Enumerative sequences of leaves in rational trees

Frédérique Bassino[1] and Marie-Pierre Béal[2] and Dominique Perrin[1]

[1] Institut Gaspard Monge, Université de Marne-la-Vallée
[2] Institut Gaspard Monge, Université Paris 7 et CNRS

http://www-igm.univ-mlv.fr/~{bassino,beal,perrin}

Abstract. We prove that any $\mathbb{N}$-rational sequence $s = (s_n)_{n\geq 1}$ of nonnegative integers satisfying the Kraft strict inequality $\sum_{n\geq 1} s_n k^{-n} < 1$ is the enumerative sequence of leaves by height of a rational k-ary tree. Particular cases of this result had been previously proven. We give some partial results in the equality case.

1 Introduction

This paper is a study of problems linked with coding and symbolic dynamics. The results can be considered as an extension of the old results of Huffman, Kraft, McMillan and Shannon on source coding. We actually prove results on rational sequences of integers that can be realized as the enumerative sequence of leaves in a rational tree.

Let s be an $\mathbb{N}$-rational sequence of nonnegative numbers, that is a sequence $s = (s_n)_{n\geq 1}$ such that s_n is the number of paths of length n going from an initial state to a final state in a finite multigraph or a finite automaton. We say that s satisfies the Kraft inequality for a positive integer k if $\sum_{n\geq 1} s_n k^{-n} \leq 1$.

A rational tree is a tree which has only a finite number of non-isomorphic subtrees. If s is the enumerative sequence of leaves of a rational k-ary tree, then s satisfies Kraft's inequality for the integer k.

In this paper, we study the converse of the above property. Consider for example the series $s(z) = \frac{3z^2}{1-z^2}$. We have $s(1/2) = 1$ and we can obtain s as the enumerative sequence of the tree of the figure below associated with the prefix code $X = (aa)^*(ab + ba + bb)$ on the binary alphabet $\{a, b\}$. We dont know however if the same can be done for the series $s(z) = z^2(\frac{1}{1-z^2} + \frac{2}{1-2z^3})$.

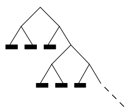

Fig. 1. Tree associated to $3z^2(z^2)^*$

Known constructions allow one to obtain a sequence s satisfying Kraft's inequality as the enumerative sequence of leaves of a k-ary tree, or as the enumerative sequence of leaves of a (perhaps not k-ary) rational tree. These two constructions lead in a natural way to the problem of building a tree both rational and k-ary. This question was already considered in [9], where it was conjectured that any $\mathbb{N}$-rational sequence satisfying Kraft's inequality is the enumerative sequence of leaves of a k-ary rational tree.

In this paper, we prove this conjecture in the case where the sequence satisfies Kraft's inequality with a strict inequality, and we give some partial results in the equality case. For example, we state the following weaker property for such a sequence: If s is an $\mathbb{N}$-rational sequence of nonnegative numbers satisfying Kraft's equality, then there is a positive integer m such that $ms = \sum_{1 \leq i \leq m} r_i$, where each r_i is the enumerative sequence of the leaves of a k-ary rational tree.

Proofs and algorithms used to establish the results are based on automata theory and symbolic dynamics. In particular, we use the state splitting algorithm which has been introduced by R. Adler, D. Coppersmith and M. Hassner in [1] to solve coding problems for constrained channels by constructing finite-state codes with sliding block decoders. This was partly based on earlier work of B. Marcus in [7].

A variant of the problem considered here consists in replacing the enumerative sequence of leaves by the enumerative sequence of all nodes. Soittola ([11]) has characterized the series which are the enumerative sequence of nodes in a rational tree. The problem of a similar characterization for rational k-ary trees remains open in the general case.

In [9], a particular case is treated. It allows to solve the problem for the enumerative sequence of leaves in the equality case under the additional assumption of a unique pole of minimal modulus.

The paper is organized as follows. We first give basic definitions and properties of rational objects, sequences and trees. We then give some definitions coming from the theory of symbolic dynamics. We define the notions of state splitting, approximate eigenvector and recall the algorithm of [1]. In section 3, we establish the announced results and give examples for the constructions.

2 Definitions and background

2.1 Rational sequences of nonnegative numbers

We denote by G a directed graph with E as its set of edges. We actually use multigraphs instead of ordinary graphs in order to be able to have several distinct edges with the same origin and end. Formally a multigraph is given by two sets E (the edges) and V (the vertices) and two functions from E to V which define the origin and the end of an edge. An edge in a multigraph going from p to q will be noted (p, x, q) where $x \in \mathbb{N}$. This is equivalent to number the edges going from p to q in order to distinguish them. We shall always say "graph" instead of "multigraph".

In this paper, we consider sequences of nonnegative numbers. Such a sequence $s = (s_n)_{n \geq 0}$ will be said to be $\mathbb{N}$-*rational* if s_n is the number of paths of length n going from a state in I to a state in F in a finite directed graph G, where I and F are two special subsets of states, the initial and final states respectively. We say that the triple (G, I, F) is a *representation* of the sequence s.

This definition is usually given for the series $\sum_{n\geq 0} s_n z^n$ instead of the sequence s. Any $\mathbb{N}$-rational sequence s satisfies a recurrence relation with integer coefficients. However, it is not true that a sequence of nonnegative integers satisfying a linear recurrence relation is $\mathbb{N}$-rational. An example can be found in [5] p. 93.

A well known result in automata theory allows us to use a particular representation of an $\mathbb{N}$-rational sequence s. One can choose a representation (G, i, F) of s with a unique initial state i and such that :

- no edge is coming in state i
- no edge is going out of any state of F.

Such a representation is called a *normalized representation.* Moreover, it is possible to reduce to one state the set of final states (see for example [10] p. 14).

We now give some basic definitions about trees. A *tree* T on a set of nodes N with a root r is a function $T : N - \{r\} \longrightarrow N$ which associates to each node distinct from the root its father $T(n)$ in such a way that, for each node n, there is a nonnegative integer h such that $T^h(n) = r$. The integer h is the height of the node n. A tree is k-ary if each node has at most k sons. A leaf is a node without son. We denote by $l(T)$ the enumerative sequence of its leaves by height, that is the sequence of numbers s_n, where s_n is the number of leaves of T at height n. A tree is said to be *rational* if it admits only a finite number of non isomorphic subtrees. If T is a rational tree, the sequence $l(T)$ is an $\mathbb{N}$-rational sequence.

The sequence $s = l(T)$ of a k-ary tree is the length distribution of a prefix code over a k-letter alphabet. The associate series $s(z) = \sum_{n\geq 1} s_n z^n$ satisfies then Kraft's inequality : $s(1/k) \leq 1$. We shall say that Kraft's strict inequality is satisfied when $s(1/k) < 1$. The equality is reached when each node of the tree has exactly zero or k sons. Conversely, the McMillan construction establishes that for any series s satisfying Kraft's inequality, there is a k-ary tree such that $s = l(T)$. Moreover, if the series satisfies Kraft's equality, then the internal nodes will have exactly k sons. But the tree obtained is not rational in general.

It is also easy to see that an $\mathbb{N}$-rational sequence is the enumerative sequence of the leaves of a rational tree. A normalized representation can be used to do that by "developing" the tree. The root will correspond to the initial state of the graph. If a node of the tree at height n corresponds to a state i in the graph which has r outgoing edges ending at states $j_1, j_2, \ldots, j_r$, it will admit r sons at height $n+1$, each of them corresponding respectively to the states $j_1, j_2, \ldots, j_r$ of the graph. The leaves of the tree will correspond to the final states of the normalized representation. The maximal number of sons of a node we get is then equal to the maximal number of edges going out of any state of the graph of this representation.

If s satisfies Kraft's inequality, the above construction does not lead in general to a k-ary rational tree. The aim of this paper is to get a k-ary rational tree T such that $s = l(T)$. This result was conjectured in [9]. We solve it for all $\mathbb{N}$-rational sequences satisfying Kraft's strict inequality and give a weaker result for the equality case.

2.2 Approximate eigenvector and state splitting

Let s be an $\mathbb{N}$-rational sequence and let (G, i, F) be a normalized representation of s. If we identify the initial state i and all final states of F in a single state still denoted i, we get a new graph denoted $\overline{G}$, which is strongly connected. The sequence s is then the length distribution of the paths of first returns to state i, that is of finite paths going from i to i without going through state i. Using the terminology of symbolic dynamics, the graph $\overline{G}$ can be seen as an irreducible shift of finite type (see, for example, [3], [4] or [6]).

We denote by M the adjacency matrix associated to the graph $\overline{G}$, that is the matrix $M = (m_{ij})_{1 \leq i,j \leq n}$, where n is the number of nodes of $\overline{G}$ and where m_{ij} is the number of edges going from state i to state j. By the Perron-Frobenius theorem (see [6]), the positive matrix M associated to the strongly connected graph $\overline{G}$ has a positive eigenvalue of maximal modulus denoted by λ, also called the spectral radius of the matrix. Actually, λ only depends on the series s, $1/\lambda$ is the minimal modulus of the poles of $\frac{1}{1-s}$. The dimension of the eigenspace of λ is equal to one. There is a positive eigenvector (componentwise) associated to λ. Moreover, if there is a positive eigenvector associated to an eigenvalue ρ, then $\rho = \lambda$.

When λ is an integer , the matrix admits a *positive integral* eigenvector. When $\lambda < k$, where k is an integer, the matrix admits a *k-approximate* eigenvector, that is, by definition, a positive integral vector $\mathbf{v}$ with $M\mathbf{v} \leq k\mathbf{v}$.

For example the left side of the figure below gives a representation (G, i, F) of the serie $s(z) = \frac{3z^2}{1-z^2}$, and the right side gives the associated graph $\overline{G}$. The adjacency matrix of $\overline{G}$ is

$$M = \begin{pmatrix} 0 & 3 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Its maximal eigenvalue is $\lambda = 2$. The components of a positive integral eigenvector are written on the nodes.

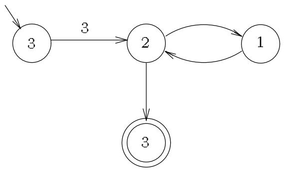

Fig. 2. Representation (G, i, F)

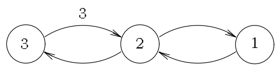

Fig. 3. Graph $\overline{G}$

Proposition 1. *If s satisfies Kraft's inequality $s(1/k) \leq 1$, then $\lambda \leq k$. In the equality case where $s(1/k) = 1$ we have $\lambda = k$.*

For a proof, we refer the reader to [3], [4] or [6].

We now define the operation of *output state splitting* in a graph $G = (V, E)$. Let q be a vertex of G and let I (resp. O) be the set of edges coming in q (resp. going out of q). Let $O = O' + O''$ be a partition of O. The operation of *(output)* *state splitting* relative to (O', O'') transforms G into the graph $G' = (V', E')$ where $V' = (V \setminus \{q\}) \cup q' \cup q''$ is obtained from V by splitting state q into two states q' and q'', and where E' is defined as follows:

1. all edges of E that are not incident to q are left unchanged.
2. the both states q' and q'' have the same input edges as q.
3. the output edges of q are distributed between q' and q'' according to the partition of O into O' and O''. We denote U' and U'' the sets of output edges of q' and q'' respectively :
   $U' = \{(q', x, p) \mid (q, x, p) \in O'\}$ and $U'' = \{(q'', x, p) \mid (q, x, p) \in O''\}$.

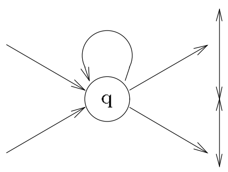

Fig. 4. Graph G

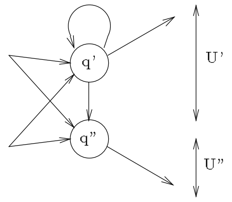

Fig. 5. Graph G'

Let us now assume that $\mathbf{v}$ is a k-approximate eigenvector for the graph G. We denote by v_p the component of index p of $\mathbf{v}$. All components v_p are positive integers. A state splitting of a state q is said to be *admissible* according to k, if the partition in O' and O'' is such that O' and O'' are not empty and:

$$k \text{ divides } \sum_{(q,x,r)\in O'} v_r$$

If the state splitting is admissible (according to k), the vector $\mathbf{v}'$ defined as follows will be a k-approximate eigenvector for the new graph G'. If p is a state distinct from q' and q'' then $v'_p = v_p$. For states q' and q'' we have:

$$v'_{q'} = \frac{1}{k} \sum_{(q,x,r)\in O'} v_r \qquad \text{and} \qquad v'_{q''} = v_q - v'_{q'}.$$

By the state splitting construction, one can check that $M'\mathbf{v}' \leq k\mathbf{v}'$, where M' is the adjacency matrix of G'.

The state splitting algorithm of [1] ensures that there is a finite number of state splittings leading to a k-ary graph, that is a graph such that at most k edges are going out of any state. For the sake of completeness, we briefly recall the proof. If there is a state q which admits more than k edges going out of it, we choose k of them and denote by $r_1, r_2, \ldots, r_k$ the sequence of end states of these edges. We then choose a subset O' of these k edges such that k divides $\sum_{(q,x,r)\in O'} v_r$. This is always possible. Indeed, by considering the $k+1$ numbers $v_{r_1}, v_{r_1}+v_{r_2}, \ldots, v_{r_1}+v_{r_2}+\cdots v_{r_k}$, we can see that at least two of them are equal modulo k, and then their difference is equal to zero modulo k. The partition of the output edges of q in O' and O'' leads to an admissible state splitting and v'_q is strictly less than v_q. This point ensures that the process stops after a finite number of splits, the final number of states being bounded by the sum of the components of the initial approximate eigenvector. The final graph obtained is k-ary.

We shall compute approximate eigenvectors for the strongly connected graphs $\overline{G}$ associated to normalized representations (G, i, F) of sequences. We shall then perform admissible state splittings that can be seen either on the graph G or on the graph $\overline{G}$. To do that, we shall associate to each node of G a *value* equal to the corresponding component of the approximate eigenvector of the graph $\overline{G}$. The initial and the final states will have same value since they correspond to the same state of $\overline{G}$.

3 The results

We now state the result in the case of Kraft strict inequality.

Theorem 2. *Let $s = (s_n)_{n \geq 1}$ be an $\mathbb{N}$-rational sequence of nonnegative integers et let k be an integer such that $\sum_{n \geq 1} s_n k^{-n} < 1$. Then there is a k-ary rational tree such that s is the enumerative sequence of its leaves.*

In order to prove this result, we first prove one lemma that remains true in the equality case. We therefore consider an $\mathbb{N}$-rational sequence s and an integer k such that $\sum_{n \geq 1} s_n k^{-n} \leq 1$. We begin with a normalized representation (G, i, F) of the $\mathbb{N}$-rational sequence s. We denote by M the adjacency matrix of G and by λ its spectral radius. Then $\lambda \leq k$. We then compute a k-approximate eigenvector $\mathbf{v} = (v_1, v_2, \ldots, v_n)^t$ of the graph $\overline{G}$. By definition, we have $M\mathbf{v} \leq k\mathbf{v}$. Without loss of generality, we can assume that state 1 is the initial state in all normalized representations.

Lemma 3. *If k divides v_1, then there is another normalized representation for s and a new corresponding approximate eigenvector v' with $v'_1 = v_1$ div k.*

Proof. We denote by P the set of states q such that there is in G an edge denoted (q, x, t) going from q to a final state t of F. Remark that, as state t is equal to state 1 in $\overline{G}$, the value of state t is equal to the value of state 1.

Let us first suppose that the initial state 1 does not belong to the set P. If there is in P a state q which admits more than one (say n) outgoing edges, we split q in q' and q'' according to partition (O', O'') where $O' = \{(q, x, t)\}$. Since k divides v_1, this state splitting is admissible and $v'_{q'} = v_1$ div k. Moreover, in the new graph G', q' admits only one outgoing edge (going to t) and q'' is either not in P or admits less than n outgoing edges. By successive state splittings of all states in P having more than one outgoing edges, we will get, in a finite number of steps, a representation such that all states with one outgoing edge ending in F have no other outgoing edges. Under the hypothesis that state 1 does not belong to P, the initial state has not been split during this processand so each new computed graph is still a normalized representation of the sequence. We denote again by $(G, 1, F)$ the final representation obtained for s and by P_{last} the set of states having one outgoing edge ending in F in this graph. Remark that the values of states of P_{last} are greater than or equal to v_1 div k. We turn all values of states of P_{last} greater than v_1 div k into v_1 div k; the vector $\mathbf{v}$ remains a k-approximate eigenvector.

We then transform the representation $(G, 1, F)$ in a new one, (H, i, P_{last}), where H is the graph obtained from G by adding a state i, an edge from i to 1 and by removing all edges of G going out of a state of P_{last}. If we look at paths in G going from 1 to F, we have just cut the last edge and added one at the beginning. We assign to state i the value v_1 div k, and the values of all states correspond now to a new k-approximate eigenvector for $\overline{H}$. We call this tranformation the "shift" transformation.

Let us now suppose that the initial state 1 belongs to P. We first split, as explained above, all states of P having more that one outgoing edge. In this case, state 1 may have been split. We denote by $1_{(1)}, 1_{(2)}, 1_{(3)}, \ldots 1_{(r)}$ the copies of state 1 obtained by successive state splittings of the initial state 1. We still denote by G the graph obtained by this transformation and by P_{last} the set of states having one outgoing edge ending in F in this graph. We then transform the representation $(G, 1, F)$ into a new one, (H, i, P_{last}), where H is the graph obtained from G by adding a state i, an edge from i to each $1_{(j)}, 1 \leq j \leq r$ and by removing all edges of G going out of a state of P_{last}. Remark that $(r-1)$ states among $1_{(1)}, 1_{(2)}, 1_{(3)}, \ldots 1_{(r)}$ belong to P_{last}. We again assign to the state i the value v_1 div k, and the values of all states correspond now to a new k-approximate eigenvector for $\overline{H}$.

Corollary 4. *If v_1 is a power of k, then there is another normalized representation and a new corresponding approximate eigenvector v' with $v'_1 = 1$.*

Proof. If $v_1 = k^m$, we iterate the construction given in previous lemma and get $v'_1 = 1$ in m steps.

Example
Let s be the following series:

$$s(z) = 2z^3 + 2z^2 \left(z^2(z^2)^*\right)^*$$

Here, $k = 2$ and $s(1/2) = 1$.
In the following pictures, the nodes are labeled with their value.

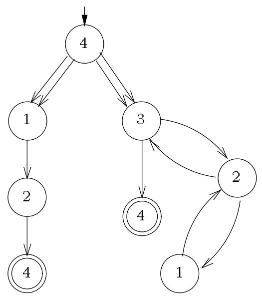

Fig. 6. Initial normalized representation

First step

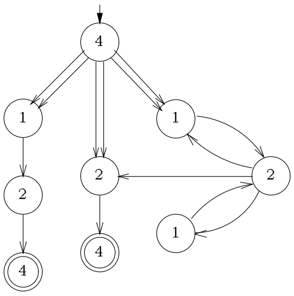

Fig. 7. First state splitting

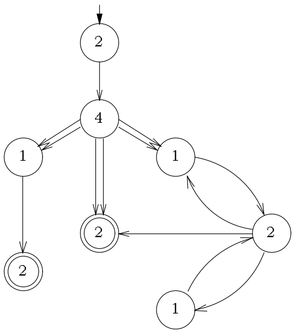

Fig. 8. First "shift"

Second step

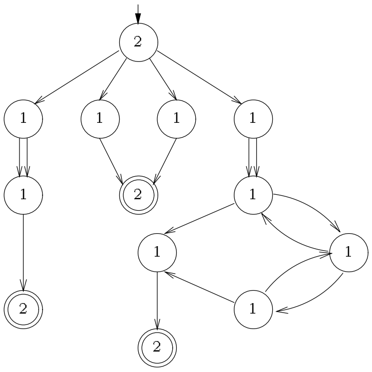

Fig. 9. Other state splittings

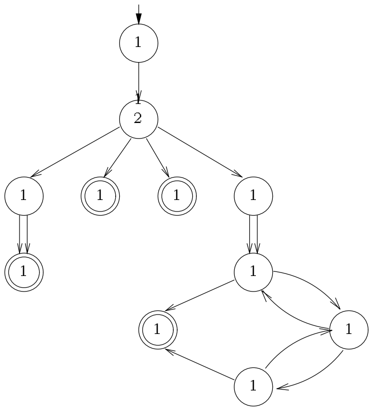

Fig. 10. Second "shift"

The *last step* is described in the proof of Theorem 2.
It corresponds here to a state splitting of all states of the graph of value different from 1.

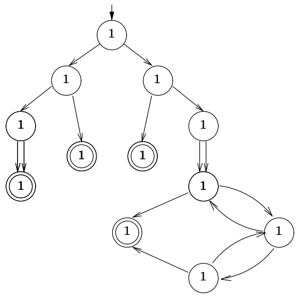

Fig. 11. Last representation

We now prove another lemma which is true only in the case of Kraft strict inequality.

Lemma 5. *Let M be a nonnegative integral matrix. If its spectral radius is strictly less than k, then there is a k-approximate eigenvector $\mathbf{w}$ of M such that w_1 is a power of k.*

Proof. Let λ ($\lambda < k$) be the positive real eigenvalue of maximal modulus of M and let $\mathbf{v}$ be an eigenvector associated to λ. We denote by P the set of positive vectors $\mathbf{w}$ such that $M\mathbf{w} < k\mathbf{w}$. The set P is an open set and $\mathbf{v}$ belongs to P. By dividing all components of $\mathbf{v}$ par v_1, we can assume that v_1 is equal to 1. As P is open, there is a positive real ϵ such that $B(v,\epsilon) \subset P$, where $B(v,\epsilon) = \{\mathbf{w} \mid v_i - \epsilon \leq w_i \leq v_i + \epsilon\}$. Let us now choose an integer m such that $1/k^m < \epsilon$. As $B(v, 1/k^m) \subset P$, we have $\{k^m \mathbf{w} \mid \mathbf{w} \in B(v, 1/k^m)\} \subset P$. This set is $\{\mathbf{w} \mid k^m v_i - 1 \leq w_i \leq k^m v_i + 1\}$ and contains $\mathbf{w}$ where $w_i = \lceil k^m v_i \rceil$. This vector is a positive integer vector $\mathbf{w}$ with $M\mathbf{w} < k\mathbf{w}$: it is a k-approximate eigenvector. Moreover $w_1 = k^m$.

Proof. (Theorem 2) We begin with a normalized representation of s and compute, by Lemma 5, a k-approximate eigenvector whose component for the initial state is a power of k. We then compute, by Corollary 4, a normalized representation $(G, 1, F)$ of s which admits a k-approximate eigenvector of component 1 for the initial state. Finally, we apply to G the state splitting algorithm described in the previous section to obtain a k-ary graph. As the component of the approximate eigenvector on the initial state is 1 and as the state splittings have to be admissible, this state will never be split during the process. A state splitting of a state of G different from state 1 leads by construction to a graph G' still representing the same sequence. The result follows then from the fact that the final normalized representation has a k-ary graph.

We can apply the construction given above to the case of Kraft equality when it is possible to find a representation of s which admits a k-eigenvector with a power of k as component on the initial state. This may perhaps not always be the case. We do not know, for example, if the series $s(z) = \frac{z^2}{(1-z^2)} + \frac{z^2}{(1-5z^3)}$ (communicated to us by Christophe Reutenauer) has such a representation for $k = 2$:

As a consequence of the previous result, we get the following proposition in the equality case, where an ultimately k-ary tree is a tree where all nodes but a finite number have at most k sons.

Proposition 6. *Let $s = (s_n)_{n\geq 1}$ be an $\mathbb{N}$-rational sequence of nonnegative integers and let k be an integer such that $\sum_{n\geq 1} s_n k^{-n} = 1$. Then there is an ultimately k-ary rational tree such that s is the enumerative sequence of its leaves.*

Proof. If we remove one term of the sequence, the remainder satisfies Kraft's strict inequality and is still $\mathbb{N}$-rational. This proves that one can construct a rational tree T for s which will be k-ary for all nodes except the root which will have $k + 1$ sons.

We now state another result for the equality case which is weaker than the previous theorem. We show that if s is an $\mathbb{N}$-rational sequence of nonnegative integers satisfying Kraft's equality for an integer k, then there is an integer m such that ms is the sum of m enumerative sequences of leaves of m k-ary rational trees.

Theorem 7. *Let s be a an $\mathbb{N}$-rational sequence satisfying $\sum_{n\geq 1} s_n k^{-n} = 1$. There is a positive integer m and m k-ary rational trees $T_1, \ldots, T_m$ such that $ms = l(T_1) + \cdots + l(T_m)$.*

Proof. We begin with a normalized represention of s : $(G, 1, F)$. Let $\mathbf{v}$ be a positive integral eigenvector associated to the spectral radius k of the adjacency matrix of G. The component v_1 on the initial state 1 is denoted by m. If $m = k^r m'$, where m' and k are relatively prime, we compute, by Corollary 4, a normalized represention of s such that $v_1 = m'$ in order to get a smaller integer m. After this step, m and k are relatively prime. If $m = 1$, we finish by the same proof as the proof of Theorem 2.

Otherwise, we denote by r the positive integer such that $k^{r-1} < m < k^r$ and $x = k^r - m$. We define a new graph H by adding to G $(r + 1)$ new states $i_1, i_2, \ldots, i_r$ and j, and with the following new edges :

$$i_1 \longrightarrow i_2 \longrightarrow \cdots i_r \longrightarrow 1$$

and x edges (in the multigraph) going from i_r to j :

$$i_r \longrightarrow j \ (x \text{ edges})$$

We assign to state i_l the value k^{l-1} and to state j the value 1 (state 1 has value m). For each state t in F we make the following transformation. We replace t by m copies of this state and we duplicate each edge entering t in m edges entering the m copies of t. We give to all copies of t the value 1. We denote by g the new sequence which admits as normalized representation $(H, i_1, F \cup \{j\})$. Note that the values of states of H correspond now to a k-integer eigenvector of $\overline{H}$ which admits 1 as component on the initial state i_1. Using the series notations, one can verify that

$$g(z) = \sum_{n\geq 1} g_n z^n = xz^r + mz^r s(z),$$

and by construction: $g(1/k) = 1$.

Since the representation of g has an eigenvector of component 1 at the initial state, g is the height distribution of leaves of a k-ary rational tree (by applying the construction of theorem 1 in the case where the value of the initial state is 1). The series g is either equal to 1 or to $zg_1 + zg_2 + \cdots + zg_k$, where g_l are again series of this type (g_l is the height distribution of the leaves of the subtree rooted by a son of the root of the tree representing g). By iterating this decomposition for each g_l, we can write

$$g(z) = z^r(f_1(z) + f_2(z) + \cdots + f_{k^r}(z)) = xz^r + mz^r s(z),$$

where f_i are height distributions of leaves of k-ary rational trees, which we simplify into:

$$f_1 + f_2 + \cdots + f_{k^r} = x + ms$$

As all f_i have nonnegative integer coefficients and satisfy $f_i(1/k) = 1$, this implies that x series among $f_1, f_2, \ldots, f_{k^r}$ are equal to 1. The m remainding series that we renumber $f_1, f_2, \ldots, f_m$ verify the equality : $f_1 + f_2 + \cdots + f_m = ms$, which is the announced result.

Example Let s be the following series:

$$s(z) = \frac{z^2}{1-z^2} + \frac{2z^2}{1-2z^3}.$$

We get that $3s = f_A + f_B + f_C$, where f_X is the height distribution of the leaves of the tree rooted by the node X in the last picture.

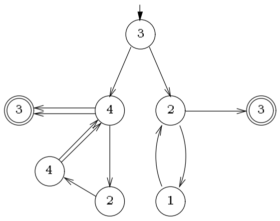

Fig. 12. The sequence s

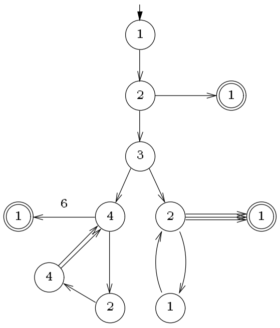

Fig. 13. The sequence g

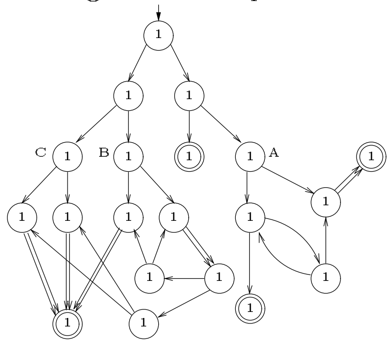

Fig. 14. The sequence g after state splittings

References

1. R. L. Adler, D. Coppersmith, and M. Hassner. Algorithms for sliding block codes. *I.E.E.E. Trans. Inform. Theory*, IT-29:5–22, 1983.
2. F. Bassino. *Séries rationnelles et distributions de longueurs.* Thèse, Université de Marne-La-Vallée, 1996.
3. M.-P. Béal. *Codage Symbolique.* Masson, 1993.
4. M.-P. Béal and D. Perrin. Symbolic dynamics and finite automata. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages.* Springer-Verlag, 1997.
5. J. Berstel and C. Reutenauer. *Rational Series and their Languages.* Springer-Verlag, 1988.
6. D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding.* Cambridge, 1995.
7. B. Marcus. Factors and extensions of full shifts. *Monats.Math*, 88:239–247, 1979.
8. D. Perrin. Arbres et séries rationnelles. *C.R.A.S. Paris, Série I*, 309:713–716, 1989.
9. D. Perrin. A conjecture on rational sequences. In R. Capocelli, editor, *Sequences*, pages 267–274. Springer-Verlag, 1990.
10. D. Perrin. Finite automata. In J. V. Leeuven, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 1. Elsevier, 1990.
11. A. Salomaa and M. Soittola. *Automata-theoretic Aspect of Formal Power Series.* Springer-Verlag, Berlin, 1978.