf{F}_6=\mathsf{K}_7 & \frac{\partial_{\mathsf{u}}}{\partial b}\,\mathsf{K}'_2 &= \mathsf{F}_6=\mathsf{K}'_1\\
\frac{\partial}{\partial a}\,\mathsf{K}_3 &= 4\,((b\cdot(ab)^*)\cdot\mathsf{F}_6)\oplus 8\,(b\cdot\mathsf{F}_6) & \frac{\partial_{\mathsf{u}}}{\partial a}\,\mathsf{K}'_3 &= ((b\cdot(ab)^*)\cdot\mathsf{F}_6)\oplus 2\,(b\cdot\mathsf{F}_6)\\
&= 4\,\mathsf{K}_6\oplus 8\,\mathsf{K}_2 & &= \mathsf{K}'_4\oplus 2\,\mathsf{K}'_2\\
\frac{\partial}{\partial b}\,\mathsf{K}_3 &= 12\,((4\,(ab)^*)\cdot\mathsf{F}_6)=12\,\mathsf{K}_3 & \frac{\partial_{\mathsf{u}}}{\partial b}\,\mathsf{K}'_3 &= 12\,((ab)^*\cdot\mathsf{F}_6)=12\,\mathsf{K}'_3
\end{aligned}$$

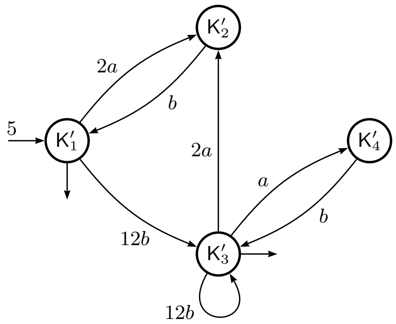
[6] The reader will note that F_6 and E_4 are *different* expressions, even if they look very much the same; E_6 is the example considered in [4].

$$\frac{\partial}{\partial a}\,\mathsf{K}_4 = 2\,(b\cdot\mathsf{F}_6) = 2\,\mathsf{K}_2$$
$$\frac{\partial}{\partial b}\,\mathsf{K}_4 = 3\,\mathsf{K}_3$$
$$\frac{\partial}{\partial b}\,\mathsf{K}_6 = ((ab)^*\cdot\mathsf{F}_6) = \mathsf{K}_7 \qquad \frac{\partial_{\mathsf{u}}}{\partial b}\,\mathsf{K}'_4 = ((ab)^*\cdot\mathsf{F}_6) = \mathsf{K}'_3$$
$$\frac{\partial}{\partial a}\,\mathsf{K}_7 = \mathsf{K}_6 \oplus 2\,\mathsf{K}_2$$
$$\frac{\partial}{\partial b}\,\mathsf{K}_7 = 3\,\mathsf{K}_3$$
$$\mathsf{c}(\mathsf{K}_1) = 5, \mathsf{c}(\mathsf{K}_2) = \mathsf{c}(\mathsf{K}_6) = 0 \qquad \mathsf{c}(\mathsf{K}'_1) = \mathsf{c}(\mathsf{K}'_3) = 1$$
$$\mathsf{c}(\mathsf{K}_3) = 4, \mathsf{c}(\mathsf{K}_4) = \mathsf{c}(\mathsf{K}_7) = 1 \qquad \mathsf{c}(\mathsf{K}'_2) = \mathsf{c}(\mathsf{K}'_4) = 0$$

Figure 7 a) shows the automaton of derived terms of E_6 (which is isomorphic to the standard automaton computed in [4]); b) shows the automaton of unitary derived terms of E_6.



(a) The automaton $\mathcal{A}_{\mathsf{E}_6}$



(b) The automaton of unitary derived terms

Fig. 7. Two $\mathbb{N}$-automata for E_6

8.2 Breaking the derived terms

The essence of derivation by a letter as defined by Antimirov and resumed here is to "break" the expression into pieces when the operator at the upper level of the expression is "+". The modification of the derivation we consider now consists in supposing that this breaking happens *spontaneously* by derivation with respect to the empty word.

The *breaking derivation* of an expression with respect to the empty word, $\dfrac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}$, is defined by induction on the depth of the expression:

$$\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{0} = \mathsf{0}, \qquad \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{1} = \mathsf{1}, \qquad \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,a = a,$$

$$\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}(k\,\mathsf{E}) = (k\,\mathsf{E}), \qquad \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}(\mathsf{E}\,k) = \left(\left[\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{E}\right]\,k\right),$$

$$\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}(\mathsf{E}+\mathsf{F}) = \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{E} \oplus \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{F}, \qquad \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{E}\right]\cdot\mathsf{F}\right),$$

$$\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}(\mathsf{E}^*) = (\mathsf{E}^*).$$

As the unitary derivation, the breaking derivation is an idempotent operation. The breaking derivation of an expression E with respect to a letter a is defined as:

$$\frac{\partial_{\mathsf{b}}}{\partial a}\,\mathsf{E} = \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\left[\frac{\partial}{\partial a}\,\mathsf{E}\right]. \tag{26}$$

This breaking derivation is then extended to words by composition.

Remark 12. It is easy to prove (by induction on the depth of the expression) that, for every $\mathbb{K}$-expression E, it holds:

$$\frac{\partial}{\partial a}\left[\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{E}\right] = \frac{\partial}{\partial a}\,\mathsf{E}. \tag{27}$$

Hence, by composing (26), it comes:

$$\frac{\partial_{\mathsf{b}}}{\partial f}\,\mathsf{E} = \frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\left[\frac{\partial}{\partial f}\,\mathsf{E}\right]. \tag{28}$$

Definition 12. *Let E be an expression in $\mathbb{K}\,\mathsf{RatE}\,A^*$. Let P''_I be the* support *of $\frac{\partial_{\mathsf{b}}}{\partial 1_{A^*}}\,\mathsf{E}$. Let $P''_{\mathsf{E}} = \{\mathsf{K}''_1, \mathsf{K}''_2, \ldots, \mathsf{K}''_n\}$ be the smallest set of expressions that contains P''_I and that is closed under breaking derivation. Let thus $\{z''^{(a)}_{i,j} \in \mathbb{K} \mid i,j \in [n], a \in A\}$ such that*

$$\frac{\partial_{\mathsf{b}}}{\partial a}\,\mathsf{K}''_i = \bigoplus_{j\in[n]} z''^{(a)}_{i,j}\,\mathsf{K}''_j\,. \tag{29}$$

The family P''_{E} is called the set of broken derived terms *of E.*

With a proof analoguous to the one of Theorem 4, we have:

Theorem 9. *Let E be an expression in $\mathbb{K}\,\mathsf{RatE}\,A^*$. The $\mathbb{K}$-automaton $\mathcal{A}''_{\mathsf{E}} = \langle P''_{\mathsf{E}}, A, Z'', I'', T''\rangle$ defined by:*

$$I''_{\mathsf{K}''_i} = \begin{cases} 1_{\mathbb{K}} & \text{if } \mathsf{K}''_i \in P''_I \\ 0_{\mathbb{K}} & \text{otherwise} \end{cases}, \qquad Z''_{\mathsf{K}''_i,\mathsf{K}''_j} = \bigoplus_{a\in A} z''^{(a)}_{i,j}\,a, \qquad T''_{\mathsf{K}''_j} = \mathsf{c}(\mathsf{K}''_j).$$

realizes the series denoted by E: $|\mathcal{A}''_{\mathsf{E}}| = |\mathsf{E}|$. The automaton $\mathcal{A}''_{\mathsf{E}}$ is called the automaton of broken derived terms *of E.* □

Example 10. Let $\mathsf{E}_7 = a^* + b^*$. Figure 8 a) shows the automaton of derived terms of E_7, b) shows the automaton of its broken derived terms.

(a) The automaton of derived terms $\mathcal{A}_{\mathsf{E}_7}$ (b) The automaton $\mathcal{B}_7$ of broken derived terms of E_7

Fig. 8. Two $\mathbb{K}$-automata for E_7

Remark 13. At first sight, breaking the derived terms increases their number; in any case, the straightforward majoration used in the proof of Theorem 2 does not hold anymore. Obvious examples—such as the one given by the expression $a + b$—shows that the number of broken derived terms may be greater than the one of “normal” derived terms. However, we conjecture, but it remains to be proved, that the number of broken derived terms of an expression E is smaller than or equal to $\ell(\mathsf{E}) + 1$.

8.3 Initial breaking of the derivatives

Clearly, the two methods above, computing unitary and broken derived terms, can be combined in various ways. In this case, we are equipped with solutions and awaiting problems. In fact, it is not a combination of these methods but just an *adaptation* of broken derived terms that we have devised recently when tackling the problem of finding an inverse to McNaughton-Yamada algorithm.

Let $\mathcal{A}$ be an automaton—that for sake of simplicity we assume here to be Boolean—and let $\mathsf{E} = \Phi(\mathcal{A})$ be an expression obtained from $\mathcal{A}$ by McNaughton-Yamada algorithm. The problem is to describe an algorithm Ψ that gives $\mathcal{A}$ back when applied to E. Although this problem is not completely solved yet, it has appeared that one of the components making up such a Ψ should be a variant of derivation that stands between the “normal” and the breaking ones (*cf.* [14]). Example 10 above illustrates this. An expression (the only one indeed) associated to $\mathcal{B}_7$ is E_7 and the broken derivation gives $\mathcal{B}_7$ back whereas the normal one produces $\mathcal{A}_{\mathsf{E}_7}$, from which it is rather difficult to recover (algorithmically) $\mathcal{B}_7$.

First we define the set of *initial derived terms* to be $\frac{\partial_{\flat}}{\partial 1_{A^*}} \mathsf{E}$: a polynomial of expressions with Boolean coefficients is a set of expressions. We then use the derivation as described at Definition 3. ¿From Equation (27), it follows that the terms obtained by derivation from the initial derived terms are indeed the derived terms. The automaton induced by this construction is defined by the following:
i) the set of states is the union of derived terms and initial derived terms,
ii) the initial states are the initial derived terms,
iii) a state K is final if and only if $\mathsf{c}(\mathsf{K}) = 1$,
iv) the triple $(\mathsf{K}, a, \mathsf{K}')$ is a transition of the automaton if and only if K' belongs to $\frac{\partial}{\partial a} \mathsf{K}$.

We then have proved in some cases (and observed in many more) that the algorithm Ψ we are looking for consists in taking the minimal co-quotient[7] of that automaton. The solution in the general case remains open.

[7] A co-quotient of $\mathcal{A}$ is the transposed automaton of a quotient of the transposed automaton of $\mathcal{A}$.

Example 11. Let us consider the automaton $\mathcal{A}_8$ of Figure 9 (a). A run of the McNaughton-Yamada algorithm may (depending on an ordering on states) give the following expression:

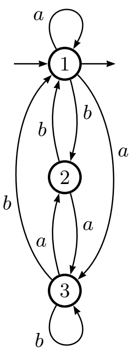
$$\mathsf{E}_8 = a^* + a^*b(ba^*b)^*ba^* + \big(a^*a + a^*b(ba^*b)^*(a+ba^*a)\big)\,\mathsf{F}_8\ , \quad \text{where}$$
$$\mathsf{F}_8 = \big(b + ba^*a + (a+ba^*b)(ba^*b)^*(a+ba^*a)\big)^*\,\big(ba^* + (a+ba^*b)(ba^*b)^*ba^*\big)\ .$$

The initial derived terms of E_8, obtained by the breaking derivation, are:

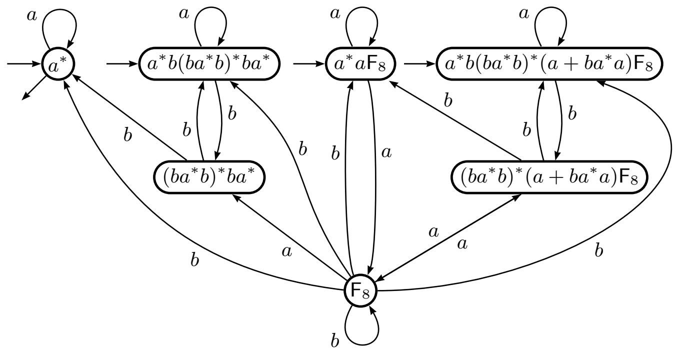
$$\{a^*,\ a^*b(ba^*b)^*ba^*,\ a^*a\mathsf{F}_8,\ a^*b(ba^*b)^*(a+ba^*a)\mathsf{F}_8\}.$$

And the derived terms of E_8 are read on the derived term automaton $\mathcal{B}_8$ itself (Figure 9 (b)).

It is quite easy to check that the minimal co-quotient of the automaton $\mathcal{B}_8$ is the automaton $\mathcal{A}_8$ itself. Actually, the morphism φ that maps every state p of $\mathcal{B}_8$ onto the state of $\mathcal{A}_8$ drawn on the same line induces a surjective mapping from the incoming transitions of p onto those of $\varphi(p)$.



(a) The automaton $\mathcal{A}_8$



(b) The derived terms automaton of E_8 with initial terms

Fig. 9. Initial breaking derivatives

References

1. V. Antimirov, Partial derivatives of regular expressions and finite automaton constructions. *Theoret. Comput. Sci. 155* (1996), 291–319.
2. J. Berstel and Ch. Reutenauer, *Les séries rationnelles et leurs langages.* Masson, 1984. Translation: *Rational Series and their Languages.* Springer, 1986.
3. J. A. Brzozowski, Derivatives of regular expressions. *J. Assoc. Comput. Mach. 11* (1964), 481–494.
4. P. Caron and M. Flouret, Glushkov construction for multiplicities. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 52–61.
5. P. Caron and M. Flouret, From Glushkov WFA's to Rational Expressions. *Proceedings of DLT'03*, Z.Ésik and Z.Fülöp, eds, LNCS 2710 (2003), 183–193.
6. P. Caron and M. Flouret, Characterization of Glushkov automata. *Theoret. Comput. Sci. 233* (2000), 75–90.
7. J.-M. Champarnaud and D. Ziadi, New finite automaton constructions based on canonical derivatives. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 36–43.
8. J.-M. Champarnaud and D. Ziadi, Canonical derivatives, partial derivatives and finite automaton constructions. *Theoret. Comput. Sci. 289* (2002), 137–163.
9. J.-M. Champarnaud and G. Duchamp , Brzozowski's derivatives extended to multiplicities. *Proceedings of CIAA'01*, LNCS 2494 (2001), 52–64.

10. S. Eilenberg, *Automata, Languages and Machines, volume A*. Academic Press, 1974.
11. V. Glushkov, The abstract theory of automata. *Russian Mathematical Surveys 16* (1961), 1–53.
12. W. Kuich, and A. Salomaa, *Semirings, Automata, Languages*. Springer, 1986.
13. S. Lombardy and J. Sakarovitch, Derivatives of rational expressions with multiplicity, in *MFCS 02*, LNCS 2420 (2002), 471–482.
14. S. Lombardy and J. Sakarovitch, How expressions can code for automata, in *Latin 04*, LNCS 2976 (2004), 242–251. Journal version to appear in *Theoret. Informatic and Applications*.
15. R. McNaughton, and H. Yamada, Regular expressions and state graphs for automata. *IRE Trans. on Electronic Computers 9* (1960), 39–47.
16. J.J.M.M. Rutten, Automata, power series, and coinduction: taking input derivatives seriously, in *ICALP 99*, J. Wiedermann, P. van Emde Boas et M. Nielsen (éd.), LNCS 1644 (1999), 645–654.
17. J.J.M.M. Rutten, Behavioural differential equations: a coinductive calculus of streams, automata, and power series. *Theoret. Comput. Sci. 308* (2003), 1–53.
18. J. Sakarovitch, *Eléments de théorie des automates*, Vuibert, 2003. Translation: *Elements of Automata Theory*, Cambridge Universiy Press, to appear.