On the Maximum Coefficients of Rational Formal Series in Commuting Variables

Christian Choffrut[1], Massimiliano Goldwurm[2], and Violetta Lonati[2]

View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by CiteSeerX

[2] Dipartimento di Scienze dell'Informazione, Universit degli Studi di Milano
Milano, Italy
{goldwurm, lonati}@dsi.unimi.it

Abstract. We study the *maximum function* of any $\mathbb{R}_+$-rational formal series S in two commuting variables, which assigns to every integer $n \in \mathbb{N}$, the maximum coefficient of the monomials of degree n. We show that if S is a power of any primitive rational formal series, then its maximum function is of the order $\Theta(n^{k/2}\lambda^n)$ for some integer $k \geq -1$ and some positive real λ. Our analysis is related to the study of limit distributions in pattern statistics. In particular, we prove a general criterion for establishing Gaussian local limit laws for sequences of discrete positive random variables.

1 Introduction

The general observation motivating this paper is the following. Consider a rational fraction $\frac{p(x)}{q(x)}$ where $p(x)$ and $q(x)$ are two polynomials with coefficients in the field of real numbers (with $q(0) \neq 0$). It is well-known that the coefficient of the term x^n of its Taylor expansion is asymptotically equivalent to a linear combination of expressions of the form $n^{k-1}\lambda^n$ where λ is a root of $q(x)$ and k its multiplicity, cf. [10, Theorem 6.8] or [16, Lemma II.9.7]. It is natural to ask whether a similar evaluation holds for formal series in two variables both in the commutative and in the noncommutative case.

The purpose of this work is to pose the problem in more general terms. Indeed, assume we are given a semiring $\mathbb{K}$ whose underlying set is contained in the reals, e.g., $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{R}_+$, etc. Assume further we are given a monoid $\mathcal{M}$ that is either the free monoid or the free commutative monoid over a finite alphabet. Consider next a rational formal series r over $\mathcal{M}$ with coefficients in $\mathbb{K}$ and denote by (r, x) the corresponding *coefficient* of the element $x \in \mathcal{M}$. Denoting by $|x|$ the length of x, we want to investigate the *maximum function* $g_r(n) = \max\{|(r,x)| : x \in \mathcal{M}, |x| = n\}$, where $|(r,x)|$ denotes the absolute value of (r,x). We note that the same definition can be extended to the formal series over trace monoids [6].

For rational formal series over a free monoid with integer coefficients, the growth of the coefficients was investigated in [17] (see also [14]), where it is

C.S. Calude, E. Calude and M.J. Dinneen (Eds.): DLT 2004, LNCS 3340, pp. 114–126, 2004.
© Springer-Verlag Berlin Heidelberg 2004

proved that for such a series r either there exists $k \in \mathbb{N}$ such that $g_r(n) = \mathrm{O}(n^k)$ or $|(r, \omega_j)| \geq c^{|\omega_j|}$ for a sequence of words $\{\omega_j\}$ of increasing length and for some constant $c > 1$. In the first case, the series is the sum of products of at most $k + 1$ characteristic series of regular languages over the free monoid (see also [2, Corollary 2.11]). When the semiring of coefficients is $\mathbb{N}$ the problem is related to the analysis of ambiguity of formal grammar (or finite automata) generating (recognizing, resp.) the support of the series; a wide literature has been devoted to this problem (see for instance [12, 20] and [21, 22] for a similar analysis in the algebraic case).

As far as the growth of coefficients is concerned, another case can be found in the literature. It is related to the tropical semiring $\mathcal{T}$ whose support is the set $\mathbb{N} \cup \{\infty\}$ and whose operations are the min for the addition and the + for the multiplication. In [19], Imre Simon proves that for all $\mathcal{T}$-rational series s over the free monoid $\{a, b\}^*$, there exists an integer k such that $(s, w) = O(|w|^{1/k})$ holds for all $w \in \{a, b\}^*$. Moreover it is proven that for each positive integer k, there exists a $\mathcal{T}$-rational series s_k such that $g_{s_k}(n) = \Theta(n^{1/k})$. Thus, the hierarchy is strict though it is not proven that all series have an asymptotic growth of this kind.

In the present work, we study the maximum function of rational formal series in two commuting variables a, b with coefficients in $\mathbb{R}_+$. As far as we know, the general problem of characterizing the order of magnitude of $g_S(n)$ for such a series S is still open, though some contributions can be found in the literature [3]. We prove the following

Theorem. *For any positive $k \in \mathbb{N}$ and any primitive $\mathbb{R}_+$-rational formal series r in the noncommutative variables a, b, if S is the commutative image of r^k then, for some $\lambda > 0$, its maximum function satisfies the relation*

$$g_S(n) = \begin{cases} \Theta(n^{k-(3/2)}\lambda^n) & \textit{if } r \textit{ is not degenerate} \\ \Theta(n^{k-1}\lambda^n) & \textit{otherwise} \end{cases}$$

This result is obtained by studying the limit distribution of the discrete random variables $\{X_n\}$ naturally associated with the series S: each X_n takes on values only in $\{0, 1, \ldots, n\}$ and $\Pr\{X_n = i\}$ is proportional to the coefficient $(S, a^i b^{n-i})$. We prove that X_n has a Gaussian limit distribution whenever S is a power of a primitive formal series r as described above. Under the same hypothesis we also give a local limit theorem for $\{X_n\}$ based on the notion of symbol-periodicity introduced in [4]. When the symbol-periodicity associated with r is 1 we just obtain a local limit theorem for $\{X_n\}$ in the sense of DeMoivre-Laplace [9]. The material we present also includes a general criterion for local limit laws that is of interest in its own right and could be useful in other contexts. It holds for sequences of discrete random variables with values in a linear progression of a fixed period d included in $\{0, 1, \ldots, n\}$.

The paper is organized as follows: after some preliminaries on rational series and probability theory, in Section 4 we present the new criterion for local limit laws. Then, we recall basic definitions and properties of the stochastic models defined via rational formal series. In Section 6 we give global and local limit theorems for the pattern statistics associated with powers of primitive rational formal series (*power model*). In the last section, we illustrate our main result on the maximum function of rational formal series in two commutative variables.

Due to space constraints, the proofs of Section 4 are omitted and can be found in Appendix for referees' convenience.

2 Preliminaries on Formal Series

Given a monoid $\mathcal{M}$ and a semiring $\mathbb{K}$ we call *formal series* over $\mathcal{M}$ any application $r : \mathcal{M} \longrightarrow \mathbb{K}$, usually denoted as an infinite formal sum $r = \sum_{w \in \mathcal{M}} (r, w)\, w$, which associates with each $w \in \mathcal{M}$ its *coefficient* $(r, w) \in \mathbb{K}$. The set of all series over $\mathcal{M}$ with coefficients in $\mathbb{K}$ is a monoid algebra, provided with the usual operations of sum, product and star restricted to the elements r such that $(r, 1_{\mathcal{M}}) = 0_{\mathbb{K}}$. These operations are called rational operations. A series r is called *rational* if it belongs to the smallest set closed under rational operations, containing the series 0 and all the series χ_w, for $w \in \mathcal{M}$, such that $(\chi_w, w) = 1$ and $(\chi_w, x) = 0$ for each $x \neq w$.

In particular, in this work we fix the alphabet $\{a, b\}$ and consider the formal series over the free monoid $\{a, b\}^*$ or the free commutative monoid $\{a, b\}^{\otimes}$ with coefficients in the semiring $\mathbb{R}_+$ of nonnegative real numbers. In the former case, the family of all formal series is denoted by $\mathbb{R}_+\langle\langle a, b\rangle\rangle$. By Kleene's Theorem [16, 15], we know that every rational $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$ admits a *linear representation* over $\{a, b\}$, i.e. a triple (ξ, μ, η) such that, for some integer $m > 0$, ξ and η are (non-null column) vectors in $\mathbb{R}_+^m$ and $\mu : \{a, b\}^* \longrightarrow \mathbb{R}_+^{m \times m}$ is a monoid morphism, satisfying $(r, w) = \xi^T \mu(w)\, \eta$ for each $w \in \{a, b\}^*$. We say that m is the *size* of the representation.

Observe that considering such a triple is equivalent to defining a (weighted) nondeterministic finite automaton over the alphabet $\{a, b\}$, where the state set is given by $\{1, 2, \ldots, m\}$ and the transitions, the initial and the final states are assigned weights in $\mathbb{R}_+$ by μ, ξ and η respectively.

Analogously, the family of all formal series over $\{a, b\}^{\otimes}$ with coefficients in $\mathbb{R}_+$ is denoted by $\mathbb{R}_+[[a, b]]$. In this case, any element of $\{a, b\}^{\otimes}$ is represented in the form $a^i b^j$. The canonical morphism $\varphi : \{a, b\}^* \to \{a, b\}^{\otimes}$, associating with each $w \in \{a, b\}^*$ the monomial $a^i b^j$ where $i = |w|_a$ and $j = |w|_b$, extends to the semiring of formal series: for every $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$

$$(\varphi(r), a^i b^j) \;=\; \sum_{|x|_a = i, |x|_b = j} (r, x)\ .$$

Notice that, since φ is a morphism, for every rational series $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$, the commutative image $\varphi(r)$ is rational in $\mathbb{R}_+[[a, b]]$.

For our purpose, a subset of rational series of particular interest is defined by the so-called primitive linear representations. We recall that a matrix $M \in \mathbb{R}_+^{m \times m}$ is *primitive* if for some $k \in \mathbb{N}$ all entries of M^k are strictly positive. The main property of these matrices is given by the Perron–Frobenius Theorem, stating that any primitive matrix M has a unique eigenvalue λ of largest modulus, which is real positive (see for instance [18, Chapter 1]). Such a λ is called the Perron-Frobenius eigenvalue of M.

Thus, a linear representation (ξ, μ, η) defined over a set of generators $\{a, b\}$ is called primitive if $\mu(a) + \mu(b)$ is a primitive matrix. We also say that a series $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$ is *primitive* if it is rational and admits a primitive linear representation. Moreover, we say that (ξ, μ, η) is *degenerate* if $\mu(\sigma) = 0$ for some $\sigma \in \{a, b\}$. Analogously, we say that $r \in R_+\langle\langle a, b\rangle\rangle$ is degenerate if, for some $\sigma \in \{a, b\}$, $(r, w) \neq 0$ implies $w \in \{\sigma\}^*$.

Given a monoid $\mathcal{M}$, let us assume that a length function $|\cdot| : \mathcal{M} \longrightarrow \mathbb{N}$ is well-defined for $\mathcal{M}$ such that the set $\{x \in \mathcal{M} : |x| = n\}$ is finite for each $n \in \mathbb{N}$. Then, for any formal series $s : \mathcal{M} \longrightarrow \mathbb{C}$ we define the *maximum function* $g_s : \mathbb{N} \longrightarrow \mathbb{R}_+$ as

$$g_s(n) = \max\{|(s, x)| \ : \ x \in \mathcal{M}, |x| = n\} \qquad \text{(for every } n \in \mathbb{N}\text{)}.$$

In the following $|\cdot|$ will denote both the modulus of a complex number and the length of a word: the meaning will be clear from the context.

Our main results concern the order of magnitude of $g_r(n)$ for some $r \in \mathbb{R}_+[[a, b]]$. To state them precisely we use the symbol Θ with the standard meaning: given two sequences $\{f_n\}, \{g_n\} \subseteq \mathbb{R}_+$, the equality $g_n = \Theta(f_n)$ means that for some pair of positive constant c_1, c_2, the relation $c_1 f_n \leq g_n \leq c_2 f_n$ holds for any n large enough.

3 Convergence in Distribution

Let X be a random variable (r.v.) with values in a set $\{x_0, x_1, \ldots, x_k, \ldots\}$ of real numbers and set $p_k = \Pr\{X = x_k\}$, for every $k \in \mathbb{N}$. We denote by F_X its distribution function, i.e. $F_X(\tau) = \Pr\{X \leq \tau\}$ for every $\tau \in \mathbb{R}$. If the set of indices $\{k \mid p_k \neq 0\}$ is finite we can consider the moment generating function of X, given by

$$\Psi_X(z) = \mathbb{E}(e^{zX}) = \sum_{k\in\mathbb{N}} p_k e^{zx_k} \ ,$$

which is well-defined for every $z \in \mathbb{C}$. This function can be used to compute the first two moments of X, since $\mathbb{E}(X) = \Psi_X'(0)$ and $\mathbb{E}(X^2) = \Psi_X''(0)$, and to prove convergence in distribution. We recall that, given a sequence of random variables $\{X_n\}_n$ and a random variable X, X_n converges to X *in distribution* (or *in law*) if $\lim_{n\to\infty} F_{X_n}(\tau) = F_X(\tau)$ for every point $\tau \in \mathbb{R}$ of continuity for F_X. It is well-known that if Ψ_{X_n} and Ψ_X are defined all over $\mathbb{C}$ and $\Psi_{X_n}(z)$ tends to $\Psi_X(z)$ for every $z \in \mathbb{C}$, then X_n converges to X in distribution [9].

A convenient approach to prove the convergence in law to a Gaussian random variable relies on the so called "quasi-power" theorems introduced in [11] and implicitly used in the previous literature [1] (see also [7]). For our purpose we present the following simple variant of such a theorem.

Theorem 1. *Let $\{X_n\}$ be a sequence of random variables, where each X_n takes values in $\{0, 1, \ldots, n\}$ and let us assume the following conditions:*

C1 *There exist two functions $r(z)$, $y(z)$, both analytic at $z = 0$ where they take the value $r(0) = y(0) = 1$, and a positive constant c, such that for every $|z| < c$*

$$\Psi_{X_n}(z) = r(z) \cdot y(z)^n \left(1 + O(n^{-1})\right); \tag{1}$$

C2 *The constant $\sigma = y''(0) - (y'(0))^2$ is strictly positive (variability condition).*

Also set $\mu = y'(0)$. Then $\frac{X_n - \mu n}{\sqrt{\sigma n}}$ converges in distribution to a normal random variable of mean 0 and variance 1, i.e. for every $x \in \mathbb{R}$

$$\lim_{n \longrightarrow +\infty} Pr\left\{\frac{X_n - \mu n}{\sqrt{\sigma n}} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \,.$$

The main advantage of this theorem, with respect to other classical statements of this kind, is that it does not require any condition of independence concerning the random variables X_n. For instance, the standard central limit theorems assume that each X_n is a partial sum of the form $X_n = \sum_{j \leq n} U_j$, where the U_j's are independent random variables [9].

4 A General Criterion for Local Convergence Laws

Convergence in law of a sequence of r.v.'s $\{X_n\}$ does not yield an approximation of the probability that X_n has a specific value. Theorems concerning approximations for expressions of the form $\Pr\{X_n = x\}$ are usually called *local limit theorems* and often give a stronger property than a traditional convergence in distribution[1]. A typical example is given by the so-called de Moivre-Laplace Local Limit Theorem [9], which intuitively states that, for certain sequences of binomial random variables X_n, up to a factor $\Theta(\sqrt{n})$ the probability that X_n takes on a value x approximates a Gaussian density at x.

In this section we present a general criterion that guarantees, for a sequence of discrete random variables, the existence of a local convergence property of a Gaussian type more general than DeMoivre-Laplace's Theorem mentioned above. In the subsequent section, using such criterion, we show that the same local convergence property holds for certain pattern statistics.

Theorem 2 (Local Limit Criterion). *Let $\{X_n\}$ be a sequence of random variables such that, for some integer $d \geq 1$ and every $n \geq d$, X_n takes on values only in the set*

$$\{x \in \mathbb{N} \mid 0 \leq x \leq n, x \equiv \rho (mod\ d)\} \tag{2}$$

for some integer $0 \leq \rho < d$. Assume that conditions C1 and C2 of Theorem 1 hold true and let μ and σ be the positive constants defined in the same theorem. Moreover assume the following property:

[1] For this reason, theorems showing convergence in distribution of a sequence of r.v.'s are sometimes called global or integral limit theorems.

C3 *For all* $0 < \theta_0 < \pi/d$ $$\lim_{n\to+\infty}\left\{\sqrt{n}\sup_{|\theta|\in[\theta_0,\pi/d]}|\Psi_{X_n}(i\theta)|\right\} = 0$$

Then, as n *grows to* $+\infty$ *the following relation holds uniformly for every* $j = 0, 1, \ldots, n$*:*

$$Pr\{X_n = j\} = \begin{cases} \dfrac{de^{-\frac{(j-\mu n)^2}{2\sigma n}}}{\sqrt{2\pi\sigma n}} \cdot (1 + o(1)) & \text{if } j \equiv \rho \ (\text{mod } d) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Due to space constraints, the proof is omitted. We only note that it extends some ideas used in [3, Section 5] and [4, Theorem 3]. However, the present approach is much more general: we drop any rationality hypothesis on the distribution of the r.v.'s X_n and only rely on conditions C1, C2, C3 together with the assumption that each X_n takes on values only in the set (2).

Observe that $\Psi_{X_n}(i\theta)$ is the so-called characteristic function of X_n, it is periodic of period 2π and it assumes the value 1 at $\theta = 0$. Condition C3 states that, for every constant $0 < \theta_0 < \pi/d$, as n grows to $+\infty$, the value $\Psi_{X_n}(i\theta)$ is of the order $o(n^{-1/2})$ uniformly with respect to $\theta \in [-\pi/d, -\theta_0] \cup [\theta_0, \pi/d]$. Note that ρ may depend on n even if $\rho = \Theta(1)$.

One can easily show that any sequence $\{X_n\}$ of binomial r.v.'s of parameters n and p, where $0 < p < 1$ (i.e. representing the number of successes over n independent trials of probability p), satisfies the hypothesis of the theorem with $d = 1$. In this case (3) corresponds to DeMoivre-Laplace Local Limit Theorem. Thus our general criterion includes the same theorem as a special case.

Relations of the form (3) already appeared in the literature. In particular in [9, Section 43], (3) is proved when X_n is the sum of n independent *lattice* r.v.'s of equal distribution and maximum span d. Note that our theorem is quite general since it does not require any condition of independence for the X_n's.

We also note that, for $d = 1$ a similar criterion for local limit laws has been proposed in [7, Theorem 9.10] where, however, a different condition is assumed, i.e. one requires that the probability generating function $p_n(u)$ of X_n has a certain expansion, for $u \in \mathbb{C}$ belonging to an annulus $1 - \varepsilon \leq |u| \leq 1 + \varepsilon$ ($\varepsilon > 0$), that corresponds to assume an equation of the form (1) for $z \in \mathbb{C}$ such that $|\Re\text{e}(z)| \leq \delta$ (for some $\delta > 0$).

5 Pattern Statistics Over Rational Series

In this section we turn our attention to sequences of random variables defined by means of rational formal series in two non-commuting variables. We recall definitions and properties introduced in [3, 4].

Let us consider the binary alphabet $\{a, b\}$ and, for $n \in \mathbb{N}$, let $\{a, b\}^n$ denote the set of all words of length n in $\{a, b\}^*$. Given a formal series $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$, let n be a positive integer such that $(r, x) \neq 0$ for some $x \in \{a, b\}^n$. Consider the probability space of all words in $\{a,b\}^n$ equipped with the probability measure given by

$$\Pr\{\omega\} = \frac{(r,\omega)}{\sum_{x\in\{a,b\}^n}(r,x)} \qquad (\omega \in \{a,b\}^n). \tag{4}$$

In particular, if r is the characteristic series χ_L of a language $L \subseteq \{a,b\}^*$, then Pr is just the uniform distribution over the set of words on length n in L: $\Pr\{\omega\} = \sharp(L\cap\{a,b\}^n)^{-1}$ if $\omega \in L$, while $\Pr\{\omega\} = 0$ otherwise. We define the random variable $Y_n : \{a,b\}^n \to \{0,1,\ldots,n\}$ such that $Y_n(\omega) = |\omega|_a$ for every $\omega \in \{a,b\}^n$. Then, for every $j = 0,1,\ldots,n$, we have

$$\Pr\{Y_n = j\} = \frac{\sum_{|\omega|=n,|\omega|_a=j}(r,\omega)}{\sum_{x\in\{a,b\}^n}(r,x)} . \tag{5}$$

For sake of brevity, we say that Y_n *counts the occurrences of* a *in the stochastic model defined by* r. If $r = \chi_L$ for some $L \subseteq \{a,b\}^*$, then Y_n represents the number of occurrences of a in a word chosen at random in $L \cap \{a,b\}^n$ under uniform distribution.

A useful tool to study the distribution of the pattern statistics Y_n is given by certain generating functions associated with formal series. Given $r \in \mathbb{R}_+\langle\langle a,b\rangle\rangle$, for every $n,j \in \mathbb{N}$ let $r_{n,j}$ be the coefficient of $a^j b^{n-j}$ in the commutative image $\varphi(r)$ of r, i.e.

$$r_{n,j} = (\varphi(r), a^j b^{n-j}) = \sum_{|x|=n,|x|_a=j} (r,x) \ .$$

Then, we define the function $r_n(z)$ and the generating function $\mathbf{r}(z,w)$ by

$$r_n(z) = \sum_{j=0}^{n} r_{n,j}\, e^{jz} \qquad \text{and} \qquad \mathbf{r}(z,w) = \sum_{n=0}^{+\infty} r_n(z)\, w^n = \sum_{n=0}^{+\infty}\sum_{j=0}^{n} r_{n,j}\, e^{jz}\, w^n$$

where z and w are complex variables. Thus, from the definition of $r_{n,j}$ and from equation (5) we have

$$\Pr\{Y_n = j\} = \frac{r_{n,j}}{r_n(0)} , \qquad \Psi_{Y_n}(z) = \frac{r_n(z)}{r_n(0)} . \tag{6}$$

Moreover, we remark that the relation between a series r and its generating function $\mathbf{r}(z,w)$ can be expressed in terms of a semiring morphism . Denoting by $\Sigma^{\oplus}$ the free commutative monoid over the alphabet Σ, consider the monoid morphism $\mathcal{H} : \{a,b\}^* \longrightarrow \{e^z, w\}^{\oplus}$ defined by setting $\mathcal{H}(a) = e^z w$ and $\mathcal{H}(b) = w$. Then, such a map extends to a semiring morphism from $\mathbb{R}_+\langle\langle a,b\rangle\rangle$ to $\mathbb{R}_+[[e^z,w]]$ so that

$$\mathcal{H}(r) = \mathbf{r}(z,w) \tag{7}$$

for every $r \in \mathbb{R}_+\langle\langle a,b\rangle\rangle$. This property translates arithmetic relations among formal series into analogous relations among the corresponding generating functions.

When r is rational, the probability spaces given by (4) define a stochastic model (called *rational* stochastic model) of interest for the analysis of pattern statistics. A typical goal in that context is to estimate the limit distribution of the number of occurrences of patterns in a word of length n generated at random according to a given probabilistic model (usually a Markovian process [13]). In the rational model, the pattern is reduced to a single letter a. However, the analysis of Y_n in such a model includes as a particular case the study of the frequency of occurrences of regular patterns in words generated at random by a Markovian process [3, Section 2.1].

The limit distribution of Y_n is studied in [3] in the global sense and in [4] in the local sense, assuming that r admits a primitive linear representation (ξ, μ, η). Set $A = \mu(a)$ and $B = \mu(b)$. Then it is easy to see that in this case

$$\mathbf{r}(z, w) = \xi^T \left(I - w(Ae^z + B)\right)^{-1} \eta \ . \tag{8}$$

It turns out that Y_n has a Gaussian limit distribution [3, Theorem 4], and this extends a similar result, earlier presented in [13] for pattern statistics in a Markovian model. A local limit property of the form (3) also holds, where d is the so-called x-period of $Ax + B$ [4, Theorem 4].

We recall (see [4]) that given a polynomial $f = \sum_k f_k x^k \in \mathbb{R}_+[x]$, the *$x$-period* of f is defined as the value $D(f) = \mathrm{GCD}\{|h - k| : f_h \neq 0 \neq f_k\}$, where we assume $\mathrm{GCD}(\{0\}) = \mathrm{GCD}(\emptyset) = +\infty$. For any matrix $M \in \mathbb{R}_+[x]^{m\times m}$ and any index $q \in \{1, 2, \ldots, m\}$, the x-period of q is the value $d(q) = \mathrm{GCD}$ $\{D((M^n)_{qq}) \mid n \geq 0\}$, assuming that every non-zero element in $\mathbb{N}\cup\{+\infty\}$ divides $+\infty$. It turns out that, for every matrix $M(x) \in \mathbb{R}_+[x]^{m\times m}$ such that $M(1)$ is primitive, all indices have the same x-period, which is called the *x-period of M*.

We conclude this section presenting some results proved in [4] we use in the following section. They give interesting properties of the x-period of a matrix of the form $M_1x + M_2$, where $M_1 \neq 0 \neq M_2$ and $M_1 + M_2$ is primitive.

Proposition 3. *Consider two non-null matrices $M_1, M_2 \in \mathbb{R}_+^{m\times m}$, assume that $M_1 + M_2$ is primitive and let d be the x-period of $M_1x + M_2$. Then d is finite and there exists an integer $0 \leq \gamma < d$ such that for any pair of indices $p, q \in \{1, 2, \cdots , m\}$ and for any integer n large enough, we have $D({(M_1x + M_2)^n}_{pq}) \equiv \gamma n + \delta_{pq}$ (mod d) for a suitable integer $0 \leq \delta_{pq} < d$ independent of n.*

Proposition 4. *Consider two non-null matrices $M_1, M_2 \in \mathbb{R}_+^{m\times m}$, assume that $M_1 + M_2$ is primitive and denote by λ its Perron-Frobenius eigenvalue. Moreover let d be the x-period of M_1x+M_2. Then, for any real number $\theta \neq 2k\pi/d$ ($k \in \mathbb{Z}$), all eigenvalues of $M_1e^{i\theta} + M_2$ are in modulus smaller than λ.*

6 Pattern Statistics in the Power Model

In this section, we consider a stochastic model defined by the *power* of any primitive rational series (note that in this case the model is not primitive anymore) and we study the central and local behaviour of the associated pattern statistics Y_n. The results we obtain extend the analysis developed in [3] and [4] concerning the primitive rational stochastic models. They also extend some results presented in [5], where the (global) limit distribution of Y_n is determined whenever r is the product of two primitive formal series.

Theorem 5. *For any positive integer k and any primitive nondegenerate $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$, let s be defined by $s = r^k$ and let Y_n be the random variables counting the occurrences of a in the model defined by s. Then the following properties hold true.*

T1 *There exist two constants α and β, satisfying $0 < \alpha$ and $0 < \beta < 1$, such that $\frac{Y_n - \beta n}{\sqrt{\alpha n}}$ converges in distribution to a normal random variable of mean value 0 and variance 1.*

T2 *If (ξ, μ, η) is a primitive linear representation for r and d is the x-period of $\mu(a)x + \mu(b)$, then there exist d many functions $C_i : \mathbb{N} \longrightarrow \mathbb{R}_+$, $i = 0, 1, \ldots, d-1$, such that $\sum_i C_i(n) = 1$ for every $n \in \mathbb{N}$ and further, as n grows to $+\infty$, the relation*

$$Pr\{Y_n = j\} = \frac{d\, C_{\langle j\rangle_d}(n)}{\sqrt{2\pi\alpha n}}\, e^{-\frac{(j-\beta n)^2}{2\alpha n}} \cdot (1 + o(1)) \tag{9}$$

holds uniformly for every $j = 0, 1, \ldots, n$ (here $\langle j\rangle_d = j - \lfloor j/d \rfloor d$).

We note that in case $k = 1$ statement T1 coincides with [3, Theorem 4] and statement T2 corresponds to [4, Theorem 4].

We split the proof of the previous theorem in two separate parts and we use the criteria presented in Theorem 1 and in Theorem 2. We still use the notation introduced in the previous section: set $A = \mu(a)$, $B = \mu(b)$, $M = A + B$ and denote by λ the Perron Frobenius eigenvalue of the primitive matrix M.

Proof of T1. Since $s = r^k$, by applying the morphism $\mathcal{H}$ defined in (7) we get

$$\mathbf{s}(z, w) = \mathbf{r}(z, w)^k \ .$$

From equation (8), since $A + B$ is primitive and both A and B are non-null, one can show [3, Section 4] that near the point $(0, \lambda^{-1})$ the function $\mathbf{r}(z, w)$ admits a Laurent expansion of the form

$$\mathbf{r}(z, w) = \frac{R(z)}{1 - u(z)w} + \mathrm{O}(1)$$

where $R(z)$ and $u(z)$ are complex functions, non-null and analytic at $z = 0$. Moreover, the constants $\alpha = (u''(0) - \beta^2)/\lambda$ and $\beta = u'(0)/\lambda$ are strictly positive. We also recall that α and β can be expressed as function of the matrix M and in particular of its eigenvectors.

As a consequence, in a neighbourhood of $(0, \lambda^{-1})$ we have

$$\mathbf{s}(z, w) = \left(\frac{R(z)}{1 - u(z)w}\right)^k + \mathrm{O}\left(\frac{1}{1 - u(z)w}\right)^{k-1}$$

and hence the associated sequence is of the form

$$s_n(z) = R(z)^k \binom{n + k - 1}{k - 1} u(z)^n + \mathrm{O}\left(n^{k-2} u(z)^n\right).$$

Now, by the definition of our stochastic model, the characteristic function of Y_n is given by $\Psi_{Y_n}(z) = s_n(z)/s_n(0)$ and hence, in a neighbourhood of $z = 0$, it has an expansion of the form

$$\Psi_{Y_n}(z) = \frac{s_n(z)}{s_n(0)} = \left(\frac{R(z)}{R(0)}\right)^k \cdot \left(\frac{u(z)}{\lambda}\right)^n \cdot \left(1 + \mathrm{O}(n^{-1})\right) \ .$$

As a consequence, both conditions of Theorem 1 hold with $\mu = \beta$ and $\sigma = \alpha$ and this proves the result. □

Proof of T2 (Outline). For every $p, q \in \{1, 2, \dots, m\}$, let $r^{(pq)}$ be the series defined by the linear representation $(\xi_p e_p,\ \mu,\ \eta_q e_q)$, where e_i is the characteristic array of entry i. Then $r = \sum_{p,q} r^{(pq)}$. Thus, since $s = r^k$, we have

$$s = \sum_{*} r^{(p_1q_1)} \cdot r^{(p_2q_2)} \cdots r^{(p_kq_k)} \tag{10}$$

where the sum is over all sequences $\ell = p_1q_1p_2q_2 \cdots p_kq_k \in \{1, 2, \dots, m\}^{2k}$. For sake of brevity, for every such ℓ, let $r^{(\ell)}$ be the series

$$r^{(\ell)} = r^{(p_1q_1)} \cdot r^{(p_2q_2)} \cdots r^{(p_kq_k)} \tag{11}$$

and let $Y_n^{(\ell)}$ denote the r.v. counting the occurrences of a in the model defined by $r^{(\ell)}$. Then, the primitivity hypothesis allows one to prove that the relation

$$\Pr\{Y_n = j\} = \sum_{*} C_\ell \Pr\{Y_n^{(\ell)} = j\} + \mathrm{O}(n^{-1})$$

holds for every $j \in \{0, 1, \dots, n\}$, where C_ℓ is a non-negative constant for each ℓ and $\sum_\ell C_\ell = 1$.

Thus, to determine the local behaviour of $\{Y_n\}$, we first study $\{Y_n^{(\ell)}\}$ for any ℓ such that $r^{(\ell)} \neq 0$. Indeed, by the previous relation, it is sufficient to prove that the equation

$$\Pr\{Y_n^{(\ell)} = j\} = \begin{cases} \dfrac{d\, e^{-\frac{(j-\beta n)^2}{2\alpha n}}}{\sqrt{2\pi\alpha n}} \cdot (1 + o(1)) & \text{if } j \equiv \rho_\ell \pmod d \\ 0 & \text{otherwise} \end{cases}$$

holds uniformly for every $j = 0, 1, \dots, n$, where α and β are defined as in T1, while ρ_ℓ is an integer (possibly depending on n) such that $0 \leq \rho_\ell < d$ (in particular $C_i(n) = \sum_{\rho_\ell = i} C_\ell$ for each i). To this aim, we simply have to show that, for every $n \in \mathbb{N}$, $Y_n^{(\ell)}$ satisfies the hypotheses of Theorem 2.

First, one can prove that $Y_n^{(\ell)}$ takes on values only in a set of the form (2). This is a consequence of the fact that the values $r_{n,j}^{(\ell)}$ are given by the convolutions

$$r_{n,j}^{(\ell)} = \sum_{\substack{n_1+n_2+\cdots+n_k=n \\ j_1+j_2+\cdots+j_k=j}} \prod r_{n_i,j_i}^{(p_iq_i)}$$

together with Proposition 3.

As far as condition C1 and C2 are concerned, we can argue (with obvious changes) as in the proof of T1 and observe that the two constants α and β are the same for all series $r^{(\ell)} \neq 0$, since they depend on the matrices A and B (not on the initial and final arrays).

To prove condition C3 let us consider the generating function of $\{r_n^{(\ell)}(z)\}$, obtained by applying the morphism $\mathcal{H}$ to (11):

$$\mathbf{r}^{(\ell)}(z,w) = \prod_{j=1}^{k} \xi_{p_j} (I - w(Ae^z + B))^{-1}{}_{p_j q_j}\ \eta_{q_j} \ .$$

Applying Proposition 4, one can prove that all singularities of $\mathbf{r}^{(\ell)}(i\theta, w)$ are in modulus greater than λ^{-1}. Hence, by Cauchy's integral formula, for any arbitrary $\theta_0 \in (0, \pi/d)$ we can choose $0 < \tau < \lambda$ such that the associated sequence $\{{r_n}^{(\ell)}(i\theta)\}$ is bounded by $\mathrm{O}(\tau^n)$ for every $|\theta| \in [\theta_0, \pi/d]$. Analogously, one gets ${r_n}^{(\ell)}(0) = \Theta(n^{k-1}\lambda^n)$ and this also implies $\Psi_{Y_n^{(\ell)}}(i\theta) = {r_n}^{(\ell)}(i\theta)/{r_n}^{(\ell)}(0) = \mathrm{O}(\epsilon^n)$ for some $0 < \epsilon < 1$. This yields condition C3 and concludes the proof. □

As a final remark, we note that Theorem 5 cannot be extended to all rational models because the "quasi-power" condition C1 does not hold for $\Psi_{Y_n}(z)$ in the general case. In fact, a large variety of limit distributions for Y_n are obtained in rational models that have two primitive components [5] and more complicated behaviours occur in multicomponent models [8].

7 Estimate of the Maximum Coefficients

The result proved in the last section provides us an asymptotic evaluation for the maximum coefficients of formal series in commuting variables that are commutative image of powers of primitive rational formal series.

Corollary 6. *For any $k \in \mathbb{N}$, $k \neq 0$ and any primitive series $r \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$, let $s = r^k$ and consider its commutative image $S = \varphi(s) \in \mathbb{R}_+[[a, b]]$. Then, for some $\lambda > 0$, the maximum function of S satisfies the relation*

$$g_S(n) = \begin{cases} \Theta(n^{k-(3/2)}\lambda^n) & \text{if } r \text{ is not degenerate} \\ \Theta(n^{k-1}\lambda^n) & \text{otherwise} \end{cases}$$

Proof. Let (ξ, μ, η) be a primitive linear representation of r and let λ be the Perron-Frobenius eigenvalue of $\mu(a) + \mu(b)$. To determine $g_S(n)$ we have to compute the maximum of the values $s_{n,j} = (S, a^j b^{n-j})$ for $j = 0, 1, \ldots, n$.

First consider the case when r is not degenerate. Then, let Y_n count the occurrences of a in the model defined by $s = r^k$ and recall that $\Pr(Y_n = j) = s_{n,j}/s_n(0)$. Now, reasoning as above, we have $s_n(0) = \Theta(n^{k-1}\lambda^n)$ and by Theorem 5, the set of probabilities $\{\Pr(Y_n = j) \mid j = 0, 1, \ldots, n\}$ has the maximum at some integer $j \in [\beta n - d, \beta n + d]$, where it takes on a value of the order $\Theta(n^{-1/2})$ and this proves the first equation.

On the other hand, if r is degenerate, then either $\mu(a) = 0$ or $\mu(b) = 0$. In the first case, all $r_{n,j}$ vanish except $r_{n,0}$ which is of the order $\Theta(\lambda^n)$. Hence for every n, the value $\max_j\{s_{n,j}\} = s_n(0)$ is given by the k-th convolution of $r_{n,0}$, which is of the order $\Theta(n^{k-1}\lambda^n)$. The case $\mu(b) = 0$ is similar. □

Example. Consider the rational function $(1 - a - b)^{-k}$. Its Taylor expansion near the origin yields the series

$$S = \sum_{n=0}^{+\infty} \binom{n+k-1}{k-1} \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$$

By direct computation, one can verify that

$$g_S(n) = \binom{n+k-1}{k-1}\binom{n}{\lfloor n/2 \rfloor} = \Theta(n^{k-3/2}2^n) \ .$$

In fact, it turns out that $S = \varphi(r^k)$ where $r = \chi_{\{a,b\}^*} \in \mathbb{R}_+\langle\langle a, b\rangle\rangle$. □

Even though the statement of Theorem 5 cannot be extended to all rational models, we believe that the property given in Corollary 6 well represents the asymptotic behaviour of maximum coefficients of all rational formal series in two commutative variables. We actually think that a similar result holds for all rational formal series in commutative variables. More precisely, let us introduce the symbol $\widehat{\Theta}$ with the following meaning: for any pair of sequences $\{f_n\}, \{g_n\} \subseteq \mathbb{R}_+$, we have $g_n = \widehat{\Theta}(f_n)$ if $g_n = \mathrm{O}(f_n)$ and $g_{n_j} = \Theta(f_{n_j})$ for some monotone strictly increasing sequence $\{n_j\} \subseteq \mathbb{N}$. Then we conjecture that the asymptotic behaviour of the maximum function of every rational formal series $t \in \mathbb{R}_+[[\sigma_1, \cdots, \sigma_\ell]]$, is of the form

$$g_t(n) = \widehat{\Theta}\left(n^{k/2}\lambda^n\right)$$

for some integer $k \geq 1 - \ell$ and some $\lambda \in \mathbb{R}_+$.

References

1. E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory*, 15:91–111, 1973.
2. J. Berstel and C. Reutenauer. *Rational series and their languages*, Springer-Verlag, New York - Heidelberg - Berlin, 1988.
3. A. Bertoni, C. Choffrut, M. Goldwurm, and V. Lonati. On the number of occurrences of a symbol in words of regular languages. *Theoretical Computer Science*, 302(1-3):431–456, 2003.
4. A. Bertoni, C. Choffrut, M. Goldwurm, and V. Lonati. Local limit distributions in pattern statistics: beyond the Markovian models. *Proceedings 21st S.T.A.C.S.*, V. Diekert and M. Habib editors, Lecture Notes in Computer Science, vol. n. 2996, Springer, 2004, 117–128.

5. D. de Falco, M. Goldwurm, and V. Lonati. Frequency of symbol occurrences in bicomponent stochastic models. Rapporto Interno n. 299-04, D.S.I. Università degli Studi di Milano, April 2004, to appear in *Theoretical Computer Science*. Preliminary version in *Proceedings of the 7th D.L.T. Conference*, Z. Esik and Z. Fülop editors, Lecture Notes in Computer Science, vol. n. 2710, Springer, 2003, 242–253.
6. V. Diekert and G. Rozenberg (editors). *The book of traces.* World Scientific, 1995.
7. P. Flajolet and R. Sedgewick. The average case analysis of algorithms: multivariate asymptotics and limit distributions. *Rapport de recherche* n. 3162, INRIA Rocquencourt, May 1997.
8. M. Goldwurm, and V. Lonati. Pattern occurrences in multicomponent models. Manuscript, September 2004, submitted for pubblication.
9. B.V. Gnedenko. *The theory of probability* (translated by G. Yankovsky). Mir Publishers - Moscow, 1976.
10. P. Henrici. *Elements of numerical analysis*, John Wiley, 1964.
11. H.K. Hwang. *Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques.* Ph.D. Dissertation, École polytechnique, Palaiseau, France, 1994.
12. W. Kuich. Finite automata and ambiguity. Technical Report n.253, I.I.G. University of Graz, 1988.
13. P. Nicodeme, B. Salvy, and P. Flajolet. Motif statistics. *Theoretical Computer Science*, 287(2):593–617, 2002.
14. C. Reutenauer. *Propriétés arithmétiques et topologiques de séries rationnelles en variables non commutatives*, These Sc. Maths, Doctorat troisieme cycle, Université Paris VI, 1977.
15. J. Sakarovitch. *Eléments de théorie des automates*, Vuibert Informatique, 2003.
16. A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series*, Springer-Verlag, 1978.
17. M.-P. Schützenberger. Finite counting automata. *Information and Control*, 5:91–107, 1962.
18. E. Seneta. *Non-negative matrices and Markov chains*, Springer–Verlag, New York Heidelberg Berlin, 1981.
19. I. Simon. The nondeterministic complexity of a finite automaton. In *Mots*, M. Lothaire editor, Lecture Notes in Mathematics, Hermes, Paris, 1990, 384–400.
20. A. Weber and H. Seidl. On the degree of ambiguity of finite automata. *Theoretical Computer Science*, 88:325–349, 1991.
21. K. Wich. Exponential ambiguity of context-free grammars. In *Proceedings of the 4th DLT*, G. Rozenberg and W. Thomas editors. World Scientific, Singapore, 2000, 125–138.
22. K. Wich. Sublinear ambiguity. In *Proceedings of the 25th MFCS*, M. Nielsen and B. Rovan editors. Lecture Notes in Computer Science, vol. n.1893, Springer-Verlag, 2000, 690–698.