

# Matrice de ramification des arbres binaires\*

J.G. Penaud

*L.A.B.R.I., Université de Bordeaux I, 33405 Talence, France*

Received 7 July 1988

Revised 27 June 1989

## Résumé

Penaud, J.G., Matrice de ramification des arbres binaires, *Discrete Applied Mathematics* 30 (1991) 1–21.

A tout sommet d'un arbre binaire on associe son *nombre de Strahler*, puis on considère une matrice dite de *ramification*, qui reflète la distribution des nombres de Strahler des fils des sommets qui ont un nombre de Strahler donné. Cette notion est alors étendue à l'ensemble des arbres binaires d'une taille donnée et on montre la forme remarquable que prend cette matrice lorsque la taille tend vers l'infini.

*Abstract.* A *Strahler number* is associated with each node of a binary tree. Then, a *ramification matrix*, reflecting the distribution of the Strahler numbers, is constructed. This notion is extended to the family of binary trees having given size and the asymptotic behaviour of this matrix is shown.

## 1. Introduction

Le paramètre *nombre de Strahler* fut introduit par les hydrogéologues Horton [7] et Strahler [13] pour rendre compte de la forme des bassins fluviaux modélisés par des arbres. Ce paramètre intervient dans de nombreux domaines où les arbres apparaissent, en particulier en informatique; appelé alors *nombre de registres*, il détermine le nombre minimal de registres nécessaires à l'évaluation d'une expression arithmétique, comme l'a établi Ershov [2]. La valeur moyenne asymptotique de ce paramètre a été étudiée par divers auteurs [11,8,4] et bijectivement en [6].

Pour le dessin d'arbres de la nature, Viennot [16] a introduit un raffinement de ce paramètre, la *matrice de ramification*; c'est une matrice stochastique dont la valeur donne des informations sur l'"allure" de l'arbre. Le lecteur trouvera dans

\* Travail financé par le PRC Mathématiques et Informatique, et le CJPP.

[17] une application de cette notion. Son application en physique à l'étude des structures ramifiées est exposée dans [14].

Le but de ce papier est de prouver une conjecture de Viennot: la matrice de ramification d'un arbre binaire aléatoire prend, lorsque le nombre de sommets croît infiniment, la forme remarquable suivante,

$$\begin{array}{cccccc}
 1/2 & 1/2 & & & & \\
 1/2 & 1/4 & 1/4 & & & \\
 1/2 & 1/4 & 1/8 & 1/8 & & \\
 1/2 & 1/4 & 1/8 & 1/16 & 1/16 & \\
 \vdots & & & & & \\
 (1/2) & (1/2)^2 & (1/2)^3 & \dots & (1/2)^{n-1} & (1/2)^{n-1} \\
 \vdots & \vdots & \vdots & & \vdots & \vdots
 \end{array}$$

Pour établir la preuve, nous montrons que les termes de la matrice sont la limite du quotient des coefficients de deux séries entières. Ces séries se déduisent de séries énumérant des arbres selon leur taille et une certaine distribution du nombre de Strahler. Pour trouver ces dernières, nous utilisons la démarche désormais classique, introduite en France par Schützenberger [12], qui consiste à coder les objets à énumérer par un langage algébrique non ambigu et déduire des équations de ce langage un système d'équations dont les solutions sont les séries entières énumérant les objets. Malheureusement, nous ne savons pas donner une formule close pour les coefficients, nous faisons alors une analyse asymptotique à la "Odlyzko" (cf. [5]), et par une application du *lemme de transfert* établi par Flajolet dans [3], nous en déduisons le résultat.

La Paragraphe 2 contient les définitions et les notations utilisées. Dans le Paragraphe 3 nous introduisons les langages algébriques et nous étudions les propriétés des séries associées aux matrices de ramification. Les résultats asymptotiques et le théorème final sont rassemblés dans le Paragraphe 4. Les preuves dans ces deux paragraphes sont assez techniques, et furent facilitées par le logiciel MACSYMA du MIT (voir à ce sujet [1] pour de nombreux exemples d'utilisation). Enfin nous concluons par quelques remarques dans le Paragraphe 5.

## 2. Définitions et notations

On considère la famille des arbres binaires complets au sens de Knuth [9], c'est-à-dire possédant une racine et tels que tout sommet interne a exactement deux fils, un *fils droit* et un *fils gauche*. Le nombre de sommets internes est appelé la *taille* de l'arbre.

Pour tout sommet d'un arbre binaire on définit un paramètre, le *nombre de Strahler* (cf. [13]), encore appelé *nombre de registre* (cf. [3]). C'est un entier compris entre 1 et  $\text{Log}_2(n+1)+1$ , où  $n$  est la taille de l'arbre, calculé récursivement de la façon suivante (on l'appellera plus brièvement *ordre*):

**Définition 1.** L'ordre d'un sommet externe est 1, celui d'un sommet interne sera égal au plus grand des ordres de ses fils si leurs ordres diffèrent, à la valeur commune augmentée de 1 sinon.

Par extension on appellera *ordre* d'un arbre  $T$ , noté  $\text{ord}(T)$ , l'ordre de sa racine.

Viennot a défini dans [16] le *bi-ordre* d'un sommet interne comme la paire non ordonnée formée par les ordres de ses fils. Pour un sommet d'ordre  $k \geq 2$ , c'est soit  $(k, i)$  avec  $1 \leq i < k$ , soit  $(k-1, k-1)$ .

Pour un arbre  $T$  posons:

$$b_{k,i}^T = \text{nombre de sommets de bi-ordre } (k, i),$$

$$o_k^T = \text{nombre de sommets d'ordre } k.$$

Cet auteur a alors introduit, pour caractériser la forme d'un arbre, la *matrice de ramification*, matrice infinie, dont les coefficients sont donnés par les expressions suivantes: pour  $2 \leq k \leq \text{ord}(T)$ ,

$$R_{k,i}^T = \frac{b_{k,i}^T}{o_k^T}, \text{ pour } 1 \leq i < k, \text{ et } R_{k,k}^T = \frac{b_{k-1,k-1}^T}{o_k^T},$$

et pour  $k > \text{ord}(T)$ ,

$$R_{k,i}^T = 0 \text{ pour tout } i.$$

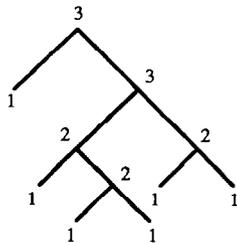
Figure 1 représente un arbre binaire et sa matrice de ramification.

Pour une famille finie d'arbres  $\mathcal{T}$ , on peut étendre cette notion de deux façons. Posons,

$$\text{ord}(\mathcal{T}) = \max_{T \in \mathcal{T}} (\text{ord}(T)).$$

**Définition 2.** On appelle *matrice de ramification d'une famille*  $\mathcal{T}$  d'arbres binaires la matrice  $P^{\mathcal{T}}$  suivante, pour  $2 \leq k \leq \text{ord}(\mathcal{T})$ ,

$$P_{k,i}^{\mathcal{T}} = \frac{\sum_{T \in \mathcal{T}} b_{k,i}^T}{\sum_{T \in \mathcal{T}} o_k^T}, \text{ pour } 1 \leq i < k, \text{ et } P_{k,k}^{\mathcal{T}} = \frac{\sum_{T \in \mathcal{T}} b_{k-1,k-1}^T}{\sum_{T \in \mathcal{T}} o_k^T},$$



		i		
k				
2	1/3	2/3		
3	1/2	0	1/2	
4	0	0	0	0

Fig. 1.

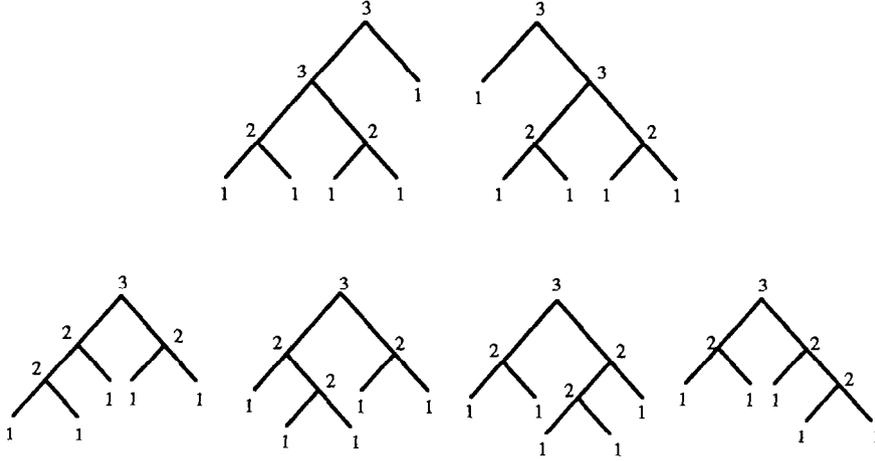


Fig. 2.

et pour  $k > \text{ord}(\mathcal{T})$ ,

$$P_{k,i}^{\mathcal{T}} = 0 \quad \text{pour tout } i.$$

**Définition 3.** On appelle *matrice de ramification d'un arbre aléatoire d'une famille*  $\mathcal{T}$  d'arbres binaires la matrice  $Q^{\mathcal{T}}$  telle que, pour tout couple d'entiers  $k$  et  $i$ ,  $k \geq 2$  et  $1 \leq i \leq k$ ,

$$Q_{k,i}^{\mathcal{T}} = \frac{1}{|\mathcal{T}|} \sum_{T \in \mathcal{T}} R_{k,i}^T.$$

Ces deux matrices sont généralement distinctes comme le montre l'exemple illustré par Fig. 2; seule la seconde permet de définir une matrice de ramification aléatoire comme moyenne des matrices de chacun des arbres qui composent la famille. Toutefois, dans le cas des arbres binaires, le résultat fondamental de ce papier est que leurs limites, lorsque la taille des arbres tend vers l'infini, sont égales. La première définition conduit à un calcul de limite plus simple, dans lequel on introduit des outils utiles pour le calcul de la seconde limite; de plus, elle permet des calculs statistiques plus précis.

**Exemple.** Il y a 6 arbres binaires complets de taille 4 et d'ordre 3. Parmi eux, 2, représentés en haut de la Fig. 2, ont 2 sommets d'ordre 3 et sont tels que,

$$R_{3,1} = R_{3,3} = 1/2, \quad R_{3,2} = 0,$$

et les 4 autres, sur le bas de la figure, ont un seul sommet d'ordre 3; la ligne 3 de leur matrice de ramification est,

$$R_{3,1} = R_{3,2} = 0, \quad R_{3,3} = 1.$$

Pour cet ensemble d'arbres, les lignes 3 des matrices de ramification sont donc respectivement,

$$\text{pour } P: \quad 2/8, 0, 6/8,$$

et

$$\text{pour } Q: \quad 1/6, 0, 5/6.$$

Il est classique de coder les arbres binaires par le langage de Łukasiewicz  $B$  sur l'alphabet  $\mathcal{A} = \{x, \bar{x}\}$ . Ce langage vérifie l'équation en variables non commutatives,

$$B = \bar{x} + xBB,$$

qui reflète la définition récursive des arbres,  $x$  codant un sommet interne et  $\bar{x}$  un sommet externe. Cette équation est non ambiguë. Pour énumérer les mots du langage  $B$  selon leur longueur, on applique le morphisme  $\xi$  défini par  $\xi(x) = x$ ,  $\xi(\bar{x}) = 1$ , puis en prenant l'image commutative, on obtient l'équation à une inconnue  $\mathcal{B}$  dans  $\mathbb{C}[[x]]$ ,

$$\mathcal{B} = 1 + x\mathcal{B}\mathcal{B},$$

dont la solution analytique à l'origine,

$$\mathcal{B}(x) = \frac{1 - \sqrt{1 - 4x}}{2x},$$

est la série énumératrice des arbres binaires selon le nombre de sommets internes. Le coefficient de  $x^n$  est le *nombre de Catalan* bien connu  $c_n$ , qui a pour expression,

$$c_n = \frac{1}{n+1} \binom{2n}{n}. \quad (1)$$

Dans ce travail où nous voulons énumérer les sommets ayant un bi-ordre donné  $(k, j)$  dans la famille des arbres binaires, nous appliquerons la même méthode issue des travaux de Schützenberger [12]. Pour cela, nous utilisons les notations suivantes:

**Notations.** (1) Un alphabet  $\mathcal{A}_k$  à  $k+2$  lettres,

$$\mathcal{A}_k = \{x, \bar{x}, x_1, x_2, \dots, x_k\},$$

où:

- $\bar{x}$  code un sommet externe,
- $x$ , un sommet interne d'ordre différent de  $k$ ,
- $x_i$  ( $2 \leq i < k$ ), un sommet interne de bi-ordre  $(k, i)$ ,
- $x_k$ , un sommet interne de bi-ordre  $(k-1, k-1)$ .

(2) Les langages ci-après,  $L_i$  le langage codant les arbres d'ordre  $i$  ( $1 \leq i < k$ ),  $\hat{L}_k$  celui codant les arbres d'ordre exactement  $k$ ,  $\hat{L}_{>k}$ , celui codant les arbres d'ordre supérieur à  $k$  et  $\hat{M}_k$  le langage codant tous les arbres binaires.

Nous déduisons de ce codage le système suivant d'équations algébriques non ambiguës qui traduit la définition récursive des arbres binaires d'un ordre donné (le caractère  $\hat{\phantom{x}}$  signalant dans un langage la présence des lettres indicées).

$$S_k = \begin{cases} L_1 = \bar{x} \\ L_2 = x(L_1 L_2 + L_2 L_1) + x(L_1)^2 \\ \vdots \\ L_{k-1} = x \sum_{1 \leq j < k-1} (L_j L_{k-1} + L_{k-1} L_j) + x(L_{k-2})^2 \\ \hat{L}_k = \sum_{1 \leq j < k} x_j (L_j \hat{L}_k + \hat{L}_k L_j) + x_k (L_{k-1})^2 \\ \hat{L}_{>k} = x \sum_{1 \leq j < k} (L_j \hat{L}_{>k} + \hat{L}_{>k} L_j) + x(\hat{L}_k \hat{L}_{>k} + \hat{L}_{>k} \hat{L}_k) + x(\hat{L}_k)^2 + x(\hat{L}_{>k})^2 \\ \hat{M}_k = \sum_{1 \leq j < k} L_j + \hat{L}_k + \hat{L}_{>k} \end{cases}$$

Les mots du langage  $\hat{M}_k$ , dernière composante de la solution du système  $S_k$ , codent donc tous les arbres binaires, et le nombre de lettres  $x_i$  dans un mot est égal au nombre de sommets de bi-ordre  $(k, i)$  dans l'arbre codé par ce mot.

### 3. Matrices de ramification de $\mathcal{F}(n)$

Appelons  $\mathcal{F}(n)$  la famille des arbres binaires complets de taille  $n$ . Nous allons dans ce paragraphe relier les entrées  $(k, j)$  des deux matrices  $P^{\mathcal{F}(n)}$  et  $Q^{\mathcal{F}(n)}$  aux coefficients de deux familles de séries entières  $\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$ . Ces séries à leur tour se déduisent simplement des séries  $\mathcal{M}_k(x)$  et  $\mathcal{N}_k(x)$  composantes de la solution de deux systèmes d'équations algébriques en variables commutatives  $\mathcal{P}_k$  et  $\mathcal{P}'_k$ , qui se déduisent de  $S_k$ .

#### 3.1. Calcul de $P^{\mathcal{F}(n)}$

Soit

$$\mathcal{U}_{k,j}(x) = \sum_{n>0} u_{n,k,j} x^n,$$

la série énumératrice à une variable où  $u_{n,k,j}$  est le nombre de sommets d'ordre  $k$  et de bi-ordre  $(k, j)$ , pour  $j < k$ , ou de bi-ordre  $(k-1, k-1)$  pour  $j = k$ , parmi tous les arbres binaires complets de taille  $n$ . De la Définition 2 et de la remarque que,

$$\sum_{T \in \mathcal{F}(n)} o_k^T = \sum_{j=1,k} u_{n,k,j},$$

on déduit immédiatement la proposition,

**Proposition 4.** *Pour tout triple d'entiers  $n, k, j$  tels que  $2 \leq k \leq \lfloor \log_2(n+1) \rfloor + 1$  et  $1 \leq j \leq k$ , on a*

$$P_{k,j}^{\mathcal{F}(n)} = \frac{u_{n,k,j}}{\sum_{i=1,k} u_{n,k,i}}. \quad (2)$$

Les séries  $\mathcal{U}_{k,j}(x)$  ne sont pas directement calculables. Considérons alors les séries à  $k+1$  variables, qui énumèrent les mots du langage  $\hat{\mathcal{M}}_k$ ,

$$\mathcal{M}_k(x, x_1, x_2, \dots, x_k) = \sum_{p, i_1, i_2, \dots, i_k} a(p, i_1, i_2, \dots, i_k) x^p x_1^{i_1} x_2^{i_2} \dots x_k^{i_k}$$

où  $a(p, i_1, i_2, \dots, i_k)$  est le nombre d'arbres d'ordre supérieur ou égal à  $k$ , ayant  $p + i_1 + i_2 + \dots + i_k$  sommets internes dont  $p$  d'ordre différent de  $k$ , et  $i_1 + i_2 + \dots + i_k$  d'ordre  $k$ ; parmi ceux-ci,  $i_j$  de bi-ordre  $(k, j)$  ( $1 \leq j \leq k$ ) et  $i_k$  de bi-ordre  $(k-1, k-1)$ .

Clairement on a,

$$u_{n,k,j} = \sum_{\substack{p, i_1, i_2, \dots, i_k \\ \text{entiers tels que} \\ p + i_1 + i_2 + \dots + i_k = n}} i_j a(p, i_1, i_2, \dots, i_k),$$

d'où la proposition.

**Proposition 5.** *Pour tout couple d'entiers  $k, j$  tels que  $k \geq 2$  et  $1 \leq j \leq k$ , on a*

$$\mathcal{U}_{k,j}(x) = x \frac{\partial \hat{\mathcal{M}}_k}{\partial x_j}(x, x, \dots, x).$$

Pour calculer la série  $\hat{\mathcal{M}}_k(x)$ , nous utilisons un système d'équations commutatives déduit du système  $\mathcal{S}_k$  et l'on peut énoncer,

**Proposition 6.** *La série  $\hat{\mathcal{M}}_k(x)$  est composante de la solution du système  $\mathcal{P}_k$  suivant,*

$$\mathcal{P}_k = \left\{ \begin{array}{l} \mathcal{L}_1 = 1 \\ \mathcal{L}_2 = x(\mathcal{L}_1)^2/d_1 \\ \mathcal{L}_3 = x(\mathcal{L}_2)^2/d_2 \\ \vdots \\ \mathcal{L}_{k-1} = x(\mathcal{L}_{k-2})^2/d_{k-2} \\ \hat{\mathcal{L}}_k = x_k(\mathcal{L}_{k-1})^2/\hat{D}_{k-1} \\ \hat{\mathcal{L}}_{>k} = \frac{\hat{d}_k - \sqrt{\hat{\Delta}_k}}{2x} \\ \hat{\mathcal{M}}_k = \sum_{1 \leq i < k} \mathcal{L}_i + \hat{\mathcal{L}}_k + \hat{\mathcal{L}}_{>k} \end{array} \right. \quad (3)$$

où,

$$d_i = 1 - 2x \sum_{j=1}^i \mathcal{L}_j, \quad (4)$$

$$\hat{d}_i = 1 - 2x \left( \hat{\mathcal{L}}_i + \sum_{j=1}^{i-1} \mathcal{L}_j \right) = d_{i-1} - 2x \hat{\mathcal{L}}_i, \quad (5)$$

$$\tilde{D}_i = 1 - 2 \sum_{j=1}^i x_j \mathcal{L}_j, \quad (6)$$

$$\hat{\Delta}_k = \hat{d}_k^2 - 4x^2 (\hat{\mathcal{L}}_k)^2. \quad (7)$$

**Démonstration.** En appliquant le morphisme  $\varphi$  défini par  $\varphi(x) = x$ ,  $\varphi(\bar{x}) = \mathbb{1}$  (le mot vide), et  $\varphi(x_i) = x_i$  pour  $1 \leq i \leq k$ , puis en prenant l'image commutative du système  $S_k$ , nous obtenons le système d'équations  $\mathcal{S}_k$  sur  $\mathbb{C}[[x, x_1, x_2, \dots, x_k]]$ , dont la solution est constituée des séries  $\mathcal{L}_i$  ( $1 \leq i < k$ )  $\hat{\mathcal{L}}_k$ ,  $\tilde{\mathcal{L}}_{>k}$ , et  $\hat{\mathcal{M}}_k$  énumérant les arbres binaires en fonction de l'ordre, du nombre de sommets de bi-ordre  $(k, i)$  et de la taille.  $\square$

**Remarque.** Dans ce système, nous indiquons encore la présence des variables indicées  $x_i$  ( $1 \leq i \leq k$ ) dans une composante de la solution, par le caractère  $\hat{\cdot}$ .

Le système réduit aux  $k - 1$  premières équations en une variable  $x$  a été étudié par divers auteurs (cf. Flajolet [4] et de Vauchaussade de Chaumont [15]). Il admet pour solutions les séries qui énumèrent les arbres d'ordre  $p$  en fonction de la taille. Ce sont des fractions rationnelles qui peuvent s'écrire:

$$\mathcal{L}_p(x) = \frac{x^{2^p - 1}}{\mathcal{F}_{2^p}(x)}, \quad (8)$$

où  $\mathcal{F}_n(z)$  sont les polynômes de Fibonacci, introduits par Kreweras [10],

$$\mathcal{F}_n(z) = \sum_{j \geq 0} \binom{n-1-j}{j} (-z)^j. \quad (9)$$

Rappelons encore que  $\mathcal{F}_n(t)$  désigne le polynôme réciproque du polynôme  $\mathcal{C}_n(t) = \mathcal{U}_n(t/2)$ , avec  $\mathcal{U}_n(t)$  polynôme de Tchebychef de 2<sup>o</sup> espèce, défini par  $\sin(n+1)\varphi = \sin \varphi \mathcal{U}_n(\cos \varphi)$  (cf. [4]).

Il n'est pas connu d'expression simple de  $\hat{\mathcal{M}}_k(x)$ , toutefois on peut établir,

**Proposition 7.** On a, pour  $k \geq 2$  et  $1 \leq j < k$ ,

$$\mathcal{U}_{k,j}(x) = \frac{2x \mathcal{L}_k \mathcal{L}_j}{\sqrt{1-4x}}, \quad \mathcal{U}_{k,k}(x) = \frac{x \mathcal{L}_{k-1}^2}{\sqrt{1-4x}}. \quad (10)$$

**Démonstration.** D'après la Proposition 5 ce calcul nécessite celui des dérivées partielles de  $\hat{\mathcal{M}}_k(x, x_1, x_2, \dots, x_k)$  au point  $(x, x, x, \dots, x)$ .

On a, pour  $1 \leq j \leq k$ ,

$$\begin{aligned}\frac{\partial \hat{\mathcal{M}}_k}{\partial x_j} &= \frac{\partial \hat{\mathcal{L}}_k}{\partial x_j} + \frac{\partial \hat{\mathcal{L}}_{>k}}{\partial x_j}, \\ \frac{\partial \hat{\mathcal{L}}_{>k}}{\partial x_j} &= \frac{1}{2x} \left[ \frac{\partial \hat{d}_k}{\partial x_j} - \frac{1}{2} \frac{\partial \hat{\Delta}_k}{\partial x_j} \frac{1}{\sqrt{\hat{\Delta}_k}} \right],\end{aligned}$$

et enfin,

$$\frac{\partial \hat{d}_k}{\partial x_j} = -2x \frac{\partial \hat{\mathcal{L}}_k}{\partial x_j}.$$

D'après les relations (5) et (7), on a,

$$\hat{\Delta}_k = \hat{d}_k^2 - 4x^2 \hat{\mathcal{L}}_k^2 = d_{k-1}(\hat{d}_k - 2x \hat{\mathcal{L}}_k) = d_{k-1}(d_{k-1} - 4x \hat{\mathcal{L}}_k)$$

et comme  $d_{k-1}$  ne contient que la variable  $x$ , on obtient:

$$\frac{\partial \hat{\Delta}_k}{\partial x_j} = -4x d_{k-1} \frac{\partial \hat{\mathcal{L}}_k}{\partial x_j}$$

d'où

$$\frac{\partial \hat{\mathcal{M}}_k}{\partial x_j} = \frac{d_{k-1}}{\sqrt{\hat{\Delta}_k}} \frac{\partial \hat{\mathcal{L}}_k}{\partial x_j} \quad (1 \leq j \leq k).$$

En dérivant l'expression de  $\hat{\mathcal{L}}_k$  dans le système  $\mathcal{S}_k$  (3), il vient, pour  $1 \leq j < k$ ,

$$\frac{\partial \hat{\mathcal{M}}_k}{\partial x_j} = \frac{d_{k-1}}{\sqrt{\hat{\Delta}_k}} \frac{2x_k \mathcal{L}_j (\mathcal{L}_{k-1})^2}{(\hat{D}_{k-1})^2} = \frac{d_{k-1}}{\sqrt{\hat{\Delta}_k}} \frac{2\mathcal{L}_j \hat{\mathcal{L}}_k}{\hat{D}_{k-1}}$$

et pour  $k$ ,

$$\frac{\partial \hat{\mathcal{M}}_k}{\partial x_k} = \frac{d_{k-1}}{\sqrt{\hat{\Delta}_k}} \frac{\mathcal{L}_{k-1}^2}{\hat{D}_{k-1}}.$$

La preuve s'achève en appliquant le lemme suivant.  $\square$

**Lemme 8.** *Pour tout entier  $k$  positif,  $\Delta_k$  a une expression indépendante de  $k$  que l'on appellera  $\Delta$ , et qui s'écrit,*

$$\Delta_k = \Delta = 1 - 4x. \quad (11)$$

**Démonstration.** Lorsque l'on se place au point  $X = (x, x, \dots, x)$ , on confond toutes les variables et alors

$$\hat{\mathcal{L}}_k = \mathcal{L}_k \quad \text{et} \quad \hat{d}_k = \hat{D}_k = d_k.$$

Dans ces conditions  $\Delta_k$  ne dépend plus de  $k$ :

$$\begin{aligned}
\Delta_k &= d_k^2 - 4x^2 \mathcal{L}_k^2, \\
&= d_{k-1}(d_{k-1} - 4x \mathcal{L}_k), \\
&= d_{k-1}^2 - 4x^2 \mathcal{L}_{k-1}^2, \\
&= \Delta_{k-1},
\end{aligned} \tag{12}$$

soit,

$$\Delta_k = \Delta_1 = 1 - 4x. \quad \square$$

### 3.2. Calcul de $Q^{\mathcal{F}(n)}$

Soit, d'autre part,

$$\mathcal{V}_{k,j}(x) = \sum_{n>0} v_{n,k,j} x^n$$

la série où  $v_{n,k,j}$  est la somme, pour tous les arbres de taille  $n$ , de la probabilité d'un sommet d'ordre  $k$  d'avoir le bi-ordre  $(k, j)$ . On a donc, par définition,

$$v_{n,k,j} = \sum_{T \in \mathcal{F}(n)} \frac{b_{k,j}^T}{o_k^T}.$$

Comme  $c_n$ , défini par relation (1), est le nombre d'arbres de taille  $n$ , on en déduit,

**Proposition 9.** *Pour tout triplet d'entiers  $n, k, j$  tels que  $2 \leq k \leq \lfloor \log_2(n+1) \rfloor + 1$  et  $1 \leq j \leq k$ , on a,*

$$Q_{k,j}^{\mathcal{F}(n)} = \frac{v_{n,k,j}}{c_n}. \tag{13}$$

Puis on construit la série énumératrice à  $k+2$  variables,

$$\tilde{\mathcal{M}}_k(x, x_1, x_2, \dots, x_k, y) = \hat{\mathcal{M}}_k(x, x_1 y, x_2 y, \dots, x_k y),$$

égale encore à,

$$\begin{aligned}
&\tilde{\mathcal{N}}_k(x, x_1, x_2, \dots, x_k, y) \\
&= \sum_{p, i_1, i_2, \dots, i_k} a(p, i_1, i_2, \dots, i_k) x^p x_1^{i_1} x_2^{i_2} \dots x_k^{i_k} y^{i_1 + i_2 + \dots + i_k},
\end{aligned}$$

l'exposant de la variable  $y$  introduite ici comptant le nombre total de sommets d'ordre  $k$ .

Pour tout arbre  $T$  admettant la suite d'entier  $(i_j)$  ( $1 \leq j \leq k$ ) comme distribution de ses bi-ordres  $(k, j)$ , on a,

$$b_{k,j}^T = i_j \quad \text{et} \quad o_k^T = i_1 + i_2 + \dots + i_k,$$

et si l'on regroupe dans la somme définissant  $v_{n,k,j}$  les arbres ayant même distribution de sommets de bi-ordre  $(k, j)$ , on obtient,

$$v_{n,k,j} = \sum_{\substack{p, i_1, i_2, \dots, i_k \\ p+i_1+i_2+\dots+i_k=n}} \frac{i_j}{i_1+i_2+\dots+i_k} a(p, i_1, i_2, \dots, i_k).$$

De cette relation entre coefficients, on déduit la relation entre les séries, d'où la

**Proposition 10.** *Pour tous entiers  $k$  et  $j$  tels que  $k \geq 2$  et  $1 \leq j \leq k$ , on a,*

$$\mathcal{V}_{k,j}(x) = x \int_0^1 \frac{1}{y} \frac{\partial \tilde{\mathcal{N}}_k}{\partial x_j}(x, x, x, \dots, x, y) dy.$$

Comme dans le cas de  $\hat{\mathcal{M}}_k(x)$ ,  $\tilde{\mathcal{N}}_k(x)$  est composante de la solution d'un système algébrique que l'on déduit de  $S_k$  défini en (3), et l'on énonce,

**Proposition 11.** *La série  $\tilde{\mathcal{N}}_k(x)$  est composante de la solution du système  $\mathcal{S}'_k$ ,*

$$\mathcal{S}'_k = \begin{cases} \mathcal{L}_1 = 1 \\ \mathcal{L}_2 = x(\mathcal{L}_1)^2/d_1 \\ \mathcal{L}_3 = x(\mathcal{L}_2)^2/d_2 \\ \vdots \\ \mathcal{L}_{k-1} = x(\mathcal{L}_{k-2})^2/d_{k-2} \\ \tilde{\mathcal{L}}_k = x_k(\mathcal{L}_{k-1})^2/\tilde{D}_{k-1} \\ \tilde{\mathcal{L}}_{>k} = \frac{\tilde{d}_k - \sqrt{\tilde{\Delta}_k}}{2x} \\ \tilde{\mathcal{N}}_k = \sum_{1 \leq i < k} \mathcal{L}_i + \tilde{\mathcal{L}}_k + \tilde{\mathcal{L}}_{>k} \end{cases} \quad (14)$$

où les nouvelles notations sont,

$$\tilde{d}_j = 1 - 2x \left( \sum_{i=1, j-1} \mathcal{L}_i + \tilde{\mathcal{L}}_j \right) = d_{j-1} - 2x\tilde{\mathcal{L}}_j, \quad (15)$$

$$\tilde{D}_j = 1 - 2 \sum_{i=1, j} x_i y \mathcal{L}_i = 1 + y(\tilde{D}_j - 1), \quad (16)$$

$$\tilde{\Delta}_j = \tilde{d}_j^2 - 4x^2(\tilde{\mathcal{L}}_j)^2 = d_{j-1}(d_{j-1} - 4x\tilde{\mathcal{L}}_j). \quad (17)$$

**Démonstration.** Soit  $\mu$  le morphisme tel que  $\mu(x) = x$ ,  $\mu(\bar{x}) = 1$  et  $\mu(x_k) = x_k y$ . Le système  $\mathcal{S}'_k$  dont  $\tilde{\mathcal{N}}_k$  est une composante de la solution se déduit de  $\mathcal{S}_k$  par application de  $\mu$ ; les expressions des  $\mathcal{L}_i$ , pour  $1 \leq i \leq k-1$ , restent inchangées et l'on prend pour  $\tilde{\mathcal{L}}_{>k}$  la solution analytique à l'origine.  $\square$

**Remarques.** (1) L'accentuation  $\sim$  exprime ici que la fonction dépend à la fois des variables indicées  $x_i$ , et de  $y$ .

(2) Les  $k-1$  premières composantes de la solution de  $\mathcal{S}'_k$  sont les mêmes que pour  $\mathcal{S}_k$ .

Il reste à exprimer les séries  $\mathcal{V}_{k,j}(x)$  à l'aide du système  $\mathcal{S}'_k$ , par la proposition suivante,

**Proposition 12.** *Si l'on pose,*

$$\alpha_k = 4x\mathcal{L}_k, \quad \beta_k = 1 - d_{k-1}, \quad \gamma_k = \alpha_k + \beta_k, \quad (18)$$

$$\theta_k = \sqrt{\frac{1-\gamma_k}{1-\beta_k}} = \frac{\sqrt{1-4x}}{d_{k-1}}, \quad (19)$$

$$\delta_k = \frac{\theta_k - \sqrt{\gamma_k/\beta_k}}{\theta_k + \sqrt{\gamma_k/\beta_k}} \frac{1 + \sqrt{\gamma_k/\beta_k}}{1 - \sqrt{\gamma_k/\beta_k}}, \quad (20)$$

alors pour tout couple d'entiers  $k, j$  tels que  $2 \leq k$  et  $1 \leq j < k$ , on peut écrire,

$$\mathcal{V}_{k,j}(x) = -\frac{d_{k-1}\mathcal{L}_j}{\beta} \left( \theta_k - 1 + \frac{\alpha_k}{2\gamma_k} \sqrt{\frac{\gamma_k}{\beta_k}} \ln(\delta_k) \right), \quad (21)$$

$$\mathcal{V}_{k,k}(x) = d_{k-1}\mathcal{L}_k \frac{1}{\gamma_k} \sqrt{\frac{\gamma_k}{\beta_k}} \ln(\delta_k).$$

**Démonstration.** Du système  $\mathcal{S}'_k$  défini en (14), on déduit,

$$\frac{\partial \tilde{\mathcal{N}}_k}{\partial x_i} = \frac{\partial \tilde{\mathcal{L}}_k}{\partial x_i} + \frac{\partial \tilde{\mathcal{L}}_{>k}}{\partial x_i}.$$

De ce système  $\mathcal{S}'_k$  et de la relation (15) on tire,

$$\tilde{\mathcal{L}}_k + \tilde{\mathcal{L}}_{>k} = \frac{1}{2x} (d_{k-1} - \sqrt{\tilde{\Delta}_k}).$$

De plus,  $\tilde{\Delta}_k$  défini par la relation (17), peut s'écrire,

$$\tilde{\Delta}_k = d_{k-1} (d_{k-1} - 4x\tilde{\mathcal{L}}_k) = d_{k-1} \left( d_{k-1} - 4x \frac{x_k y (\mathcal{L}_{k-1})^2}{\tilde{D}_{k-1}} \right),$$

soit,

$$\tilde{\Delta}_k = d_{k-1} \left( d_{k-1} - 4x_k y \frac{d_{k-1}\mathcal{L}_k}{\tilde{D}_{k-1}} \right) = (d_{k-1})^2 \left( 1 - 4x_k y \frac{\mathcal{L}_k}{\tilde{D}_{k-1}} \right).$$

Calculons les dérivées partielles,

$$\frac{\partial \tilde{D}_{k-1}}{\partial x_j} = -2y\mathcal{L}_j, \quad \frac{\partial \tilde{D}_{k-1}}{\partial x_k} = 0$$

d'où, pour  $1 \leq j < k$ , on obtient,

$$\frac{\partial \tilde{\mathcal{N}}_k}{\partial x_j} = \frac{2d_{k-1}x_k \mathcal{L}_k \mathcal{L}_j}{x} \frac{y^2}{(\sqrt{1+y(\tilde{D}_{k-1}-1)})^3 \sqrt{1+y(\tilde{D}_{k-1}-1-4x_k \mathcal{L}_k)}},$$

$$\frac{\partial \tilde{\mathcal{N}}_k}{\partial x_k} = \frac{d_{k-1} \mathcal{L}_k}{x} \frac{y}{\sqrt{1+y(\tilde{D}_{k-1}-1)} \sqrt{1+y(\tilde{D}_{k-1}-1-4x_k \mathcal{L}_k)}}.$$

Evaluons ces fonctions au point  $(x, x, \dots, x, y)$  alors  $\tilde{D}_{k-1} = d_{k-1}$ . En utilisant les relations (18) il vient,

$$\frac{\partial \tilde{\mathcal{N}}_k}{\partial x_j}(x, x, \dots, x, y) = 2d_{k-1} \mathcal{L}_k \mathcal{L}_j \frac{y^2}{(\sqrt{1-\beta_k y})^3 \sqrt{1-\gamma_k y}},$$

$$\frac{\partial \tilde{\mathcal{N}}_k}{\partial x_k}(x, x, \dots, x, y) = \frac{d_{k-1} \mathcal{L}_k}{x} \frac{y}{\sqrt{1-\beta_k y} \sqrt{1-\gamma_k y}},$$

et, en reportant dans les expressions de la Proposition 10,

$$\mathcal{V}_{k,j}(x) = 2xd_{k-1} \mathcal{L}_k \mathcal{L}_j \int_0^1 \frac{y}{(\sqrt{1-\beta_k y})^3 \sqrt{1-\gamma_k y}} dy,$$

$$\mathcal{V}_{k,k}(x) = d_{k-1} \mathcal{L}_k \int_0^1 \frac{1}{\sqrt{1-\beta_k y} \sqrt{1-\gamma_k y}} dy.$$

Pour trouver la primitive de ces intégrales elliptiques, posons selon l'usage,

$$t(y) = \sqrt{\frac{1-\gamma_k y}{1-\beta_k y}}, \quad \theta_k = t(1),$$

soit,

$$\theta_k = \sqrt{\frac{1-\gamma_k}{1-\beta_k}} = \frac{\sqrt{1-4x}}{d_{k-1}}.$$

Les deux intégrales s'écrivent alors,

$$\mathcal{V}_{k,j}(x) = -4 \frac{xd_{k-1} \mathcal{L}_k \mathcal{L}_j}{\alpha_k} \int_1^{\theta} \frac{1-t^2}{\gamma_k - \beta_k t^2} dt,$$

$$\mathcal{V}_{k,k}(x) = -2d_{k-1} \mathcal{L}_k \int_1^{\theta} \frac{1}{\gamma_k - \beta_k t^2} dt.$$

Tous calculs faits, on obtient le résultat énoncé.  $\square$

#### 4. Etude asymptotique des matrices de ramification

Le calcul des limites de  $P_{k,j}^{\mathcal{A}(n)}$  et  $Q_{k,j}^{\mathcal{A}(n)}$  nécessite la connaissance du comportement de  $u_{n,k,j}$  et de  $v_{n,k,j}$  à l'infini. Pour cela, il est classique d'utiliser les

ressources de l'analyse complexe en considérant les séries  $\mathcal{U}_{k,j}(x)$ , et  $\mathcal{V}_{k,j}(x)$  comme des fonctions de la variable complexe  $x$  dans leur domaine de convergence. Ainsi, le théorème de Cauchy, appliqué par exemple à  $\mathcal{U}_{k,j}(x)$ , stipule que,

$$u_{n,k,j} = \frac{1}{2i\pi} \int_{\Gamma} \mathcal{U}_{k,j}(x) \frac{dx}{x^{n+1}}$$

où  $\Gamma$  est une courbe simple fermée dans le domaine d'analyticit  de  $\mathcal{U}_{k,j}(x)$  et qui entoure l'origine.

Nous allons montrer dans la Proposition 14 que les singularit s de  $\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$  sont r elles et de module sup rieur ou  gal   1/4. Les relations  tablies au paragraphe pr c dent permettent d' valuer les fonctions  $\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$  au voisinage de ce point. On  crira, par exemple pour la premi re,

$$\mathcal{U}_{k,j}(x) = \mathcal{U}_{k,j}^1(x) + O(\mathcal{U}_{k,j}^2(x)),$$

o   $\mathcal{U}_{k,j}^1(x)$  est une fonction  l mentaire dont le d veloppement est connu. Pour traiter l'incidence du terme en grand O sur  $u_{n,k,j}$ , on appliquera le *lemme de transfert*,  tabli par Flajolet dans [3], dont nous rappelons l' nonc .

**Notations.** (1) Dans la suite de l'article, l'expression  $(g(x), x^n)$  d signe le coefficient de  $x^n$  dans le d veloppement en s rie de  $g(x)$ .

(2) Si une fonction de la variable  $x$  a une expression  $f$  d pendant de plusieurs autres fonctions, on notera  $[f](x)$  la valeur de cette fonction au point  $x$ .

**Lemme 13** (dit lemme de transfert [3]). (1) Si  $g(x)$  est analytique dans le domaine

$$\mathcal{D}_1(a) = \{x: |x| \leq a, x \neq a\},$$

et si lorsque  $x$  tend vers  $a$    l'int rieur de  $\mathcal{D}_1$

$$g(x) = O(|a-x|^{-s})$$

avec  $s > 1$ , alors:

$$(g(x), x^n) = O(a^{-n} n^{s-1}).$$

(2) Si  $g(x)$  est analytique dans le disque indent ,

$$\mathcal{D}_2(a) = \{x: |x| \leq a + \lambda, \omega < |\text{Artg}(x-a)| < 2\pi\}$$

avec  $\lambda > 0$  et  $0 < \omega < \pi/2$ , et si lorsque  $x$  tend vers  $a$    l'int rieur de  $\mathcal{D}_2$ ,

$$g(x) = O(|a-x|^r),$$

avec  $r > 0$ , alors,

$$(g(x), x^n) = O(a^{-n} n^{-r-1}).$$

Le point 1/4 du plan complexe est clairement un point singulier pour les fonctions

$\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$  en raison du facteur  $\sqrt{1-4x}$ . La proposition suivante a pour objet de préciser le domaine d'analyticit  de ces fonctions.

**Proposition 14.** (1) Pour tout couple d'entiers  $(k, j)$  tel que  $k > 2$  et  $1 \leq j \leq k$ , les fonctions  $\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$  sont analytiques dans le domaine  gal au plan complexe priv  de la demi-droite des r els sup rieurs ou  gaux    $1/4$ , que l'on notera  $\mathbb{C} \setminus \mathbb{R}_{\geq 1/4}$ .

(2) De plus les fonctions  $\mathcal{L}_k(x)$ ,  $\alpha_k(x)$ ,  $\beta_k(x)$ ,  $d_{k-1}(x)$  et  $\gamma_k(x)$  sont analytiques dans un voisinage de  $1/4$  et valent en ce point,

$$\mathcal{L}_k(1/4) = \alpha_k(1/4) = d_{k-1}(1/4) = 2^{1-k}, \quad (22)$$

$$\beta_k(1/4) = 1 - 2^{1-k}, \quad \gamma_k(1/4) = 1. \quad (23)$$

**D monstration.** Elle se fait en trois  tapes, d'abord  tude pour  $|x| \neq 0$  et  $|x| \neq 1/4$ , puis au voisinage de l'origine et enfin au voisinage de  $1/4$ .

(a)  tude pour  $|x| \neq 0$  et  $|x| \neq 1/4$ .

La fa on la plus simple est d'utiliser les expressions trigonom triques obtenues en [4]. En faisant le changement de variable,

$$x = \frac{1}{2 \cos \varphi + 2} = \frac{1}{4 \cos^2 \varphi / 2},$$

on obtient les expressions, pour  $k \geq 2$ ,

$$\mathcal{L}_k = \frac{\sin \varphi}{\sin 2^{k-1} \varphi}. \quad (24)$$

On en d duit aussit t,

$$\beta_k = 2x \sum_{i=1}^{k-1} \mathcal{L}_i = \frac{1}{\cos \varphi / 2} \frac{\sin(2^{k-1} - 1)\varphi / 2}{\sin 2^{k-1} \varphi / 2},$$

et apr s quelques manipulations trigonom triques,

$$d_{k-1} = \frac{1}{4 \cos^2 \varphi / 2} \frac{\sin \varphi \sin 2^{k-1} \varphi}{(\sin 2^{k-1} \varphi / 2)^2},$$

$$\alpha_k = \frac{1}{\cos^2 \varphi / 2} \frac{\sin \varphi}{\sin 2^{k-1} \varphi},$$

$$\gamma_k = \frac{1}{d_{k-1}} \frac{\sin \varphi / 2 \cos(2^{k-1} - 1)\varphi / 2}{\cos^2 \varphi / 2 \sin 2^{k-1} \varphi / 2}.$$

On remarque que, pour toute d termination de  $\varphi$ , les p les et z ros  $\varphi$  de ces fonctions sont donc tous r els, et comme pour tout  $\varphi$  r el,

$$0 \leq 2 \cos \varphi + 2 \leq 4,$$

les z ros et p les en  $x$  sont donc r els et sup rieurs    $1/4$ , c'est- -dire dans  $\mathbb{R}_{\geq 1/4}$ .

De plus pour  $x \neq 0$ ,  $x \notin \mathbb{R}_{\geq 1/4}$ , les arguments des racines carrées et du logarithme sont définis et non nuls:

- pour  $\gamma_k/\beta_k$  et pour  $(1-\gamma_k)/(1-\beta_k)$  c'est évident d'après les expressions ci-dessus et (19);
- pour  $\delta_k$ , car d'après sa définition (20),

$$\{\delta_k = 0 \text{ ou } 1/\delta_k = 0\} \Rightarrow \beta_k = \gamma_k \Rightarrow \alpha_k = 0,$$

ce qui est impossible dans ce domaine et donc ces fonctions sont uniformes et analytiques pour  $x \neq 0$ ,  $x \notin \mathbb{R}_{\geq 1/4}$ .

(b) Étude au voisinage de l'origine.

Pour  $k \geq 2$ , on déduit de (8) et (9) que  $\mathcal{L}_k(0) = 0$ , ce qui implique d'après la Proposition 7 que les fonctions  $\mathcal{U}_{k,j}(x)$  sont définies et nulles en zéro, pour  $1 \leq j \leq k$ . Par contre les fonctions  $\mathcal{V}_{k,j}(x)$  calculées par la Proposition 12, contiennent des dénominateurs nuls pour ce point. Il convient donc de vérifier qu'elles ont une limite en ce point afin d'étendre par continuité leur définition. De (8) et (9) on obtient l'expression suivante de  $\mathcal{L}_k$  valable dans un voisinage de zéro,

$$\mathcal{L}_k(x) = x^{2^k-1} (1 + (2^k - 2)x + O(|x|)),$$

d'où l'on tire immédiatement que les fonctions définies par les relations (18), (19), et (4), vérifient,

$$\alpha_k(x) = 4x^{2^k-1} (1 + (2^k - 2)x + O(|x|^2)),$$

$$d_k(x) = 1 - 2x(1 + x + O(|x|^2)),$$

$$\beta_2(x) = 2x, \quad \beta_k(x) = 2x(1 + x + O(|x|^2)) \quad (k \geq 3),$$

et en définissant la fonction  $\varepsilon(k)$  par  $\varepsilon(2) = 1$ ,  $\varepsilon(k) = 0$  pour  $k \geq 3$ ,

$$\gamma_k(x) = 2x(1 + (1 + \varepsilon(k))x + O(|x|^2)),$$

$$\theta_k(x) = 1 - 2x^{2^k-1} (1 + O(|x|)),$$

d'où, pour un voisinage de 0, 0 exclu,

$$\left[ \frac{\alpha_k}{\beta_k} \right] (x) = 2x^{2^k-1} (1 + (2^k - 3 + \varepsilon(k))x + O(|x|)),$$

et par continuité, pour tout  $k \geq 2$ , on peut poser,

$$\left[ \frac{\alpha_k}{\beta_k} \right] (0) = 0, \quad \left[ \frac{\gamma_k}{\beta_k} \right] (0) = 1.$$

De même on déduit des développements de Taylor précédents la valeur suivante de la fonction  $\delta_k$  définie par (20), valable dans un voisinage de 0, 0 exclu,

$$\delta_k(x) = 1 + 2x + O(|x|^2).$$

On peut donc définir par continuité, au point zéro, pour tout  $k \geq 2$ ,  $\delta_k(0) = 1$ .

Il résulte de cette étude que les fonctions

$$\left[ \sqrt{\frac{\gamma_k}{\beta_k}} \right] (x), \quad \theta_k(x), \quad \delta_k(x) \quad \text{et} \quad \ln(\delta_k(x)),$$

sont donc bien uniformes et analytiques dans un voisinage de zéro. Il suffit alors de développer l'expression suivante, pour  $k \geq 2$ , dans un voisinage de zéro,

$$\left[ \theta_k - 1 + \frac{\alpha_k}{2\beta_k} \sqrt{\frac{\gamma_k}{\beta_k}} \ln(\delta_k) \right] (x) = -2x^{2^k-1} + O(|x|^{2^k-1+2}),$$

et, en reportant ces développements dans (21), on obtient, dans un voisinage de l'origine, zéro exclu, pour  $k \geq 2$  et  $1 \leq j < k$ ,

$$\mathcal{V}_{k,j}(x) = 2x^{2^k-1+2^{j-1}-1}(1 + O(|x|)),$$

$$\mathcal{V}_{k,k}(x) = x^{2^k-1-1}(1 + O(|x|)).$$

On pose alors par continuité, pour  $k \geq 2$  et  $1 \leq j \leq k$ ,

$$\mathcal{V}_{k,j}(0) = 0, \quad \mathcal{V}_{k,k}(0),$$

et la première partie de la proposition est démontrée.

**Remarque.** Les fonctions  $\mathcal{U}_{k,j}(x)$  et  $\mathcal{V}_{k,j}(x)$  sont donc équivalentes dans un voisinage de zéro, mais je ne connais pas de preuve plus directe de ce fait.

(c) Étude dans un voisinage de  $1/4$ .

Il est clair que pour  $k$  fixé, les fonctions  $\mathcal{L}_k(x)$ ,  $\alpha_k(x)$ ,  $\beta_k(x)$ ,  $\gamma_k(x)$  et  $d_{k-1}(x)$  sont définies dans un voisinage de  $1/4$ ,  $1/4$  inclus. En appliquant la règle de l'Hospital à la relation (24) on obtient, pour  $k \geq 2$ ,

$$\mathcal{L}_k(1/4) = 2^{1-k},$$

d'où d'après (18) et (12),

$$\alpha_k(1/4) = d_{k-1}(1/4) = 2^{1-k},$$

et,

$$\beta_k(1/4) = 1 - 2^{1-k}, \quad \gamma_k(1/4) = 1,$$

ce qui achève la preuve de la proposition.  $\square$

Dans les deux propositions suivantes, nous appliquons le lemme de transfert avec comme domaine  $\mathcal{D}(1/4) = \{x: |x| \leq 1/4 + \lambda, \omega < |\text{Arctg}(x - 1/4)| < 2\pi\}$ , avec  $\lambda > 0$  et  $0 < \omega < \pi/2$ ,  $x \neq 1/4$ , contenant le contour d'intégration représenté sur Fig. 3.

**Proposition 15.** Lorsque  $n$  tend vers l'infini, les coefficients des séries  $\mathcal{U}_{k,j}$  pour  $k \geq 2$  et  $1 \leq j \leq k$ , vérifient,

$$\lim_{n \rightarrow \infty} \frac{u_{n,k,j}}{\sum_{1 \leq i < k} u_{n,k,i} + u_{n,k,k}} = \frac{1}{2^j}, \quad \lim_{n \rightarrow \infty} \frac{u_{n,k,k}}{\sum_{1 \leq i < k} u_{n,k,i} + u_{n,k,k}} = \frac{1}{2^{k-1}}.$$

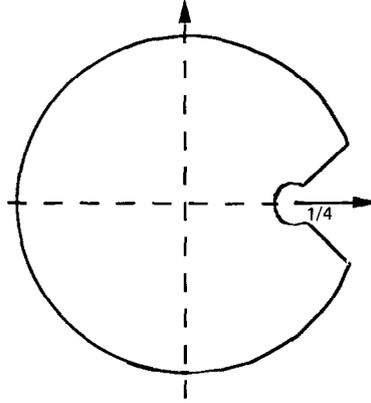


Fig. 3.

**Démonstration.** Au voisinage de la plus petite singularité  $1/4$ , les fonctions  $\mathcal{L}_i$  sont régulières et admettent un développement de Taylor; on en déduit un développement de  $\mathcal{U}_{k,j}(x)$  à l'intérieur de  $\mathcal{D}_2$ , soit, pour  $1 \leq j < k$ ,

$$\mathcal{U}_{k,j}(x) = \frac{1}{2} \mathcal{L}_k(1/4) \mathcal{L}_j(1/4) \frac{1}{\sqrt{1-4x}} + O(\sqrt{|1-4x|}),$$

$$\mathcal{U}_{k,k}(x) = \frac{1}{4} (\mathcal{L}_{k-1}(1/4))^2 \frac{1}{\sqrt{1-4x}} + O(\sqrt{|1-4x|}).$$

Nous sommes alors dans le second cas du lemme de transfert ce qui donne, en utilisant les valeurs au point  $1/4$ , calculées dans la Proposition 14,

$$u_{n,k,j} = 2^{1-k-j} \left( \frac{1}{\sqrt{1-4x}}, x^n \right) + O\left( \frac{4^n}{n\sqrt{n}} \right), \quad \text{pour } 1 \leq j < k,$$

$$u_{n,k,k} = 2^{2-2k} \left( \frac{1}{\sqrt{1-4x}}, x^n \right) + O\left( \frac{4^n}{n\sqrt{n}} \right).$$

Or il est bien connu que dans le développement de Newton de  $(1-4x)^{-1/2}$ , on a,

$$\left( \frac{1}{\sqrt{1-4x}}, x^n \right) = \frac{(2n)!}{(n!)^2},$$

dont un équivalent est donné par la formule de Stirling,

$$\left( \frac{1}{\sqrt{1-4x}}, x^n \right) \approx \frac{4^n}{\sqrt{\pi n}}.$$

Donc il l'emporte sur le terme en  $O$ , et l'on peut écrire,

$$u_{n,k,j} \approx 2^{1-k-j} \left( \frac{1}{\sqrt{1-4x}}, x^n \right), \quad 1 \leq k < j,$$

$$u_{n,k,k} \approx 2^{2-2k} \left( \frac{1}{\sqrt{1-4x}}, x^n \right),$$

et la limite,

$$\lim_{n \rightarrow \infty} \frac{u_{n,k,j}}{\sum_{1 \leq i < k} u_{n,k,i} + u_{n,k,k}} = \frac{2^{1-k-j}}{\sum_{1 \leq i < k} 2^{1-k-i} + 2^{2-2k}} = \frac{1}{2^j}.$$

En recommençant le même calcul pour  $u_{n,k,k}$  on achève la preuve.  $\square$

De même, pour les séries  $\mathcal{V}_{k,j}(x)$ , on a la proposition suivante,

**Proposition 16.** *Lorsque  $n$  tend vers l'infini, les coefficients des séries  $\mathcal{V}_{k,j}(x)$  sont, pour  $k \geq 2$  et  $1 \leq j < k$  équivalents à,*

$$v_{n,k,j} \approx {}^{-j}c_n, \quad v_{n,k,k} \approx 2^{1-k}c_n,$$

où  $c_n$  désigne le nombre de Catalan défini par la relation (1).

**Démonstration.** Au voisinage du point  $1/4$ , la fonction  $[\theta_k \sqrt{\beta_k/\gamma_k}](x)$  est définie et voisine de zéro; en développant le logarithme  $\ln(\delta_k)$ , on a,

$$\ln(\delta_k) = -2\theta_k \sqrt{\frac{\beta_k}{\gamma_k}} + \mathcal{O}\left(\left|\theta_k \sqrt{\frac{\beta_k}{\gamma_k}}\right|^3\right) + \varrho,$$

où  $\varrho$  est une fonction analytique en  $1/4$ . En développant le premier terme de cette expression au voisinage de  $1/4$ , on peut alors écrire,

$$\ln(\delta_k(x)) = -2 \left[ \frac{1}{d_{k-1}} \sqrt{\frac{\beta_k}{\gamma_k}} \right] (1/4) \times \sqrt{1-4x} + \mathcal{O}(\sqrt{|1-4x|^3}) + \varrho(x).$$

En reportant cette valeur dans l'expression (21) de  $\mathcal{V}_{k,j}(x)$  établie dans la Proposition 12 et en remarquant que  $\gamma_k - \alpha_k = \beta_k$ , on a,

$$\mathcal{V}_{k,j}(x) = \left[ \frac{\mathcal{L}_j}{\gamma_k} \right] (1/4) \times \sqrt{1-4x} + \mathcal{O}(|\sqrt{1-4x}|^3) + \varphi_1(x),$$

où  $\varphi_1(x)$  regroupe tous les termes réguliers en  $1/4$ ; en développant cette dernière selon Taylor, on obtient finalement,

$$\mathcal{V}_{k,j}(x) = \varphi_1(1/4) - \left[ \frac{\mathcal{L}_j}{\gamma_k} \right] (1/4) \times \sqrt{1-4x} + \mathcal{O}(|1-4x|).$$

La même démarche appliquée à  $\mathcal{V}_{k,k}(x)$  donne

$$\mathcal{V}_{k,k}(x) = \varphi_2(1/4) - 2 \left[ \frac{\mathcal{L}_k}{\gamma_k} \right] (1/4) \times \sqrt{1-4x} + \mathcal{O}(|1-4x|),$$

où  $\varphi_2(x)$  regroupe ici encore les termes réguliers en  $1/4$ .

Nous sommes une nouvelle fois dans le second cas du lemme de transfert (avec  $a=1/4$  et  $r=1$ ) ce qui donne, en évaluant les expressions au point  $1/4$ ,

$$\begin{aligned} v_{n,k,j} &= 2^{1-j}(-\sqrt{1-4x}, x^n) + O(4^n n^{-2}), \quad 1 \leq j < k, \\ v_{n,k,k} &= 2^{2-k}(-\sqrt{1-4x}, x^n) + O(4^n n^{-2}). \end{aligned}$$

Or le développement de Newton de  $-\sqrt{1-4x}$  donne,

$$(-\sqrt{1-4x}, x^n) = \frac{2(2n-2)!}{n!(n-1)!} \simeq 2 \frac{4^{n-1}}{\sqrt{\pi n^3}},$$

coefficient très voisin des nombres de Catalan  $c_n$ ,

$$(-\sqrt{1-4x}, x^n) = 2 \left( \frac{1-\sqrt{1-4x}}{2x}, x^n \right) = 2c_{n-1};$$

comme pour  $n$  grand,  $c_n \simeq 4c_{n-1}$ , on obtient les expressions de la proposition.  $\square$

Nous sommes alors en mesure d'établir le théorème principal de ce papier, qui se déduit immédiatement des deux propositions précédentes appliquées aux relations (2) et (13).

**Théorème 17.** (1) *Pour tous entiers  $k$  et  $j$  tels que  $k \geq 2$  et  $1 \leq j \leq k$ ,*

$$\lim_{n \rightarrow \infty} P_{k,j}^{\mathcal{F}(n)} \quad \text{et} \quad \lim_{n \rightarrow \infty} Q_{k,j}^{\mathcal{F}(n)}$$

*existent.*

(2) *Ces deux limites sont égales. Leur valeur commune est:*

$$R_{k,j} = \frac{1}{2^j}, \quad \text{pour } 1 \leq j < k, \quad \text{et} \quad R_{k,k} = \frac{1}{2^{k-1}}.$$

La matrice constituée par ces coefficients sera appelée *matrice de ramification des arbres binaires complets*.

## 5. Conclusion

Un autre approche pour établir ce résultat d'égalité serait l'étude des comportements limites des systèmes dynamiques dans lesquels les populations seraient les arbres évoluant suivant la taille, et les variables étudiées les probabilités d'un sommet d'avoir tel ordre, tel bi-ordre, tel quotient de ces deux paramètres. Il suffirait alors de montrer que ces variables sont faiblement mélangées, ce qui est cependant une condition plus forte.

Un autre objectif serait d'examiner les matrices de ramification pour différentes familles d'arbres, par exemple celle des tournois binaires, énumérés par  $n!$

**Bibliographie**

- [1] M.P. Delest, Utilisation des langages algébriques et du calcul formel pour le codage et l'énumération des polyominos, Thèse d'État, Université de Bordeaux I (1987).
- [2] A.P. Ershov, On programming of arithmetic operations, *Comm. ACM* 1 (8) (1958) 3-6.
- [3] P. Flajolet, Mathematical methods in the analysis of algorithms and data structures, INRIA Res. Rept. 400 (1985).
- [4] P. Flajolet, J. Françon et J. Vuillemin, The number of registers required for evaluating arithmetic expressions, *Theoret. Comput. Sci.* 9 (1979) 99-125.
- [5] P. Flajolet et A. Odlyzko, The average height of binary trees and other simple trees, *J. Comput. System. Sci.* 25 (1982) 171-213.
- [6] J. Françon, Sur le nombre de registres nécessaires à l'évaluation d'une expression arithmétique, *RAIRO Inform. Théor. Appl.* 18 (1984) 335-364.
- [7] R.E. Horton, Erosional development of streams and their drainage basins, hydrophysical approach to quantitative morphology, *Bull. Geol. Soc. Amer.* 56 (1945) 275-370.
- [8] R. Kemp, The average number of registers needed to evaluate a binary tree optimally, *Acta Inform.* 11 (1979) 363-372.
- [9] D.E. Knuth, *The Art of Computer Programming: Fundamental Algorithms* (Addison-Wesley, Reading, MA, 1968).
- [10] G. Kreweras, Sur les éventails de segments, *Cahiers BURO* 15 (1970) 1-41.
- [11] A. Meir, J.W. Moon et J.R. Ponder, On the order of random channel networks, *SIAM J. Algebraic Discrete Methods* 1 (1980) 25-32.
- [12] M.P. Schützenberger, Certain elementary families of automata, *Proceedings Symposium on Mathematical Theory of Automata*, Polytechnic Institute of Brooklyn (1962) 139-153.
- [13] A.N. Strahler, Hypsometric (area-altitude) analysis of erosional topology, *Bull. Geol. Soc. Amer.* 63 (1952) 1117-1142.
- [14] J. Vannimenus et X.G. Viennot, Combinatorials tools for the analysis of ramified patterns, *J. Statist. Phys.* 54 (1989) 1529-1538.
- [15] M. Vauchassade de Chaumont et X.G. Viennot, Enumeration of RNAs secondary structure by complexity, *Lecture Notes in Biomathematics* 37 (Springer, Berlin, 1985) 360-363.
- [16] X.G. Viennot, Trees, in: M.P. Schützenberger, D. Perrin et A. Lascoux, eds., *Mélanges* (1989).
- [17] X.G. Viennot, G. Eyrolles, N. Janey et D. Arques, Combinatorial analysis of ramified patterns and computer imagery of trees, *Proceedings of SIGGRAPH 89*, Boston (1989).