A SURVEY ON NONCOMMUTATIVE RATIONAL SERIES

CHRISTOPHE REUTENAUER

Université du Québec à Montréal

Département de mathématiques et d'Informatique

C.P. 8888, succursale Centre Ville

Montréal (Québec) H3C 3P8

§1. INTRODUCTION

Since the title of this conference is *"Formal power series ..."*, it seems justified to give a survey on noncommutative series. We shall concentrate only on rational series. They are not very known in algebraic or enumerative combinatorics, although their one-variable version, the linear recursive sequences, are well-known.

Noncommutative rational series, as they are considered in this survey, appear as a generalization of languages which are recognizable by a finite automaton; they count the multiplicities of the recognition process. On the other hand, Kleene's theorem **(a language is recognizable if and only if it is regular)** lead Schützenberger to the observation that the Kleene star operation on languages is actually an inversion, and that his theorem was quite close to the characterization of linear recursive sequences through rational fractions.

We try here to give an overview of the algebraic, arithmetic and combinatorial aspects of rational series, together with applications to variable-length codes. No proof is given. If not precisely indicated, the proofs of the theorems may be found in [BR88], or [E74], [SS78].

§2. RATIONAL AND RECOGNIZABLE SERIES

Let $\mathbb{Q}\langle\langle A\rangle\rangle$ denote the $\mathbb{Q}$-algebra of noncommutative formal **series** in the variables in A (we take Q for simplicity). A **polynomial** is an element of $\mathbb{Q}\langle A\rangle$, which is the subalgebra of $\mathbb{Q}\langle\langle A\rangle\rangle$ generated by the variables. A **rational** series is an element obtained from polynomials by algebraic operations (sum, product), and inversion (recall that a series is invertible if and only its constant term is invertible).

Example : $A = \{a, b\}$; $(1 - a - a^2)^{-1} = \sum_{n\geq 0} F_n a^n$, $(1 - a - b(1 - a)^{-1} b)^{-1}$ = sum of the words in A^* (the free monoid generated by A) which have an even number of b's,

$(1 - a - b)^{-1} b (1 - 2a - 2b)^{-1} = \sum_{w\in A^*} \alpha_w w$, where α_w is the integer represented in binary system

by the word w (a for 0, b for 1). Note that a theory of series generalizing the third example is presented in [AS92], under the name "k-regular sequences".

A series S is **recognizable** if $\delta(S) = \sum_{\text{finite}} S_i \otimes T_i$, where δ is the mapping such that for any word w, one has $\delta(w) = \sum_{w=uv} u \otimes v$ ($w = uv$ in the free monoïd A^*). Observe that there is a natural pairing $\mathbb{Q}\langle\langle A\rangle\rangle \times \mathbb{Q}\langle A\rangle \to \mathbb{Q}$, defined by $\langle S,P\rangle = \sum_{w\in A^*} \langle S,w\rangle\langle P,w\rangle$ (where $\langle S,w\rangle$ is the coefficient of w in S). In this way, $\mathbb{Q}\langle\langle A\rangle\rangle$ is the dual space of $\mathbb{Q}\langle A\rangle$, and a recognizable series is a **representative function**, in the sense of Hopf algebraists. A more combinatorial, or computer-scientist, look at recognizable series comes from automata theory. We illustrate if by some examples. Look at the graphs of figure 1.

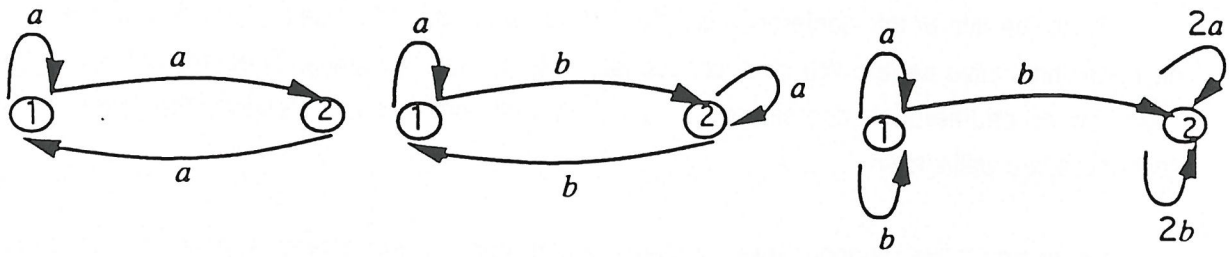

Figure 1

Let w a fixed word and count the number of paths from 1 to 1 in the graphs (in the third one, the paths from 1 to 2, and one counts each path with its multiplicity, which for a given path is the product of the multiplicities of the edges). Then this number is the coefficient in the corresponding series of the previous example.

An equivalent definition of recognizable series involves **Hankel matrices** : a series is recognizable if and only the infinite matrix $(\langle S, uv\rangle)_{u,v\in A^*}$ has finite rank (cf.[CP71], [F74], [J75]).

The fundamental result is due to Schützenberger [S61] : **a series if rational if and only it is recognizable.**

§3. SYNTACTIC ALGEBRA AND LINEAR RECURSION

The previous Hopf-algebraic approach, and the automata theory point of view (through the minimal automaton and the syntactic monoïd), both lead to the following definition : the **syntactic algebra** of a series S is the quotient of the algebra $\mathbb{Q}\langle A\rangle$ by the biggest two-sided ideal contained in Ker S (S being viewed as a linear function on $\mathbb{Q}\langle A\rangle$). The Schützenberger theorem then takes the form : **a series is**

rational if and only if its syntactic algebra is finite-dimensional.

In the one-variable case, it is well-known that $\sum_{n\geq 0} r_n a^n$ is rational if and only if the sequence (r_n) satisfies a linear recursion : $r_{n+k} + \alpha_1 r_{n+k-1} + \ldots + \alpha_k r_n = 0$ for any n in $\mathbb{N}$ ($\alpha_1, \ldots, \alpha_k$ are constants in $\mathbb{Q}$). Taking k minimum (i.e. the shortest such recursion), the syntactic algebra is $\mathbb{Q}[a]/(a^k + \alpha_1 a^{k-1} + \ldots + \alpha_k)$.

In the noncommutative case, linear recursion also exist [S61, BR88]. Their existence is in strong relationship with the fact that **each right ideal of $\mathbb{Q}\langle A\rangle$ is a free $\mathbb{Q}\langle A\rangle$ - module** (Cohn [C61]). We only give an example : let $S = \sum_{w\in\{a,b\}^*} \alpha_w w$, where α_w is as before the number represented in binary system by w (a for 0, b for 1). Then one has for any word w :

$$\alpha_{aw} = \alpha_w, \alpha_{baw} = 2\alpha_{bw} - \alpha_w, \alpha_{bbw} = 3\alpha_{bw} - 2\alpha_w$$

The fact that these relations, together with the initial conditions $\alpha_1 = 0$, $\alpha_b = 1$, completely define S, is a consequence of the fact that $\{a, ba, bb\}$ is a complete prefix code, with set of left factors $\{1, b\}$ (for the definition of a code, see Sect. 9).

The syntactic algebra $\mathcal{A}$ of a rational series S reflects the properties of S. For instance, the coefficient $\langle S, w\rangle$ depends only on the commutative image of w if and only if $\mathcal{A}$ is commutative; this kind of result leads to the theory of pseudo-varieties of finite dimensional algebras and rational series ([R80]), analog to that of monoids and languages (see [E74], [P84]).

Another result illustrating the relationship between S and $\mathcal{A}$ is the following : if $S = (1 - \underline{C})^{-1}$, where C is a maximal code and $\underline{C} = \sum_{w\in C} w$, then $\mathcal{A}$ is semi-simple if and only if C is a biprefix code (see [BP85]).

§4. LANGUAGES

A **language** (i.e. a subset of A^*) is **rational** (regular) if it may be obtained from the finite languages by applying finitely many times the operations $L_1 \cup L_2$, $L_1 L_2 = \{w_1 w_2 \mid w_i \in L_i\}$ and $L^* = \bigcup_{n\geq 0} L^n$ = submonoid generated by L.

Furthermore, a language is called **recognizable** (by a finite automaton) if it is the set of labels of paths in a given directed graph (with edges labeled in A), where these paths go from a given vertex to some vertex in a given set of vertices.

The **characteristic series** of a language L is $\underline{L} = \sum_{w \in L} w$. Then, **a language is rational (resp. recognizable) if and only if its characteristic series is**. From this, one can deduce Kleene's theorem : a **language is rational if and only if it is recognizable**. Actually, this theorem was proved before Schützenberger's one, and the path from the first to the second was certainly guided by the observation of the following striking fact : the operation $L^{\cdot} = \bigcup L^n$ has a close analogy with the inversion $(1 - S)^{-1} = \sum_{n \geq 0} S^n$.

Besides rational languages, there is a bigger family of languages, closely related to rational series, consisting of the **supports** of these series. We define, for a series S, $\mathrm{supp}(S) = \{w \in A^{\cdot} \mid \langle S, w \rangle \neq 0\}$. This family satisfies the usual closure properties of language theory (direct and inverse homomorphism, intersection, union). The analogy with complement of varieties (in algebraic geometry) leads to the following question : what happens when two supports are disjoint ? The conjecture is that **two supports are disjoint if and only they are rationally separated**, i.e. if there is a rational language containing one and not intersecting the other. Note that rational languages are closed under complementation. A particular case of this conjecture has been proved [RR84] : **if a language and its complement are both supports, then they are rational**. This shows also that the family of supports is not closed under complementation, except in trivial cases.

Another open question, raised in [SS78], is the following : **if L is the support of some rational series in $\mathbb{R}\langle\langle X \rangle\rangle$, is it also the support of some rational series in $\mathbb{Q}\langle\langle X \rangle\rangle$** ?

§5. EXTENSIONS

The previous question leads to the problem of extending the field of coefficients. Actually, rational and recognizable series may be defined over any semi-ring (inversion is replaced by the **star** operation, with $S^{\cdot} = \sum_{n \geq 0} S^n$ if S has no constant term). Then Schützenberger's theorem is still valid. The question is to know if $K \subseteq L$ is an extension of semi-ring and if $S \in K\langle\langle X \rangle\rangle$ is rational in $L\langle\langle X \rangle\rangle$, is it then rational in $K\langle\langle X \rangle\rangle$? The answer is known to be positive when K, L are fields, or if K is a noetherian ring and L its field of fractions. A nice result of Fliess [F] also shows that it is true for the extension of semi-rings $\mathbb{N} \subseteq \mathbb{Q}_+$. However, it is not true for $\mathbb{Q}_+ \subseteq \mathbb{R}_+$, nor $\mathbb{N} \subseteq \mathbb{Z}$.

This question is even interesting even for one-variable series : in this case, one defines rational series in $\mathbb{N}[[a]]$, which are shown to be exactly the generating series of rational languages. It was shown by Berstel [B71] that these series have "dominating roots", thus are closely related to the Pisot-Vijanaragayan numbers. The converse was also proved ([S76], [KOE78]), and these series are thus completely characterized.

§6. UNAMBIGUOUS SERIES

The **Hadamard product** of two series S, T is $S \odot T = \sum_w \langle S, w\rangle \langle T, w\rangle w$. **If these series are both rational, then so is their Hadamard product** ([S]). The problem is that of invertibility. In general, the Hadamard inverse of a rational series is not rational, e.g. $\sum_{n\geq 0}(n+1)^{-1}\, a^n$ is not rational (the primes appearing as factors of the denominators of the coefficients of a rational series are in finite number : Eisenstein criterion). Let us be slightly more general. We say that a rational series S in $\mathbb{Q}\langle\langle X\rangle\rangle$ is **Hadamard sub-invertible** if the series $\sum_{w,\langle S,w\rangle\neq 0}\langle S,w\rangle^{-1}\, w$ is rational. Equivalently, S is a regular element of the Hadamard algebra of rational series ; note that the indempotents of this algebra are by Sect. 4 the characteristic series of rational languages, and S is regular if and only if there exists a rational series T such that $S \odot T = \underline{L}$, with $L = \mathrm{supp}(S)$ a rational language.

Denote by $P(S)$ the set of primes p such that p divides the numerator or the denominator of some nonzero coefficient (in reduced form) of S. Then S is called a **Pólya series** if $P(S)$ is finite.

The problem is to characterize Hadamard sub-invertible series, and Pólya series. For this, we need one definition more. Define the **unambiguous rational operations** : the sum $S + T$ is unambiguous if $\mathrm{supp}(S) \cap \mathrm{supp}(T) = \varnothing$; the product $S\,T$ is unambiguous if each word in $\mathrm{supp}(S) \cdot \mathrm{supp}(T)$ has a unique factorization $u\,v$ with $u \in \mathrm{supp}(S\,)$, $v \in \mathrm{supp}(T)$; the star $S^* = \sum_{n\geq 0} S^n$ is unambiguous if $\mathrm{supp}(S)$ is a code. Then a series is called **unambiguously rational** if it is obtained from polynomials by a finite number of unambiguous rational operations.

It is not difficult to verify that each unambiguously rational series is both Hadamard sub-invertible and a Pólya series. The conjecture is the converse : **each Hadamard sub-invertible rational series, and each rational Pólya series, is unambiguously rational.**

In the one-variable case, this conjecture is a theorem of Pólya [P21] (for Pólya series), and a theorem of Benzaghou [Bn70] (for the Hadamard invertibility); see also [Bz84], [LT90]. In several noncommuting variables, particular cases are known [R80a], [R79].

§7. DECIDABILITY

Most of the questions concerning rational series, or their supports, are undecidable. This is because that one can easily encode the 10th Hilbert problem (undecidability of the solvability of a diophantine equation). For example, it is undecidable if a given support is equal to the whole free monoïd A^*. In the one-variable case however, this is still open and known under the name of Pisot problem.

However, the question to know if the set of coefficients of a given rational series is finite is decidable, and related to the Burnside problem for monoids of matrices (see [J78], [MS77], [S62]).

§8. RATIONAL IDENTITIES

Each rational series may be written as a **rational expression**, which is a well-formed expression using the letters in A, the scalars in $\mathbb{Q}$, the symbols $+$, $\times$ and * (one takes usually * instead of inversion). For example, the sum of words in $\{a, b\}^*$ having an even number of b's is $(a + ba^*b)^*$. For a given series, the rational expression is far from

being unique. This leads to the problem of **rational identities**. For example, one has $(ab)^* = 1 + a(ba)^*b$, $(a + b)^* = (a^*b)^*a^*$.

Observe that both identities have evident combinatorial interpretations : the first one corresponds to the two ways one can look at a word of the form $ab\ ab\ ...ab$ (try to say repetitively ab, then after a while you do not know any more if you repeat ab or ba); the second describes how to cut a word on a, b after each b.

But these identities may also be proved by using algebraic operations and the fact that S^* is the inverse of $1 - S$. Indeed, if we multiply both sides of the first identity by $1 - ab$ on the left, then the left -hand side becomes 1, and the right-hand side becomes

$$1 + a(ba)^*b - ab - aba(ba)^*b = 1 - ab + a(1 - ba)(ba)^*b = 1 - ab + ab = 1.$$

For the second, we multiply by $1 - a - b$ on the right ; the right-hand side becomes

$$(a^*b)^*a^* - (a^*b)^*a^*a - (a^*b)^*a^*b = (a^*b)^*a^*(1 - a) - (a^*b)^*a^*b = (a^*b)^* - (a^*b)^*a^*b$$
$$= (a^*b)^*(1 - a^*b) = 1,$$

what was to be shown.

These calculations show that the previous identities are "trivial", i.e. are algebraic consequences of the fact that $S^* = (1 - S)^{-1}$. It is shown in [K91] that this holds for all rational identities. This result is closely related to Amitsur's theorem on rational identities in a skew field, see [A66], [Be70], [C77]. When the field of coefficients is replaced by a semi-ring, the theory is quite different, see [Co71], [K92].

§9. CODES

A **code** is the (unique) basis of some free submonoïd of A^*. Series are well adapted for the study of codes. A simple reason for that is that if C is a subset of $A^*\backslash 1$, then C is a code if and only if $\underline{C}^* = \underline{C^*}$.

The main conjecture in the theory of codes is the **factorization conjecture** : given a **finite and maximal code** C, **there exist finite sets** S **and** P **of words such that** $\underline{A}^* = \underline{S}\,\underline{C}^*\,\underline{P}$, i. e. each word in A^* has a unique decomposition $s\,c_1 \dots c_n p$ with $s \in S, c_1 \dots, c_n \in C, p \in P$. Note the analogy with the coset decomposition of a group with respect to a subgroup. It has been shown that there exist always a decomposition $\underline{A}^* = \underline{S}\,\underline{C}^*\,\underline{P} + \underline{F}$ with F a finite set ([BR88], [ZG92]). The previous conjecture, by inversion of series, may also be stated as follows : there exist polynomials P_1, S_1 in $\mathbb{N}\langle A\rangle$ such that $\underline{C} - 1 = P_1(\underline{A} - 1)\,S_1$. A weak form of this has been proved, with $P_1, S_1 \in \mathbb{Z}\langle A\rangle$, see [BR88].

The factorization conjecture is easy to prove in the case of a **prefix** code (i.e. no word in the code is a left factor of another word of the code). A code is **biprefix** if it is prefix, and if also its reversal is prefix. When C is a **thin** (thin is a generalization of rational) maximal biprefix code, one defines its **indicator** to be the series L such that $\langle L, w\rangle$ = number of factorizations $w = sc_1 \dots c_n\, p$, where s(resp. p) is a right (resp. left) factor of some word in C, and where the c_i are in C (cf. figure 2)

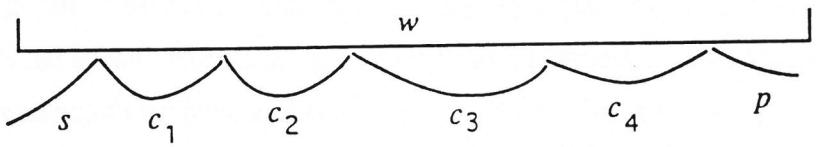

Figure 2

Then L has only finitely many coefficients, and the **degree** d of C is defined to be the maximal coefficient. The **tower** of C is the series $T = \sum_w (d - \langle L, w\rangle)\, w$, and one shows that $\underline{C} - 1 = (\underline{A} - 1)\, T\, (\underline{A} - 1) + d(\underline{A} - 1)$.

Example : $C = \{a^3, a^2ba, a^2b^2, ab, ba^2, baba, bab^2, b^2a, b^3\}$. Then

$$\underline{C} - 1 = (a + b - 1)\,(2 + a + b + 1)\,(a + b - 1) + 3(a + b - 1).$$

If T' is defined from T by substracting 1 to each nonzero coefficient of T, then T' is the tower of a new biprefix code, called the **derived code**, of degree $d-1$. In the previous example, the derived code is therefore $\underline{C'} - 1 = (a+b-1)\ (1)\ (a+b-1) + 2(a+b-1) = (a+b)^2 - 1$, hence $C' = \{a^2, ab, ba, b^2\}$.

One can construct all biprefix codes by reversing the previous process. This is a short overview of the Césari [Cé79] theory of biprefix codes, as it has been exposed and extended in the book by Berstel and Perrin [BP85].

References

[A66] Amitsur S.A., *"Rational identities, and applications to algebra and geometry"*, J. Algebra 3, 1966, 304-359.

[AS92] Allouche J. P., Shallit J., *"The ring of k-regular sequences"*, Theor. Comput, Sci. 98, 163-197, 1992.

[B71] Berstel J., *"Sur les pôles et le quotient de Hadamard de séries N-rationelles"*, C.R. Acad, Sci. Paris Sér. A 272, 1079-1081, 1971.

[BP85] Berstel J., Perrin D., *"Theory of codes"*, Academic Press, 1985.

[BR88] Berstel J., Reutenauer C. , *"Rational series and their languages"*, Springer Verlag, 1988.

[Be70] Bergman G. M., *"Skew fields of noncommutative rational functions, after Amitsur"*, Séminaire Schützenberger-Lentin Nivat 1969/70, exposé n. 16, 1970.

[Bn70] Benzaghou B., *"Algèbres de Hadamard"*, Bull. Soc. Math. France 98, 209-252, 1970.

[Bz84] Bezivin J.-P., *"Factorisation de suites récurrentes linéaires et applications"*, Bull. Soc. Math France 112, 365-376, 1984.

[C61] Cohn P.M., *"On a generalization of the Euclidean algorithm"* Proc. Cambridge Philos. Soc. 57, 18-30, 1961.

[C77] Cohn P.M., *"Skew field constructions"*, London Mathematical Lect. Note Series 27, Cambridge Univ. Press, 1977.

[Cé79] Césari Y., *"Propriétés combinatoires des codes bipréfixes"*, in : Théorie des Codes, D. Perrin ed., LITP, ENSTA, Paris, 29-46, 1979.

[Co71] Conway J. H., *"Regular algebras and finite machines"* Chapman & Hall, London, 1971.

[CP71] Carlyle J. W., Paz A., *"Realizations by stochastic finite automaton"*, J. Comput, System Sci. 5, 26-40, 1971.

[E74S] Eilenberg S., *"Automata, languages and machines"*, vol. A, Academic Press, 1974.

[E76] Eilenberg S. , *"Automata, languages and machines"*, vol. B, Academic Press, 1976."

[F74] Fliess M., *"Matrices de Hankel"*, J. Maths Pures Appl. 53, 197-222, 1974.

[F75M] Fliess M., *"Séries rationnelles positives et processus stochastiques"*, Ann Inst. H. Poincaré Sect. B 11(2), 1-21, 1975.

[J75] Jacob G., *"Représentations et substitutions matricielles dans la théorie algébrique des transductions"*, Thesis, Univ. of Paris, 1975.

[J78] Jacob G., *"La finitude des représentations linéaires des semi-groupes est décidable"*, J. Algebra 52, 437-459, 1978.

[K91] Krob D., *"Expressions rationnelles sur un anneau"*, Lect. Notes Math. 1478 (Topics in invariant therory, M.-P. Malliavin, ed.), 215-243, Springer Verlag, 1991.

[K92] Krob D., *"Models of a K-rational identity system"*, J. Comput. System Sci. 45, 396-434, 1992.

[KOE78] Katayama T., Okamoto M., Enomoto H., *"Characterization of the structure generating functions of regular sets and DOL growth functions"*, Information and Control 36, 85-101, 1978.

[LT90] Larson R., Taft E., *"The algebraic structure of linearly recursive sequences under Hadamard product"*, Israel J. Math 72, 118-132, 1990.

[MS77] Mandel A., Simon I., *"On finite semi-groups of matrices"*, Theor, Comput. Sci. 5, 101-111, 1977.

[P21] Pólya G., *"Arithmetische Eigenschaften der Reihenentwicklungen rationaler Funktionen"*, J. reine angew. Math 151, 1-31, 1921.

[P86] Pin J.-E. , *"Varieties of formal languages"*, North Oxford, London and Plenum, New York, 1986.

[R79] Reutenauer C., *"Séries de Pólya en variables noncommutatives"*, Proc. of the conference "Fundamentals of Computer Science", 391-396, 1979, Akademic Verlag, Berlin.

[R80a] Reutenauer C., *"Séries formelles et algèbres syntactiques"*, Thèse, Université Paris 6, 1980.

[R80] Reutenauer C., *"Séries formelles et algèbres syntactiques"*, J. Algebra 66, 448-483, 1980.

[RR84] Restivo A., Reutenauer C., *"On cancellation properties of languages which are support of rational power series"*, J. Comput. System Sci. 5, 153-159, 1984.

[S61] Schützenberger M.-P., *"On the definition of a family of automata"*, Information and Control 4, 245-270, 1961, 36, 167-186, 1985.

[S62] Schützenberger M.-P., *"Finite counting automata"*, Information and Control 5, 91-107, 1962.

[So76] Soittola M., *"Positive rational sequences"*, Theor. Comput. Sci. 2, 317-322, 1976.

[SS78] Salomaa A., Soittola M., *"Automata-theoretic aspects of formal power series"*, Springer Verlag, 1978.

[ZG92] Zhang L., Gu C.K., *"On factorization of finite maximal codes", in : "Words, languages and combinatorics"*, M. Ito ed., World Scientific, 534-542, 1992.

A SURVEY ON NONCOMMUTATIVE RATIONAL SERIES

Christophe Reutenauer
Université du Québec à Montréal

Reading List

1. S. Eilenberg, Automata, Languages and Machines, Acad. Press, Vol. A, 1974.

2. A. Salomaa, M. Soittola, Automata - Theoretic Aspects of Formal Power Series, Springer Verlag, 1978.

3. J. Berstel, C. Reutenauer, Rational Series and Their Languages, Springer Verlag, 1988.