Codes and noncommutative stochastic matrices

Sylvain Lavallée, Dominique Perrin,
Vladimir Retakh, and Christophe Reutenauer

Abstract. Given a matrix over a skew field fixing the column $^t(1,\ldots,1)$, we give formulas for a row vector fixed by this matrix. The same techniques are applied to give noncommutative extensions of probabilistic properties of codes.

Mathematics Subject Classification (2010). 16Z05, 05E15.

Keywords. Stochastic matrices, codes, eigenvectors, quasi-determinants.

1. Introduction

By a *noncommutative stochastic matrix* we mean a square matrix $S = (a_{ij})_{1 \le i,j \le n}$ over a skew field, satisfying $\sum_{j=1}^{n} a_{ij} = 1$ for any i; in other words, the row-sums are all equal to 1. Equivalently the vector $^t(1,\ldots,1)$ is fixed by S. We are answering here the following problem: find a row vector fixed by S.

In the commutative case, formulas are known. They occur in probability theory, where this problem is relevant. Indeed, it amounts to finding the stationary distribution of the finite Markov chain whose transition matrix is S. See Appendix 1 for details.

But this problem may also be considered as a problem of noncommutative linear algebra: given a square matrix over some skew field, which has a given column as eigenvector for some eigenvalue, find a corresponding row vector. It is easy to reduce this general problem to the previous one, where the eigenvector is $^t(1,\ldots,1)$ and the eigenvalue is 1.

In order to give formulas, which necessarily involve inverses of elements of the skew field and thus may be undefined, we take a *generic noncommutative stochastic matrix*: this is the matrix (a_{ij}) of noncommuting variables a_{ij} subject only to the condition that this matrix fixes $^t(1,\ldots,1)$.

We seek now a row vector fixed by the matrix. We work in the *free field* generated by these variables (in the sense of Paul Cohn), which we call the *stochastic free field*. The formula giving the row eigenvector uses the theory of variable-length codes. Considering the complete digraph on the set $\{1,\ldots,n\}$, let M_i be the set of paths from i to i. This is a free monoid and its basis C_i is a prefix code. Let P_i be the set of proper prefixes of C_i, that is, the paths starting from i and not passing through i

again. We identify P_i with the noncommutative power series which is equal to the sum of all the words in P_i and we still denote this series by P_i. Then we show that the elements P_i^{-1} can be evaluated in the stochastic free field and that the vector $(P_1^{-1}, \ldots, P_n^{-1})$ is fixed by our matrix; moreover, the P_i^{-1} sum to 1, hence they form a kind of noncommutative limiting probability. See Theorem 1 and Example 1 to have a flavor of the result.

The second part of the article deals with general variable-length codes, not necessarily prefix. One motivation is the fact that the proofs are quite similar. The other motivation is that we obtain noncommutative generalization of well-known probabilistic results in the theory of codes, mostly due to Schützenberger (see [4] and [5]), who generalized the recurrent events of Feller.

In Appendix 1, we review the commutative case. In Appendix 2, we show how the theory of quasideterminants may be used to obtain our results on noncommutative stochastic matrices.

2. Basics

2.1. Languages and codes. A *language* is a subset of a free monoid A^*; the latter is generated by the *alphabet A*. A language is *rational* if it is obtained from finite languages by the operations (called rational) *union*, *product* (concatenation) and *star*. The product of two languages $L_1 L_2$ is $\{w_1 w_2 \mid w_1 \in L_1, \ w_2 \in L_2\}$, and the star of L is $L^* = \{w_1 \ldots w_n \mid w_i \in L, \ n \geq 0\} = \bigcup_{n \geq 0} L^n$.

It is well known that rational languages may be obtained by using only *unambiguous rational operations*; these are: disjoint union, unambiguous product (meaning that if $w \in L_1 L_2$, then w has a unique factorization $w = w_1 w_2$, $w_i \in L_i$) and the star L^* restricted to languages which are *codes*, or equivalently bases of free submonoids of A^*.

2.2. Series. By a *series* we mean an element of the $\mathbb{Q}$-algebra of noncommutative series $\mathbb{Q}\langle\!\langle A \rangle\!\rangle$, where A is a set of noncommuting variables. A *rational* series is an element of the least subalgebra of $\mathbb{Q}\langle\!\langle A \rangle\!\rangle$, which contains the $\mathbb{Q}$-algebra of noncommutative polynomials $\mathbb{Q}\langle A \rangle$, and which is closed under the operation

$$S \mapsto S^* = \sum_{n=0}^{\infty} S^n = (1 - S)^{-1},$$

which is defined if S has zero constant term. We denote by $\mathbb{Q}\langle\!\langle A \rangle\!\rangle^{\mathrm{rat}}$ the $\mathbb{Q}$-algebra of rational series. Each such series has a $*$-*rational expression*: this is a well-formed

expression involving scalars, letters (elements of A), products and star operations, the latter restricted to series with 0 constant term. We say that a $*$-rational expression is *positive* if the scalars involved are all ≥ 0.

Let L be a rational language. Since L may be obtained by unambiguous rational expressions, it follows that its *characteristic series* $\sum_{w \in L} w \in \mathbb{Q}\langle\!\langle A \rangle\!\rangle$ is rational. We shall identify a language and its characteristic series. For all this, see [6] or [7].

2.3. Free fields. The ring $\mathbb{Q}\langle\!\langle A \rangle\!\rangle^{\mathrm{rat}}$ contains $\mathbb{Q}\langle A \rangle$; it is not a skew field. There exist skew fields containing $\mathbb{Q}\langle A \rangle$. Among them is the so called *free field*. We denote it $\mathcal{F}$. It is generated by $\mathbb{Q}\langle A \rangle$ and has the following universal property (which characterizes it): for each $\mathbb{Q}$-algebra homomorphism $\mu \colon \mathbb{Q}\langle A \rangle \to D$, where D is a skew field, there exists a $\mathbb{Q}$-subalgebra O_μ of $\mathcal{F}$ and a homomorphism $\bar{\mu} \colon O_\mu \to D$ extending μ and such that

$$f \in O_\mu, \ \bar{\mu} f \neq 0 \implies f^{-1} \in O_\mu.$$

The free field $\mathcal{F}$ is also characterized by the following property: say that a square matrix $M \in \mathbb{Q}\langle A \rangle^{n \times n}$ is *full* if there exists no factorization $M = PQ$, $P \in \mathbb{Q}\langle A \rangle^{n \times r}$, $Q \in \mathbb{Q}\langle A \rangle^{r \times n}$, with $r < n$. Then the square matrices over $\mathbb{Q}\langle A \rangle$ which are invertible in $\mathcal{F}$ are exactly the full matrices. See [10].

By a *rational expression* over $\mathbb{Q}\langle A \rangle$ we mean a well-formed expression involving elements of $\mathbb{Q}\langle A \rangle$ and the operations sum, product and inversion. Such an expression can be naturally evaluated in the free field $\mathcal{F}$, provided one never inverts 0. For example, $(a + b^{-1})^{-1}$ can be evaluated in the free field whereas $(ab - (b^{-1}a^{-1})^{-1})^{-1}$ cannot. If a rational expression can be evaluated in the free field, we say it is *evaluable in $\mathcal{F}$*.

We shall consider also rational expressions over any skew field D and say that such an expression is *evaluable in D* if it can be evaluated without inversion of 0. If the elements of D appearing in the rational expression are actually in a subring R of D, we say that the expression is *over R*.

There is a canonical embedding of $\mathbb{Q}\langle\!\langle A \rangle\!\rangle^{\mathrm{rat}}$ into $\mathcal{F}$, which can be seen as follows: let S be any rational series; it has a $*$-rational expression; replace in it the operation T^* by $(1 - T)^{-1}$; then one obtains a rational expression in $\mathcal{F}$, which is evaluable in $\mathcal{F}$ and represents the image of S under the embedding $\mathbb{Q}\langle\!\langle A \rangle\!\rangle^{\mathrm{rat}} \hookrightarrow \mathcal{F}$. Thus, each rational language and each rational series is naturally an element of the free field. In this way, each $*$-rational expression is equivalent to an evaluable rational expression over $\mathbb{Q}\langle A \rangle$. See [13].

In the sequel, we use the notation x^* for $(1 - x)^{-1}$, when x is in a ring and $1 - x$ is invertible.

2.4. The derivation λ of the free field. There is a unique derivation λ of $\mathbb{Q}\langle A \rangle$ such that $\lambda(a) = a$ for any $a \in A$. It maps each word $w \in A^*$ onto $|w|w$, where $|w|$ is the length of w. It has a unique extension to the free field $\mathcal{F}$, which we still denote λ. Indeed, this follows from Theorem 7.5.17, p. 451 in [12]; see also [11].

2.5. $D[t]$ and other rings. Let D be a skew field and t be a central variable. It is well known that the ring of polynomials in t over D is a left and right Euclidean ring, and thus an Ore ring. It has a field of fractions $D(t)$, each element of which is of the form PQ^{-1} and $R^{-1}S$ for suitable P, Q, R, S in $D[t]$. The ring of series in t over D is denoted $D[[t]]$. It is contained in the skew field of Laurent series $D((t))$. The latter also contains canonically $D(t)$, and we may identify $D(t)$ with a subset of $D((t))$. A series $S \in D[[t]]$ is called *rational* if S is in $D(t)$. The ring of rational series is denoted by $D[[t]]^{\text{rat}}$. Thus $D[[t]]^{\text{rat}} = D[[t]] \cap D(t)$.

Each polynomial $P \in D[t]$ may uniquely be written $P = (1-t)^n Q$, with $n \in \mathbb{N}$, $Q \in D[t]$ and $Q(1) \neq 0$. Thus if $S \in D(t)$, one has $S = (1 - t)^n QR^{-1}$, with $n \in \mathbb{Z}$, $Q, R \in D[t]$ and $Q(1), R(1) \neq 0$. We say that S is *evaluable at $t = 1$* if $n \geq 0$, and in this case, its *value at $t = 1$* is $Q(1)R(1)^{-1}$ if $n = 0$, and 0 if $n \geq 1$. This value is a well-defined element of D, which does not depend on the fraction chosen to represent S.

We extend this to matrices: a matrix over $D(t)$ is said to be *evaluable at $t = 1$* if all his entries are, and then its value at $t = 1$ is defined correspondingly.

Consider a rational expression $E(t)$ over the skew field $D(t)$. We obtain a rational expression over the skew field D by putting $t = 1$ in $E(t)$. Suppose that the rational expression $E(1)$ obtained in this way is evaluable in D and evaluates to $\alpha \in D$; then the rational expression $E(t)$ is evaluable in $D(t)$, evaluates to an element $P(t)Q(t)^{-1}$ in $D(t)$, with $P, Q \in D[t]$, and PQ^{-1} is evaluable at $t = 1$ with value $\alpha \in D$. The standard details are left to the reader.

2.6. Central eigenvalues of matrices over a skew field. Let M be a square matrix over D. Then $1 - Mt$ is invertible over $D[[t]]$, hence over $D((t))$. Since $D(t)$ is a skew field, contained in $D((t))$, and containing the coefficients of $1 - Mt$, the coefficients of its inverse $(tM)^* = (1-tM)^{-1}$ lie also in $D(t)$ and finally in $D[[t]]^{\text{rat}}$. Recall that a square matrix over a skew field is left singular (that is, has a nontrivial kernel when acting at the left on column vectors) is and only if it is right singular. Thus M has an eigenvector for the eigenvalue 1 at the left if and only if it holds on the right.

By the *multiplicity* of the eigenvalue 1 of M we mean the maximum of the nullity (that is, dimension of kernel) of the positive powers of $1 - M$. Observe that this coincides with the usual multiplicity (of 1 in the characteristic polynomial) if D is commutative. Note that the same properties hold for any nonzero central eigenvalue λ by considering $1 - \lambda^{-1}M$; we treat only the case $\lambda = 1$ for the future application.

Lemma 1. *Let M be a square matrix over the skew-field D and t be a central variable.*

(i) *M has the eigenvalue 1 if and only if $(1 - tM)^{-1}$ is not evaluable at $t = 1$.*

(ii) *If M has the eigenvalue 1 with multiplicity 1, then $(1-t)(1-tM)^{-1}$ is evaluable at $t = 1$, is non-null and its rows span the eigenspace for the eigenvalue 1.*

Proof. (i) Suppose that M has the eigenvalue 1. Then M is conjugate over D to a matrix of the form $N = \begin{bmatrix} 1 & 0 \\ P & Q \end{bmatrix}$, where Q is square. Then, computing in $D[[t]]$, we have $(1 - tN)^{-1} = \begin{bmatrix} (1-t)^{-1} & 0 \\ \times & \times \end{bmatrix}$. This is clearly not evaluable for $t = 1$, and therefore $(1 - tM)^{-1}$ is also not evaluable for $t = 1$.

Conversely, suppose $(1 - tM)^{-1}$ is not evaluable for $t = 1$. Then, we have $(1 - tM)^{-1} = ((1 - t)^{n_{ij}} P_{ij}/Q_{ij})_{i,j}$, with $P_{ij}, Q_{ij} \in D[t]$, $P_{ij}(1), Q_{ij}(1) \neq 0$, $n_{ij} \in \mathbb{Z}$ and some $n_{ij} < 0$. Let $-n$ be the minimum of the n_{ij}. Then $n > 0$ and $(1 - t)^n (1 - tM)^{-1}$ is evaluable at $t = 1$ and its value P at $t = 1$ is non-null. Now, we have

$$(1 - tM)^{-1} = 1 + (1 - tM)^{-1} tM,$$

thus

$$(1 - t)^n (1 - tM)^{-1} = (1 - t)^n + (1 - t)^n (1 - tM)^{-1} tM.$$

Since $n > 0$, we obtain for $t = 1$ that

$$P = PM,$$

which shows that M has the eigenvalue 1, since each row of P is fixed by M and $P \neq 0$.

(ii) We write as before $N = \begin{bmatrix} 1 & 0 \\ P & Q \end{bmatrix}$, where N is conjugate to M over D. Then

$$(1 - tN)^{-1} = (tN)^* = \begin{bmatrix} t^* & 0 \\ (tQ)^* tP t^* & (tQ)^* \end{bmatrix}.$$

This formula is easily verified by checking that the right-hand side is the inverse of $1 - tN$.

We claim that $(tQ)^*$ is evaluable at $t = 1$. Indeed, otherwise, by (i), Q has the eigenvalue 1 and is conjugate to a matrix $N = \begin{bmatrix} 1 & 0 \\ R & S \end{bmatrix}$, S square. Then M is conjugate to $N = \begin{bmatrix} 1 & 0 & 0 \\ \times & 1 & 0 \\ \times & R & S \end{bmatrix}$ and the square of $1 - M$ has nullity ≥ 2, contradiction.

Now we see that

$$(1 - t)(tN)^* = \begin{bmatrix} 1 & 0 \\ (tQ)^* tP & (1 - t)(tQ)^* \end{bmatrix}$$

is evaluable at $t = 1$ and that its value at $t = 1$ is non-null. Thus, by the first part of the proof, its rows span the eigenspace for the eigenvalue 1. $\qquad\square$

2.7. Rational series in one variable.

Let R be a ring and t a central variable. In the ring of formal power series $R[[t]]$, we consider the subring $R[[t]]^{\mathrm{rat}}$, which is the smallest subring containing $R[t]$ and closed under inversion. If R is a skew field, then $R[[t]]^{\mathrm{rat}}$ canonically embeds into the skew field $R(t)$. If $R \to S$ is a ring homomorphism, then it induces a ring homomorphism $R[[t]]^{\mathrm{rat}} \to S[[t]]^{\mathrm{rat}}$ fixing t.

3. Generic noncommutative stochastic matrices

3.1. Generic matrices. Let $M = (a_{ij})_{1 \le i, j \le n}$ be a *generic noncommutative matrix*; in other words, the a_{ij} are noncommuting variables. We denote by $\mathcal{F}$ the corresponding free field. Associated to M is the matrix S: it is the same matrix, but this time we assume that the a_{ij} are noncommuting variables subject to the *stochastic identities*

$$\sum_{j=1}^{n} a_{ij} = 1 \quad \text{for all } i = 1, \dots, n. \tag{1}$$

In other words, the row sums of S are equal to 1; equivalently, S has ${}^t(1, \dots, 1)$ as column eigenvector with the eigenvalue 1. We call S a *generic noncommutative stochastic matrix*. The algebra over $\mathbb{Q}$ generated by its coefficients (hence subject to the relations (1)) is a free associative algebra, since it is isomorphic with $\mathbb{Q}\langle a_{ij}, i \ne j \rangle$. Indeed, we may eliminate a_{ii} using (1). We denote this algebra by $\mathbb{Q}\langle a_{ij}/(1) \rangle$, referring to the relations (1). Hence there is a corresponding free field, which we call the *stochastic free field*, denoted $\mathcal{S}$.

3.2. Existence of elements and identities in the stochastic free field. We want to verify that certain rational expressions make sense in the stochastic free field $\mathcal{S}$. For example, taking a 2 by 2 matrix $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we want to show that $(1 + bd^*)^{-1} = (1 + b(1-d)^{-1})^{-1}$ makes sense in $\mathcal{S}$ (hence under the hypothesis $a + b = c + d = 1$). It is necessary to take care of this existence problem, since otherwise, one could invert 0 (and our proved identities will be meaningless). The idea is to prove the existence of certain specializations of the variables, compatible with the identities in $\mathcal{S}$ (identities (1) above), such that the specialized rational expression makes sense. In our example, we could take $b = 0$: then bd^* specializes to 0 and $1 + bd^*$ to 1, hence $(1 + bd^*)^{-1}$ is evaluable under the specialization. A fortiori, since $\mathcal{S}$ is a free field, $(1 + bd^*)^{-1}$ is evaluable in $\mathcal{S}$.

By a *Bernouilli morphism* we mean a $\mathbb{Q}$-algebra morphism π of the free associative algebra $\mathbb{Q}\langle a_{ij} \rangle$ into $\mathbb{R}$ such that

(i) $\sum_{j=1}^{n} \pi(a_{ij}) = 1$ for any $i = 1, \dots, n$;

(ii) $\pi(a_{ij}) > 0$ for any $i, j = 1, \dots, n$.

Clearly, such a morphism induces naturally a $\mathbb{Q}$-algebra morphism form $\mathbb{Q}\langle a_{ij}/(1) \rangle$ into $\mathbb{R}$.

Lemma 2. *There exists a subring $\mathcal{S}_\pi$ of the stochastic free field $\mathcal{S}$ such that*

(i) $\mathcal{S}_\pi$ *contains* $\mathbb{Q}\langle a_{ij}/(1) \rangle$;

(ii) *there is an extension of π to $\mathcal{S}_\pi$ (we still denote it by π);*

(iii) *if $f \in \mathcal{S}_\pi$ and $\pi(f) \ne 0$, then $f^{-1} \in \mathcal{S}_\pi$.*

Proof. This is a consequence of the fact that $\mathcal{S}$ is a free field, corresponding to the free associative algebra $\mathbb{Q}\langle a_{ij}/(1)\rangle$, hence is the universal field of fractions of $\mathbb{Q}\langle a_{ij}/(1)\rangle$. This implies that there exists a specialization $\mathcal{S} \to \mathbb{R}$ extending $\pi : \mathbb{Q}\langle a_{ij}/(1)\rangle \to \mathbb{R}$, and the lemma follows from [10], 7.2 and Cor. 7.5.11. $\qquad\square$

Corollary 1. *Suppose that π is a Bernouilli morphism and that $S = \sum_{w \in L} w$, where L is a rational subset of the free monoid $\{a_{ij}\}^*$, generated by the noncommuting variables a_{ij}, such that $\sum_{w \in L} \pi(w) < \infty$. Then any positive $*$-rational expression for S is evaluable in the stochastic free field $\mathcal{S}$.*

Proof. This is proved inductively on the size of the rational expression for S. Note that for each subexpression and corresponding series S', $\pi(S')$ converges and is > 0. Hence, we apply inductively the lemma and see that for each subexpression, the corresponding element is in $\mathcal{S}_\pi$. $\qquad\square$

Lemma 3. *Let S be a $*$-rational series in $\mathbb{Q}\langle\!\langle a_{ij}\rangle\!\rangle$ having a $*$-rational expression which is evaluable in $\mathcal{S}$. Then it is evaluable in the free field $\mathcal{F}$. If moreover $S = 0$ in $\mathcal{F}$, then $S = 0$ in $\mathcal{S}$.*

Proof. There exists a specialization $\mathcal{F} \to \mathcal{S}$, since $\mathcal{F}$ is the universal field of fractions of $\mathbb{Q}\langle a_{ij}\rangle$; see [10], Chapter 7. Hence there is a subring H of $\mathcal{F}$ and a surjective $\mathbb{Q}$-algebra morphism $\sigma : H \to \mathcal{S}$ such that: $\forall f \in H, \sigma f \neq 0 \implies f^{-1} \in H$, and such that H contains $\mathbb{Q}\langle a_{ij}\rangle$.

We may therefore prove, by induction on the size of the rational expression, that S exists in H and that $\sigma(S)$ is the element of $\mathcal{S}$ defined by the rational expression. It follows that if $S = 0$ in $\mathcal{F}$, then $S = 0$ in $\mathcal{S}$. $\qquad\square$

3.3. Paths. Each path in the complete directed graph with set of vertices $\{1, \ldots, n\}$ defines naturally an element of the free associative algebra $\mathbb{Q}\langle a_{ij}\rangle$, hence of the free field $\mathcal{F}$. This is true also for each rational series in $\mathbb{Q}\langle\!\langle a_{ij}\rangle\!\rangle$.

We define now several such series. First, consider the set of nonempty paths $i \to i$ which do not pass through i; we denote by C_i the sum in $\mathbb{Q}\langle\!\langle a_{ij}\rangle\!\rangle$ of all the corresponding words. It is classically a rational series, and thus defines an element of the free field $\mathcal{F}$. Now let P_i be the sum of the paths (that is, the corresponding words) from i to any vertex j, which do not pass again through i; this set of words is the set of proper prefixes of the words appearing in C_i. Likewise, P_i defines an element of $\mathcal{F}$.

Example 1. $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The graph is
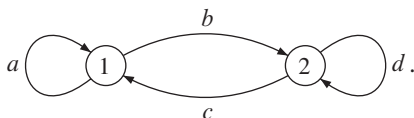
Then

$$C_1 = a + bd^*c, \quad C_2 = d + ca^*b,$$
$$P_1 = 1 + bd^*, \qquad P_2 = 1 + ca^*.$$

3.4. Results

Theorem 1. *The elements P_i can be evaluated in the stochastic free field S and $(P_1^{-1}, \ldots, P_n^{-1})$ is a left eigenvector of the noncommutative generic stochastic matrix S. Moreover, in S,*

(i) $\sum_{i=1}^{n} P_i^{-1} = 1$;

(ii) C_i *can be evaluated in S and is equal to* 1;

(iii) $\lambda(C_i)$ *can be evaluated in S and is equal to* P_i.

Recall that λ is the unique derivation of the free field $\mathcal{F}$ which extends the identity on the set $\{a_{ij}\}$.

Example 1 (continued). We verify that $(P_1^{-1}, P_2^{-1}) \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (P_1^{-1}, P_2^{-1})$. It is enough to show that $P_1^{-1}a + P_2^{-1}c = P_1^{-1}$. This is equivalent to

$$P_2^{-1}c = P_1^{-1}(1 - a) \iff c^{-1}P_2 = a^*P_1 \iff c^{-1} + a^* = a^* + a^*bd^*.$$

Now we use the stochastic identities:

$$a + b = 1 \implies 1 - a = b \implies a^* = b^{-1} \implies a^*b = 1,$$
$$c + d = 1 \implies d^* = c^{-1}.$$

Thus, we may conclude.

(i) Similarly,

$$P_1^{-1} + P_2^{-1} = 1 \iff P_2 + P_1 = P_1 P_2$$
$$\iff 2 + bd^* + ca^* = 1 + bd^* + ca^* + bd^*ca^*,$$

and we conclude since $d^*c = ba^* = 1$.

(ii) In S, $C_1 = a + bd^*c = a + b = 1$.

(iii) In $\mathcal{F}$, $\lambda(C_1) = a + bd^*c + b\lambda(d^*)c + bd^*c$ since λ is a derivation such that $\lambda(a) = a$, $\lambda(b) = b$ and $\lambda(c) = c$. Now

$$\lambda(d^*) = \lambda((1 - d))^{-1} = -(1 - d)^{-1}\lambda(1 - d)(1 - d)^{-1} = d^*dd^*.$$

Thus, this time in S,

$$\lambda(C_1) = a + 2bd^*c + bd^*dd^*c$$
$$= a + 2b + bd^*d$$
$$= a + b + b(1 + d^*d)$$
$$= 1 + bd^* = P_1.$$

3.5. Proof of the theorem

Lemma 4. *Consider the matrix* $(tS)^*$ *in* $S(t)$. *Then* $(1-t)(tS)^*$ *can be evaluated for* $t = 1$ *and is nonzero.*

Proof. By Lemma 1 (ii), it is enough to show that S has the eigenvalue 1 with multiplicity 1. Now, by a change of basis over $\mathbb{Q}$ (replace the canonical basis of column vectors $e_1, \ldots, e_n$ by $e_1, \ldots, e_{n-1}, e_1 + \cdots + e_n$), we bring S to the form

$$T = \begin{bmatrix} N & 0 \\ \lambda & 1 \end{bmatrix},$$

where $n_{ij} = a_{ij} - a_{nj}$ for $1 \le i, j \le n-1$ and $\lambda_j = a_{nj}$ for $j = 1, \ldots, n-1$. We claim that $N - 1$ is invertible in S. It is enough to show that it is full in $\mathbb{Q}\langle a_{ij}/(1)\rangle$. Suppose that $N - 1$ is not full: then $N - 1 = PQ$, with P, Q over $\mathbb{Q}\langle a_{ij}/(1)\rangle$ of size $n \times (n-1)$ and $(n-1) \times n$. By replacing a_{nj} by 0 and a_{ii} by $a_{ii} + 1$, we find that the matrix $(a_{ij})_{1 \le i, j \le n-1}$ is nonfull over $\mathbb{Q}\langle a_{ij}, 1 \le i, j \le n-1\rangle$, which is absurd, since it is a generic matrix. Thus $N - 1$ is invertible and no power of it has a nontrivial kernel. Consequently, the positive powers of $T - 1$ have all rank $n - 1$. Therefore the multiplicity of 1 as eigenvalue of T, hence of S, is 1. $\qquad\square$

Proof of Theorem 1. We identify paths in the complete directed graph on $\{1, \ldots, n\}$, and corresponding words in the free monoid $\{a_{ij}\}^*$. Let us also identify an infinite sum of paths with the corresponding series in $\mathbb{Q}\langle\langle a_{ij}\rangle\rangle$. Let P_{ij} denote the set of paths from i to j that do no pass through i again. We therefore have $P_i = \sum_j P_{ij}$. Denote by $D(u_1, \ldots, u_n)$ the diagonal matrix whose diagonal elements are $u_1, \ldots, u_n$. Observe that each path from i to j may be decomposed as the concatenation of a path from i to i (thus, an element of C_i^*) and a path from i to j that does not pass again through i (thus, an element of P_{ij}). Since $(M^*)_{ij}$ is the sum of all paths from i to j, we obtain the identity in $\mathbb{Q}\langle\langle a_{ij}\rangle\rangle$: $(M^*)_{ij} = C_i^* P_{ij}$. Thus we have the matrix identity: $M^* = D(C_1^*, \ldots, C_n^*)(P_{ij})$. Now $P_{ii} = 1$ and P_{ij} has no constant term. Hence (P_{ij}) is invertible over $\mathbb{Q}\langle\langle a_{ij}\rangle\rangle$.

Inverting, we obtain $D(C_1 - 1, \ldots, C_n - 1) = (P_{ij})(M - 1)$ since $M^* = (1 - M)^{-1}$ and similarly $C_i^* = (1 - C_i)^{-1}$. If we multiply by the column vector $\gamma = {}^t(1, \ldots, 1)$, we obtain ${}^t(C_1 - 1, \ldots, C_n - 1) = (P_{ij})(M - 1)\gamma$.

This equality holds in $\mathbb{Q}\langle\langle a_{ij}\rangle\rangle$, and actually, in its subalgebra of rational series since C_i, P_{ij} are rational series. Hence it holds in the free field $\mathcal{F}$.

Applying the derivation λ of $\mathcal{F}$ we also obtain

$${}^t(\lambda(C_1), \ldots, \lambda(C_n)) = (\lambda(P_{ij}))(M - 1)\gamma + (P_{ij})M\gamma.$$

Now we claim that C_i, P_{ij} and $\lambda(C_i)$ can be evaluated in the stochastic free field S. Thus, since $M\gamma = \gamma$ in S, we obtain that in S

$${}^t(C_1 - 1, \ldots, C_n - 1) = 0,$$

and

$$^t(\lambda(C_1), \ldots, \lambda(C_n)) = (P_{ij})\gamma = {}^t(P_1, \ldots, P_n),$$

which proves parts (ii) and (iii) of the theorem.

In order to prove the claim, we take a Bernouilli morphism π. Let i be some element of $\{1, \ldots, n\}$ and consider the set E of paths not passing through i. Then $\pi(E) < \infty$; indeed, the matrix N, which is obtained from M by removing row and column i, satisfies: $\pi(N)$ has row sums < 1; therefore, $\pi(N)^* = \sum_{k \geq 0} \pi(N)^k$ converges (as is well known) and by graph theory, $\pi(E) = \pi(N^*) = \pi(\bar{N})^*$.

It follows that $\pi(C_i)$, $\pi(P_{ij})$ are finite. For $\lambda(C_i)$, it is easy to see inductively on the size of a rational expression of C_i that since C_i can be evaluated in $\mathcal{S}$, so is $\lambda(C_i)$; one has simply to use the identity $\lambda(H^*) = H^*\lambda(H)H^*$. Note also that $\pi(P_i) > 0$, hence P_i is nonzero in $\mathcal{S}$, and P_i^{-1} is an element in $\mathcal{S}$, by Corollary 1.

We now prove (i). Let Q_i denote the set of paths from 1 to some vertex, that do not pass by i; in particular, $Q_1 = 0$. Then, for any i, j, we have

$$(M^*)_{1i} P_i + Q_i = (M^*)_{1j} P_j + Q_j$$

since both sides represent all the paths departing from 1. Let t be a central variable. Replacing each path w by $t^{|w|}w$ and writing correspondingly $P_1(t), \ldots, P_n(t)$, we obtain

$$(tM)^*_{1i} P_i(t) + Q_i(t) = (tM)^*_{1j} P_j(t) + Q_j(t).$$

This holds in $\mathbb{Q}\langle A\rangle[[t]]$ and actually in its subalgebra of rational elements $\mathbb{Q}\langle A\rangle[[t]]^{\mathrm{rat}}$. Now we have canonical homomorphisms (see 2.5 and 2.7)

$$\mathbb{Q}\langle A\rangle[[t]]^{\mathrm{rat}} \to \mathbb{Q}\langle A/(1)\rangle[[t]]^{\mathrm{rat}} \to \mathcal{S}[[t]]^{\mathrm{rat}} \to \mathcal{S}(t).$$

The composition maps the matrix M onto S. Hence, we have in $\mathcal{S}(t)$

$$(tS)^*_{1i} P_i(t) + Q_i(t) = (tS)^*_{1j} P_j(t) + Q_j(t),$$

where we keep the notation $P_i(t) \in \mathcal{S}(t)$ for the image under the composition. Observe that P_i, by Corollary 1, has a rational expression which can be evaluated in $\mathcal{S}$. Hence $P_i(t)$ can be evaluated for $t = 1$ and is equal to P_i. Similarly, $Q_i(t)$ can be evaluated for $t = 1$ and evaluates to Q_i.

Multiply the last equality by $1 - t$. By Lemma 4, $(1 - t)(tS)^*_{1i}$ can be evaluated for $t = 1$ and is equal to α_i say. Thus, we obtain

$$\alpha_i P_i = \alpha_j P_j.$$

Now $(tS)^* = 1 + (tS)^*tS$ so that, multiplying by $1 - t$ and putting $t = 1$, we obtain that each row of $(1 - t)(tS)^*|_{t=1}$ is fixed by S. In particular, so is $(\alpha_1, \ldots, \alpha_n)$. Since, by Lemma 4, $(1 - t)(tS)^*|_{t=1}$ is nonzero, some row of it is nonzero, and by symmetry, each row is nonzero. Hence, since we already know that each P_i is nonzero in $\mathcal{S}$, we see that each α_i is $\neq 0$. Thus, since $P_i^{-1}\alpha_i^{-1} = P_1^{-1}\alpha_1^{-1}$,

$$(P_1^{-1}, \ldots, P_n^{-1}) = P_1^{-1}\alpha_1^{-1}(\alpha_1, \ldots, \alpha_n),$$

which shows that $(P_1^{-1}, \ldots, P_n^{-1})$ is fixed by S.

Now $\sum_{i=1}^{n} (M^*)_{1i} = M_{11}^* P_1$, since both sides represent the paths departing from 1. Thus we deduce that $\sum_{i=1}^{n} \alpha_i = \alpha_1 P_1$ in S by the same technique as above. Thus,

$$\sum_{i=1}^{n} P_i^{-1} = \sum_{i=1}^{n} P_1^{-1} \alpha_1^{-1} \alpha_i = 1. \qquad \square$$

4. Unambiguous automata

4.1. Unambiguous automata. An *unambiguous automaton* is equivalent to a multiplicative homomorphism μ from the free monoid A^* into $\mathbb{Q}^{n \times n}$ such that each matrix μw, $w \in A^*$, has entries in $\{0, 1\}$. This may be expressed by associating to μ the directed graph with edges labelled in A with vertices $1, \ldots, n$, and edges $i \xrightarrow{a} j$ if and only if $(\mu a)_{ij} = 1$. Then the non-ambiguity means that for any vertices i, j and any word w, there is at most one path from i to j labelled w (the label of a path is the product of the label of the edges). The *matrix of the automaton* is by definition $M = \sum_{a \in A} a\mu a$.

We say that the unambiguous automaton is *complete* if the zero matrix does not belong to the monoid μA^*. Equivalently, for each word w there is some path labelled w. We say that the automaton is *transitive* if the underlying graph is strongly connected. This means that for any vertices i, j, there is some path $i \rightarrow j$; equivalently, $(\mu w)_{ij} \neq 0$ for some word w.

The monoid μA^* is finite. Hence it has a unique minimal ideal I. There is a *rank* function on μA^*, and the elements of minimum rank are precisely the elements of I. Since $\mu A^* \subseteq \{0, 1\}^{n \times n}$, the rows of an element in μA^* are ordered by inclusion (by identifying a subset of $\{1, 2, \ldots, n\}$ and its characteristic row vector). It is shown that the nonzero rows of elements of the minimal ideal are precisely the maximal rows of elements of μA^*. Similarly for columns. The ideal I is the disjoint union of the minimal right (resp. left) ideals of μA^*, and the intersection of a minimal left and a minimal right ideal is a group. For this, see [4], Chapter VI, and [5], Chapter VI, especially Exercise 3.4 and also [3].

We shall use the following result.

Proposition 1. *Let c be a maximal column and R be the sum of the distinct rows of some element in the minimal ideal. Then $Rc = 1$.*

Proof. There exist x, y in I such that c is a column of x and R is the sum of the distinct rows of y. The element xy is in I and belongs therefore to a group with neutral element e, say. Then e has a column-row decomposition $e = st$, where s (resp. t) is a $n \times r$ (resp. $r \times n$) matrix with entries in $\{0, 1\}$, with r the minimal rank of μA^*, where $ts = I_r$ (the identity matrix), and where the set of nonzero rows of

e is the set of rows of t, which has distinct rows, and similarly for the columns of s (see [4] Prop. IV.3.3 or [5], Prop. VI.2.3).

Now xM is a minimal right ideal of μA^*, containing e, hence $xM = eM$ and therefore $x = em = stm$. Hence c is a sum of columns of s, and since c is a maximal column, c is a column of s. Similarly, $y = nst$ and each nonzero row of y is a row of t. We have also $e = n'y$, hence each nonzero row of t, being a row of e, is a row of y. Thus R is the sum of the rows of t: $R = \lambda t$, with $\lambda = (1, \ldots, 1)$. Finally $Rc = \lambda tc$ and since $ts = I_r$ and c is a column of s, tc is a column of I_r and $\lambda tc = 1$. □

Example 2. The unambiguous automaton is as follows:
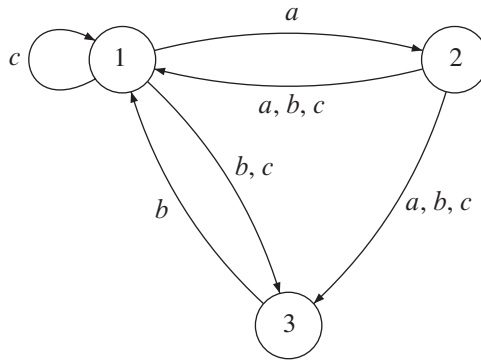
The associated representation μ is defined by

$$\mu a = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mu b = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mu c = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrix of the automaton is

$$M = \begin{bmatrix} c & a & b+c \\ a+b+c & 0 & a+b+c \\ b & 0 & 0 \end{bmatrix}.$$

Idempotents in the minimal ideal are for example μc and $\mu ba = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. The maximal rows are $(1, 0, 1)$ and $(0, 1, 0)$ and the maximal columns are $^t(1, 1, 0)$ and $^t(0, 1, 1)$.

4.2. Codes. Recall that a code is the basis of some free submonoid of the free monoid. Given an unambiguous automaton with associated representation μ, and some vertex i, the language $\{w \in A^* \mid (\mu w)_{ii} = 1\}$ is a free submonoid of A^*; we denote by C_i its unique basis, which is therefore a code. Explicitly, C_i is the set of

labels of paths $i \rightarrow i$ which do not contain i as internal vertex. Note that C_i is a rational code and that each rational code is obtained in this way. We shall use also the set P_i of labels of paths starting at i and not passing again through i. See [4].

Example 2 (continued). Write $A = a + b + c$. Then

$$C_1 = c + aA(1 + b) + (b + c)b,$$
$$C_2 = A(c + b^2 + cb)^*a + Ab(c + b^2 + cb)^*a,$$
$$C_3 = b(c + aA)^*(b + c + aA),$$
$$P_1 = 1 + a + aA + b + c,$$
$$P_2 = 1 + A(c + b^2 + cb)^*(1 + b + c)$$
$$\quad + A(1 + b(c + b^2 + cb)^*(b + c)) + Ab(c + b^2 + cb)^*,$$
$$P_3 = 1 + b(c + aA)^*(1 + a).$$

We shall use the following property of rational maximal codes: let C be such a code; then there exists rational languages P, S, F, whose elements are factors of words of C, such that in $\mathbb{Q}\langle\langle A \rangle\rangle$

$$A^* = SC^*P + F.$$

Moreover $1 \notin F$, $1 \in S$, $1 \in P$. This property is proved in [5], Lemma XII.4.3, for finite codes. The proof is easily extended to rational codes.

4.3. Bernouilli morphisms.
A Bernouilli morphism is a multiplicative morphism $\pi: A^* \to \mathbb{R}_+$ such that $\pi|A$ is a probability on A such that $\pi(a) > 0$ for any a in A.

It is known that if L is a language having the property that it does not intersect some ideal in A^*, then $\pi(L) = \sum_{w \in L} \pi(w) < \infty$. This property is true if L is a rational code. See [4], Propositions I.5.6 and I.5.12.

From this, we deduce that $\pi(L) < \infty$ for each language $L = C_i, P_i, S, P, F$ considered in Section 4.2.

4.4. Probabilistic free field.
We know that $\mathbb{Q}\langle A \rangle$ is embedded in the corresponding free field denoted $\mathcal{F}$. Consider now the $\mathbb{Q}$-algebra $\mathbb{Q}\langle A \rangle / (A - 1)$, which is the quotient of $\mathbb{Q}\langle A \rangle$ by its two-sided ideal generated by $A - 1 = \sum_{a \in A} a - 1$. This $\mathbb{Q}$-algebra is a free associative algebra, since the relation $A = 1$ allows to eliminate one variable. We denote it $\mathbb{Q}\langle A/(A - 1) \rangle$ Hence, there is a corresponding free field, denoted $\mathcal{P}$ and which we call the *probabilistic free field*.

Theorem 2. *Let $\mu: A^* \to \mathbb{Q}^{n \times n}$ be the homomorphism corresponding to a complete and transitive unambiguous automaton. Let $M = \sum_{a \in A} a\mu a$ be its matrix, P the image of M in the probabilistic free field $\mathcal{P}$, C_i the code generating the fixpoints of vertex i, P_i the sum of the labels of all paths starting at i and not passing again*

through i. *Then* P_i, C_i *and* P_i^{-1} *can be evaluated in* $\mathcal{P}$. *Moreover, the following equalities hold in* $\mathcal{P}$:

  (i) $C_i = 1$;

  (ii) $(1 - t)(tP)^* \in \mathcal{P}(t)$ *can be evaluated at* $t = 1$ *and its diagonal elements are* $\lambda(C_i)^{-1}$, $i = 1, \dots, n$;

  (iii) $(P_1^{-1}, \dots, P_n^{-1})P = (P_1^{-1}, \dots, P_n^{-1})$;

  (iv) $\sum_{i=1}^{n} P_i^{-1} = 1$;

  (v) *for any maximal columns* ℓ, ℓ', $(P_1^{-1}, \dots, P_n^{-1})\ell = (P_1^{-1}, \dots, P_n^{-1})\ell'$.

Example 2 (continued). $C_1 = 1$ holds in $\mathcal{P}$ since one has even in $\mathbb{Q}\langle A \rangle$: $C_1 - 1 = (1 + a)(a + b + c - 1)(1 + b)$. Moreover, in $\mathcal{P}$ we have

$$C_2 = (c + b^2 + cb)^*a + b(c + b^2 + cb)^*a = (1 + b)(c + b^2 + cb)^*a.$$

Now, in $\mathbb{Q}\langle\langle b, c \rangle\rangle$, one has $(1 + b)(c + b^2 + cb)^* = (b + c)^*$, since $\{c, b^2, cb\}$ is a complete suffix code with set of suffixes $\{1, b\}$ (see [4]). Thus $C_2 = (b + c)^*a = 1$ since $a = 1 - b - c$. Also,

$$C_3 = b(c + a)^*(b + c + a) = b(c + a)^* = 1.$$

In $\mathcal{P}$, we have $S = \begin{bmatrix} c & a & b+c \\ 1 & 0 & 1 \\ b & 0 & 0 \end{bmatrix}$. We show that $P_2 = a^{-1}P_1$; indeed

$$
\begin{aligned}
P_2 &= 1 + (c + b^2 + cb)^*(1 + b + c) + 1 \\
&\quad + b(c + b^2 + cb)^*(b + c) + b(c + b^2 + cb)^* \\
&= 2 + (1 + b)(c + b^2 + cb)^*(1 + b + c) \\
&= 2 + (b + c)^*(1 + b + c) \\
&= 2 + (b + c)^* + (b + c)^*(b + c) \\
&= 1 + 2(b + c)^* = 1 + 2a^{-1} = a^{-1}P_1
\end{aligned}
$$

since $P_1 = 2 + a$. We deduce that $P_1^{-1}a = P_2^{-1}$. Moreover,

$$P_3 = 1 + b(c + a)^*(1 + a) = 2 + a = P_1$$

since $b(c + a)^* = 1$. Thus

$$
\begin{aligned}
P_1^{-1}(b + c) + P_2^{-1} &= P_1^{-1}(a + b + c) = P_1^{-1} = P_3^{-1}, \\
P_1^{-1}c + P_2^{-1} + P_3^{-1}b &= P_1^{-1}(c + a + b) = P_1^{-1}.
\end{aligned}
$$

This shows that $(P_1^{-1}, P_2^{-1}, P_3^{-1})S = (P_1^{-1}, P_2^{-1}, P_3^{-1})$. Furthermore, $P_1^{-1} + P_2^{-1} + P_3^{-1} = P_1^{-1}(1 + a + 1) = 1$. Now the only two maximal columns are ${}^t(1, 1, 0)$ and ${}^t(0, 1, 1)$. We have $(P_1^{-1}, P_2^{-1}, P_3^{-1})\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = (P_1^{-1}, P_2^{-1}, P_3^{-1})\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ since $P_1 = P_3$.

4.5. Proof of theorem. We need the following lemma.

Lemma 5. *Let $P = \sum_{a \in A} a \mu a \in \mathcal{P}^{n \times n}$ be the image in $\mathcal{P}^{n \times n}$ of the matrix M of some complete and transitive unambiguous automaton, with associated homomorphism $\mu \colon A^* \to \mathbb{Q}^{n \times n}$. Then P has the eigenvalue 1 with associated eigenspace of dimension 1. Moreover, if t is a central variable, then in $\mathcal{P}(t)$, $(1 - t)(tP)^*$ can be evaluated for $t = 1$ and its rows span the eigenspace above.*

Proof. Consider the (left) $\mathcal{P}$-subspace E of $\mathcal{P}^{1 \times n}$ spanned by the maximal rows. It has as subspace the subspace E' spanned by the differences of such rows. Let C be the sum of the distinct columns of some element of the minimal ideal of μA^*. Then $rC = 1$ if r is a maximal row (dual statement of Proposition 1). Thus E' is strictly included in E.

By Section 4.1, for each maximal row r and each $a \in A$, $r \mu a$ is a maximal row, denoted r_a. Then

$$rP = \sum_{a \in A} r(\mu a)a = \sum_{a \in A} r_a a = r + \sum_{a \in A} (r_a - r)a,$$

since $\sum_{a \in A} a = 1$ in $\mathcal{P}$. Thus r is fixed by P modulo the subspace E'. Hence P has 1 as eigenvalue.

We show that its multiplicity is 1. Indeed the multiplicity does not decrease under specialization. For the latter, we take a positive Bernouilli morphism π; then $\pi(P)$ is an irreducible matrix because the automaton is transitive; it has nonnegative coefficients. We claim that its eigenvalues are of modulus ≤ 1. Thus, we may apply the Perron–Frobenius theorem ([19], Section 15.3, Theorem 1), and since 1 is an eigenvalue of $\pi(P)$ by the previous calculations, it is a root of multiplicity 1 of the characteristic polynomial. But we know that 1 is an eigenvalue of P, hence it has multiplicity 1. We conclude by using Lemma 1.

It remains to prove the claim. Since the automaton is unambiguous, the matrix $M^n = (\sum_{a \in A} a \mu a)^n = \sum_{w \in A^n} w \mu w$ has the property that each entry is a subsum of $\sum_{w \in A^n} w$. Hence each entry of $\pi(M^n) = \pi(P^n)$ is bounded by 1. Hence each eigenvalue of $\pi(P)$ has modulus ≤ 1. $\qquad\square$

Proof of Theorem 2. The set C_i is a rational maximal code. So we may use the result at the end of Section 4.2: $A^* = SC_i^*P + F$, where S, P, F are rational languages contained in the set of factors of C_i. Then, by Section 4.3, $\pi(S), \pi(P)$ and $\pi(F)$ are $< \infty$ for any Bernouilli morphism. This implies that S, P, F can be evaluated in $\mathcal{P}$ (cf. the proof of Corollary 1). The same holds for C_i and P_i. Now the equality in $\mathbb{Q}\langle\!\langle A \rangle\!\rangle$ above may be rewritten:

$$A^* - F = SC_i^*P$$
$$\implies 1 - (1 - A)F = (1 - A)SC_i^*P$$
$$\implies (1 - (1 - A)F)^{-1} = P^{-1}(1 - C_i)S^{-1}A^*$$
$$\implies 1 - C_i = P(1 - (1 - A)F)^{-1}(1 - A)S.$$

This holds in $\mathbb{Q}\langle\!\langle A \rangle\!\rangle^{\mathrm{rat}}$ and all these rational expressions can be evaluated in $\mathcal{P}$. Thus, in $\mathcal{P}$, we obtain $1 - C_i = 0$.

Let Q_i be the set of paths from 1 to any vertex, that do not pass again through i. Then we have, as in the proof of Theorem 1, for any i, j,

$$(M^*)_{1i} P_i + Q_i = (M^*)_{1j} P_j + Q_j.$$

Arguing as in the latter proof, we find that, denoting α_i the value of $(1 - t)(tP)^*_{1i}$ at $t = 1$ (which exists by Lemma 5), we obtain $\alpha_i P_i = \alpha_j P_j$. Note that $(tP)^*_{11} = C_1(t)^*$, where $C_1(t)^*$ denotes the canonical image of $\sum_{w \in C_1^*} t^{|w|} w \in \mathbb{Q}\langle A \rangle[[t]]^{\mathrm{rat}}$ under the composition of homomorphisms

$$\mathbb{Q}\langle A \rangle[[t]]^{\mathrm{rat}} \to \mathbb{Q}\langle A/(A-1)\rangle[[t]]^{\mathrm{rat}} \to \mathcal{P}[[t]]^{\mathrm{rat}} \to \mathcal{P}(t).$$

Thus α_1 is the value at $t = 1$ of $(1-t)C_1(t)^*$. Now taking the previous notations with $i = 1$, we have in $\mathcal{P}(t)$: $(tA)^* = S(t)C_1(t)^* P(t) + F(t)$. Multiplying by $(1 - t)$ and putting $t = 1$, we obtain, since S, P, F can be evaluated in $\mathcal{P}$: $1 = S\alpha_1 P$. Thus $\alpha_1 = S^{-1}P^{-1}$.

Now we have also $C_1 - 1 = P(1 - (1 - A)F)^{-1}(A - 1)S$. Thus, in $\mathcal{F}$, letting $P' = P(1 - (1 - A)F)^{-1}$,

$$\lambda(C_1) = \lambda(P')(A - 1)S + P'\lambda(A)S + P'(A - 1)\lambda(S).$$

We deduce that, in $\mathcal{P}$, $\lambda(C_1) = P'\lambda(A)S = PS$. This shows that $\alpha_1 = \lambda(C_1)^{-1}$. This proves (ii) and in particular, $\alpha_1 \neq 0$. Thus, since $P_1 \neq 0$ in $\mathcal{P}$, all α_i and P_i are $\neq 0$ in $\mathcal{P}$. Then (iii) and (iv) are proved as in the proof of Theorem 1.

In order to prove (v), we observe that the elements of the minimal ideal I of μA^* are those of this monoid which have a minimal number of distinct non-null rows (see [5], Exercise VI.3.5, or [3], Proposition 1). This implies that if $r_1, \ldots, r_k$ are the distinct non-null rows of some element μw of I, then for any letter a, $r_1 \mu a, \ldots, r_k \mu a$ are the distinct non-null rows of $\mu(wa)$. We deduce that the span of the elements $r_1 + \cdots + r_k$ is invariant under the matrices μa; indeed, $(r_1 + \cdots + r_k)\mu a = r'_1 + \cdots + r'_k$, where $r'_1, \ldots, r'_k$ are the distinct non-null rows of $\mu(wa)$. Let F denote this subspace, and F' the subspace spanned by the difference of such elements. By Proposition 1, we have that F' is strictly included in F. Hence, there is a vector in F fixed by each μa. This implies that the eigenvector for eigenvalue 1 of the matrix P is in F and is therefore orthogonal to each difference of maximal columns of μA^*. This proves (v). □

1. Appendix 1: the commutative case

The following result in an exercise on determinants, which is left to the reader (see also Remark 2 in Appendix 2).

Lemma 6. *If the column eigenvector $^t(1, \ldots, 1)$ is in the right kernel of a square matrix over a commutative ring, then the row vector $(M_1, \ldots, M_n)$ is in its left kernel, where M_i is the i-th principal minor of the matrix.*

From this, one may deduce the so-called *Markov chain tree theorem*, by using, as suggested in [20], p. 4, the matrix-tree theorem. This may be done as follows: let a_{ij} for $i \neq j$, $i, j = 1, \ldots, n$ be commuting variables and $M = (m_{ij})$ be the n by n matrix with $m_{ij} = -a_{ij}$ if $i \neq j$ and with diagonal elements $m_{ii} = -\sum_{j \neq i} m_{ij}$, so that M has the column vector $^t(1, \ldots, 1)$ in its right kernel. Then by [9] (where a generalization of the classical matrix-tree theorem is proved), the principal i-th minor M_i of M is equal to the sum of the weights of all spanning trees of the complete digraph on $\{1, \ldots, n\}$, rooted at i (the edges of the tree all pointing toward i). Here the weight of a subgraph is the product of the a_{ij} for all edges (i, j) in the subgraph. Note that for a subtree, its weight involves only a_{ij} with $i \neq j$.

From this, we may deduce the Markov chain tree theorem: let (a_{ij}) be a stochastic matrix (that is, fixing $^t(1, \ldots, 1)$). Then the row vector $(b_1, \ldots, b_n)$ is fixed by this matrix, where b_i is the sum of the weights of all spanning trees of the complete digraph on $\{1, \ldots, n\}$, rooted at i. As before, the weight of a subgraph is the product of the a_{ij} for all edges (i, j) in the subgraph. We deduce it from the argument above since the diagonal entries depend only on the others and since the spanning trees do not involve loops.

The Markov chain tree theorem gives a formula for the stationary distribution of a finite Markov chain. Equivalently, this formula gives a row vector fixed by a matrix fixing $^t(1, \ldots, 1)$. This theorem is attributed to Kirchoff by Persi Diaconis, who gives a probabilistic proof of it (see [8] p. 443–444). See also [2], [1].

Using our Theorem 1, one may deduce that, if $M = (a_{ij})$ is a generic commutative matrix, then $B = \sum_{i=1}^n b_i$, the sum of the weights of all rooted trees, is equal to the derivative of $-\det(1 - (a_{ij}))$, with respect to the derivation fixing each a_{ij}, and then substituting in it each a_{ii} by $1 - \sum_{j \neq i} a_{ij}$. Indeed, letting the variables in Theorem 1 commute, we know that $P_i = \lambda(C_i)$ and that C_i^* is the i-th element in $(1 - M)^{-1}$, hence $(1 - C_i)^{-1} = N_i/D$, where $D = \det(1 - (a_{ij}))$ and where N_i is the i-th principal minor of $1 - M$. Thus $C_i - 1 = -D/N_i$. By Theorem 1, this element is mapped onto 0 if we apply the above substitution, since the latter implies the relations (1). Now apply derivation λ to the last equation. We find $\lambda(C_i) = (-\lambda(D)N_i + D\lambda(b_i))/N_i^2$. Applying the substitution, we find that $\lambda(C_i) = -\lambda(D)/N_i$. Thus, by Theorem 1 again, $1 = \sum P_i^{-1} = \sum \lambda(C_i)^{-1} = -\sum N_i/\lambda(D)$ and therefore $-\lambda(D) = \sum N_i$. Finally we note that, with the above notation, $b_i = N_i$ after the substitution.

2. Appendix 2: Quasideterminants of stochastic matrices

2.1. Theory of quasideterminants. The theory of quasideterminants was developed as a tool for linear algebra over noncommutative rings replacing the theory of determinants over commutative rings. Quasideterminants were introduced in [14] and developed in [15], [16], [17] and [18]. Let R be an associative unital ring and let $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$ be a matrix over R. Denote by A^{ij} the submatrix of A obtained from A by deleting its i-th row and its j-th column. Set $r_i = (a_{i1}, a_{i2}, \ldots, \hat{a}_{ij}, \ldots, a_{in})$ and $c_j = (a_{1j}, a_{2j}, \ldots, \hat{a}_{ij}, \ldots, a_{nj})$. Recall, that for any matrix C we denote by ${}^t C$ the transposed matrix.

Definition. Suppose that the matrix A^{ij} is invertible. Then the quasideterminant $|A|_{ij}$ is defined as

$$|A|_{ij} = a_{ij} - r_i (A^{ij})^{-1} c_j.$$

Example. If $n = 2$, then $|A|_{12} = a_{12} - a_{11} a_{21}^{-1} a_{22}$.

Let matrix A be invertible and $A^{-1} = (b_{pq})$. If the quasideterminant $|A|_{ij}$ is invertible then $b_{ji} = |A|_{ij}^{-1}$.

In the commutative case, $|A|_{ij} = (-1)^{i+j} \det A / \det A^{ij}$.

It is sometimes convenient to use another notation for quasideterminants $|A|_{ij}$ by boxing the leading entry, i.e.,

$$|A|_{ij} = \begin{vmatrix} \ldots & \ldots & \ldots \\ \ldots & \boxed{a_{ij}} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}.$$

We now recall the basic properties of quasideterminants (see [14]). An equality $|A|_{pq} = |B|_{rs}$ means that the first quasideterminant is defined if and only if the second quasideterminant is defined and that both quasideterminants are equal. The properties are:

(i) *Permutations of rows and columns.* Let σ, τ be permutations of $\{1, 2, \ldots, n\}$. Set $B = (a_{\sigma(i), \tau(j)})$. Then $|A|_{pq} = |B|_{\sigma(p), \tau(q)}$.

(ii) *Multiplication of row and columns.*

Let the matrix $B = (b_{ij})$ be obtained from matrix A by multiplying the i-th row by $\lambda \in R$ from the left, i.e., $b_{ij} = \lambda a_{ij}$ and $b_{kj} = a_{kj}$ for all j and $k \neq i$. Then $|B|_{kj} = \lambda |A|_{kj}$ if $k = i$, and $|B|_{kj} = |A|_{kj}$ if $k \neq i$ and λ is invertible.

Let the matrix $C = (c_{ij})$ be obtained from matrix A by multiplying the j-th column by $\mu \in R$ from the right, i.e., $c_{ij} = a_{ij} \mu$ and $c_{il} = a_{il}$ for all i and $l \neq j$. Then $|C|_{il} = |A|_{il} \mu$ if $l = j$, and $|C|_{il} = A|_{il}$ if $l \neq j$ and μ is invertible.

(iii) *Addition of rows and columns.*

Let the matrix B be obtained from A by replacing the k-th row of A with the sum of k-th and l-th row, i.e., $b_{kj} = a_{kj} + a_{lj}$, $b_{ij} = a_{ij}$ for $i \neq k$. Then $|A|_{ij} = |B|_{ij}$, $i = 1, 2, \ldots, \hat{k}, \ldots, n$, $j = 1, 2, \ldots, n$.

Let the matrix C be obtained from A by replacing the k-th column of A with the sum of k-th and l-th column, i.e., $c_{ik} = a_{ik} + a_{il}$, $b_{ij} = a_{ij}$, $c_{ij} = a_{ij}$ for $j \neq k$. Then $|A|_{ij} = |C|_{ij}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, \hat{l}, \ldots n$.

In [14] a noncommutative analogue of the Cramer's rule for systems of left linear equations, i.e., the systems when coefficients are at the left of the unknowns, was formulated. The analogue for systems of right linear equations can be formulated as follows.

Let $B = (b_{ij})$ an $n \times n$-matrix over R, $\xi = (\xi_i)$ be a row-matrix over R and $x = (x_i)$ be a row-matrix of unknowns. Here $i, j = 1, 2, \ldots, n$. For $1 \leq k \leq n$ denote by $B(\xi, k)$ the matrix obtained from B by replacing the k-th row of B by ξ.

Proposition 2. *If $xB = \xi$ then*

$$x_k |B|_{kq} = |B(\xi, k)|_{kq}$$

for any k provided that the both quasideterminants are defined.

Example. For $n = 2$ one has

$$x_1(b_{12} - b_{11}b_{21}^{-1}b_{22}) = \xi_2 - \xi_1 b_{21}^{-1} b_{22}$$

and also

$$x_1(b_{11} - b_{12}b_{22}^{-1}b_{21}) = \xi_1 - \xi_1 b_{22}^{-1} b_{21}.$$

2.2. Results

Lemma 7. *Let $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$, be a stochastic matrix over R. Consider the system of $n + 1$ equations*

$$\sum_{i=1}^{n} x_i a_{ij} = x_j, \quad j = 1, 2, \ldots, n, \tag{2}$$

together with the equation

$$\sum_{i=1}^{n} x_i = 1.$$

Then any of the n first equations of the system is a corollary of the other n equations.

Proof. Take any $1 \leq k \leq n$ and add all equations (2) for $j \neq k$. The right hand side of the sum can be written as $1 - x_k$ and the left hand side can be written as $\sum_{i=1}^{n} x_i (1 - a_{ik})$. As a result we have

$$\sum_{i=1}^{n} x_i (1 - a_{ik}) = 1 - x_k,$$

which implies

$$\sum_{i=1}^{n} x_i a_{ik} = x_k.$$

The lemma is proved. □

Theorem 3. *Let* $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$, *be a stochastic matrix. The system*

$$\sum_{i=1}^{n} x_i a_{ij} = x_j, \quad j = 1, 2, \ldots, n,$$

$$\sum_{i=1}^{n} x_i = 1$$

has a unique solution over the algebra of series in variables a_{ij} *satisfying the relations* $\sum_{j=1}^{n} a_{ij} = 1$. *The solutions are given by the formula*

$$x_k^{-1} = 1 + \sum a_{ki_1} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{s-1} i_s},$$

where the sum is taken over all sets of naturals $i_1, i_2, \ldots, i_s$, *with* $s \geq 1$ *and* $i_p \neq k$, $p = 1, 2, \ldots, s$.

Proof. Lemma 7 implies that $x_1, \ldots, x_n$ are solutions of the system

$$\sum_{i=1}^{n} x_i = 1,$$

$$\sum_{i=1}^{n} x_i (a_{ij} - \delta_{ij}) = 0, \quad j \neq k.$$

Write the system in the form $xB = \xi$, where $x = (x_1, \ldots, x_n)$ and $\xi = (1, 0, \ldots, 0)$, and apply Proposition 2. Note that $|B(\xi, k)|_{k1} = 1$: indeed, the k-th row of $B(\xi, k)$ is ξ and therefore, by the definition of quasideterminants, $r_k = (0, \ldots, 0)$ and $|B(\xi, k)|_{k1} = B(\xi, k)_{k1} = 1$. Therefore,

$$x_k |B|_{k1} = 1.$$

Recall that A^{kk} is the submatrix of A obtained from A by omitting its k-th row and k-th column. Set $C = A^{kk}$. Let I be the unit matrix of order $n - 1$ and

$a(k) = (a_{k1}, a_{k2}, \ldots, \hat{a}_{kk}, \ldots, a_{kn})$. Note that the first column of matrix B is $\gamma = {}^t(1, 1, \ldots, 1)$. By the definition of quasideterminants

$$|B|_{k1} = 1 - a(k)(C-I)^{-1}\gamma = 1 + a(k) \sum_{p=0}^{\infty} C^p \gamma = 1 + \sum a_{ki_1} a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_{s-1} i_s}.$$

This proves the theorem. □

Remark 1. Note that monomials $a_{ki_1} a_{i_1 i_2} \ldots a_{i_{s-1} i_s}$ can be interpreted as paths in the complete graph with vertices $1, 2, \ldots, n$.

2.3. Stochastic matrices and main quasiminors. Observe that a matrix M is stochastic if and only if ${}^t(1, 1, \ldots, 1)$ is in the kernel of $M - I$. This justifies the next results.

Lemma 8. *Let $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$, over an associative unital ring annihilate the column vector ${}^t(1, 1, \ldots, 1)$. For $p \neq q$ one has*

$$|A^{pq}|_{qp} = -|A^{pp}|_{qq}$$

if the right-hand side is defined.

Proof. Without loss of generality one can assume that $p = 1$ and $q = n$. Then

$$|A^{1n}|_{n1} = \begin{vmatrix} a_{21} & a_{22} & \ldots & a_{2n-1} \\ a_{31} & a_{32} & \ldots & a_{3n-1} \\ & & \ldots & \\ \boxed{a_{n1}} & a_{n2} & \ldots & a_{nn-1} \end{vmatrix}.$$

Since A is stochastic by hypothesis, we can rewrite the elements a_{k1}, $k = 2, 3, \ldots, n$, as $a_{k1} = -a_{k2} - a_{k3} - \cdots - a_{kn}$. By adding the columns in the last quasideterminant to the first one and using property (iii), property (ii) for the first column and $\mu = -1$ and property (i) we get the expression

$$- \begin{vmatrix} a_{2n} & a_{22} & \ldots & a_{2n-1} \\ a_{3n} & a_{32} & \ldots & a_{3n-1} \\ & & \ldots & \\ \boxed{a_{nn}} & a_{n2} & \ldots & a_{nn-1} \end{vmatrix} = - \begin{vmatrix} a_{22} & a_{23} & \ldots & a_{2n} \\ a_{32} & a_{33} & \ldots & a_{3n} \\ & & \ldots & \\ a_{n2} & a_{n3} & \ldots & \boxed{a_{nn}} \end{vmatrix}$$

which is $-|A^{11}|_{nn}$. Our computations also show the existence of $|A^{1n}|_{n1}$. The lemma is proved. □

Theorem 4. *Let $A = (a_{ij})$, $i, j = 1, 2, \ldots, n$, annihilate the column vector ${}^t(1, 1, \ldots, 1)$. Assume that all quasideterminants $|A^{ii}|_{jj}$ are defined for $i \neq j$. Then $xA = 0$ where $x = (x_1, x_2, \ldots, x_n)$ if and only if*

$$x_i |A^{jj}|_{ii} = x_j |A^{ii}|_{jj}, \quad i \neq j.$$

Remark 2. In the commutative case $|A i i|_{jj} = \det A^{ii}/\det A^{ij,ij}$ provided that the denominator is defined. Here $A^{ij,ij}$ is the submatrix of A obtained from A by removing its rows and columns with the indices i and j. Thus the theorem implies

$$x_i M_j = x_j M_i,$$

where the M_i's are the main minors of the matrix and we may choose $x_i = M_i$, $i = 1, 2, \ldots, n$, as a solution of the equation $xA = 0$ obtaining Lemma 6.

Proof of Theorem 4. We will prove the "if" part. The "only if" part can be proved by reversing the arguments. Without loss of generality, we assume that $i = 1$ and $j = n$. Note that x_i's satisfy the system of linear equations

$$\sum_{p=1}^{n-1} x_p a_{pq} = -x_n a_{nq}, \quad q = 1, 2, \ldots, n-1.$$

Cramer's rule give us the equality

$$x_1 |A^{nn}|_{11} = \begin{vmatrix} \boxed{-x_n a_{n1}} & -x_n a_{n2} & \ldots & -x_n a_{nn-1} \\ a_{21} & a_{22} & \ldots & a_{2n-1} \\ & & \ldots & \\ a_{n-11} & a_{n-12} & \ldots & a_{n-1n-1} \end{vmatrix}.$$

By properties (i) and (ii), the right-hand side is equal to $-x_n |A^{1n}|_{n1}$.
The theorem now follows from Lemma 8.   □

By using the results from [15] and [16] we can show that Theorem 4 implies Theorem 3 provided the corresponding quasideterminants are invertible.

References

[1] D. J. Aldous, The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.* **3** (1990), 450–465. Zbl 0717.05028 MR 1069105

[2] V. Anantharam and P. Tsoucas, A proof of the Markov chain tree theorem. *Statist. Probab. Lett.* **8** (1989), 189–192. Zbl 0679.60069 MR 1017890

[3] M-P. Béal, E. Czeizler, J. Kari, and D. Perrin, Unambiguous automata. *Math. Comput. Sci.* **1** (2008), 625–638. Zbl 1158.68019 MR 2415162

[4] J. Berstel and D. Perrin, *Theory of codes.* Pure Appl. Math. 117, Academic Press Inc., Orlando, FL, 1985. Zbl 0587.68066 MR 0797069

[5] J. Berstel, D. Perrin, and C. Reutenauer, *Codes and automata.* Encyclopedia Math. Appl. 129, Cambridge University Press, Cambridge 2010. Zbl 1187.94001 MR 2567477

[6] J. Berstel and C. Reutenauer, *Rational series and their languages.* Monogr. Theoret. Comput. Sci. EATCS Ser. 12, Springer-Verlag, Berlin 1988. Zbl 0668.68005 MR 0971022

[7] J. Berstel and C. Reutenauer, *Noncommutative rational series with applications*. Encyclopedia Math. Appl. 137, Cambridge University Press, Cambridge 2010.

[8] A. Broder, Generating random spanning trees. In *30th Annual symposium on foundations of computer science*, IEEE Computer Society Press, Washington, DC, 1989, 442–447.

[9] S. Chaiken and D. J. Kleitman, Matrix tree theorems. *J. Combinatorial Theory Ser. A* **24** (1978), 377–381. Zbl 0376.05032 MR 0480115

[10] P. M. Cohn, *Free rings and their relations*. 2nd ed., London Math. Soc. Monogr. 19, Academic Press, London 1985. Zbl 0659.16001 MR 0800091

[11] P. M. Cohn, From Hermite rings to Sylvester domains. *Proc. Amer. Math. Soc.* **128** (2000), 1899–1904. Zbl 0949.16007 MR 1646314

[12] P. M. Cohn, *Free ideal rings and localization in general rings*. New Math. Monogr. 3, Cambridge University Press, Cambridge 2006. Zbl 1114.16001 MR 2246388

[13] M. Fliess, Sur le plongement de l'algèbre des séries rationelles non commutatives dans un corps gauche. *C. R. Acad. Sci. Paris Sér. A* **271** (1970), 926–927. Zbl 0211.36002 MR 0272833

[14] I. M. Gel′fand and V. S. Retakh, Determinants of matrices over noncommutative rings. *Funktsional. Anal. i Prilozhen.* **25** (1991), 13–25; English transl. *Funct. Anal. Appl.* **25** (1991), 91–102. Zbl 0748.15005 MR 1142205

[15] I. M. Gel′fand and V. S. Retakh, Theory of noncommutative determinants, and characteristic functions of graphs. *Funktsional. Anal. i Prilozhen.* **26** (1992), 1–20. English transl. *Funct. Anal. Appl.* **26** (1992), 231–246. Zbl 0799.15003 MR 1209940

[16] I. M. Gelfand and V. Retakh, A theory of noncommutative determinants and characteristic functions of graphs. Publications du LACIM 14, UQAM, Montréal (1993), 1–26.

[17] I. Gelfand and V. Retakh, Quasideterminants, I. *Selecta Math.* (*N.S.*) **3** (1997), 517–546. Zbl 0919.16011 MR 1613523

[18] I. Gelfand, S. Gelfand, V. Retakh, and R. L. Wilson, Quasideterminants. *Adv. Math.* **193** (2005), 56–141. Zbl 1079.15007 MR 2132761

[19] T. Lancaster and M. Tismenetsky, *Theory of matrices*. 2nd ed., Academic Press, Orlando 1985. Zbl 0558.15001 MR 0792300

[20] L. Lovász and P. Winkler, Exact mixing in an unknown Markov chain. *Electron. J. Comb.* **2** (1995), Research paper R15. Zbl 0823.60057 MR 1345248

S. Lavallée, Département de Mathématiques (LaCIM), Université du Québec à Montréal, CP 8888, Succ. Centre-ville, Montréal (Québec) H3C 3P8, Canada

E-mail: lavallee.sylvain.2@courrier.uqam.ca

D. Perrin, Institut Gaspard-Monge, Université de Paris-Est, 77454 Marne-la-Vallée Cedex 2, France

E-mail: perrin@univ-mlv.fr

V. Retakh, Department of Mathematics, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, U.S.A.

E-mail: vretakh@math.rutgers.edu

C. Reutenauer, Département de Mathématiques (LaCIM), Université du Québec à Montréal, CP 8888, Succ. Centre-ville, Montréal (Québec) H3C 3P8, Canada

E-mail: reutenauer.christophe@uqam.ca