Quelques techniques analytiques pour l'étude asymptotique des paramètres dans les arbres

Helmut Prodinger

Department of Algebra and Discrete Mathematics
Technical University of Vienna
Wiedner Hauptstraße 8–10
A–1040, Vienna, Austria

Written for the Proceedings of the Summerschool in Ile de Ré 1983
Retyped 1991

1. Introduction

La structure d'arbre est l'une des structures de données les plus répandues en informatique. Les algorithmes de manipulation d'arbres sont ainsi très nombreux. L'analyse de ces algorithmes au sens de Knuth nous conduit souvent à des problèmes combinatoires de dénombrement des configurations; l'étape suivante consiste en l'analyse asymptotique de ces dénombrements. Voir par exemple [4], [15], [16].

L'étude de Flajolet et Odlyzko [7] est paradigme pour un point de vue un peu différent: Au lieu d'éxtraire des coefficients de fonctions génératrices et d'approximer ceux-ci avec des méthodes de l'analyse réelle, on étudie des singularités des series génératrices. Plus précisement on recherche à établir un développement asymptotique de la fonction autour de la singularité la plus proche de l'origine; ensuite des lemmes de traduction permettent de traduire ce développement en un développement asymptotique des coefficients de Taylor de la série génératrice en question. Par exemple, si $f(z)=\sum_{n\ge0} f_n z^n$ a sa singularité la plus proche de l'origine en 1, avec

$$f(z)\sim A-B\sqrt{1-z}+\dots,$$

on déduit que

$$[z^n]\,f(z):=f_n\sim\frac{B}{2\sqrt{\pi}}n^{-3/2}.$$

De même, si l'on a

$$f(z)\sim -A\log(1-z)+\dots,$$

alors

$$[z^n]f(z)\sim\frac{A}{n},\quad \text{etc.}$$

Enfin, plus on connait de termes du développement de la série génératrice, plus on a de termes du développement des coefficients de Taylor.

Cette traduction est un processus assez automatique. Le problème principal est de trouver le développement autour de la (des) singularité(s). Dans de nombreux

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX

cas, on peut se servir de la transformation de Mellin, méthode très populaire dans l'école française d'algorithmes.

Le but de ce travail est d'expliquer cette stratégie générale dans trois cas particuliers: le problème d'allocation de registres et d'évaluation d'expressions arithmétiques; puis la hauteur des feuilles dans les arbres planaires. Enfin, le troisième problème est la recherche de la valeur asymptotique d'une certaine expression (très similaire au cas de registres); cette valeur est intéressante dans l'analyse d'un algorithme de tri (voir [6], [21]).

Avant de s'engager dans des détails techniques, on formule quelques points fondamentaux:

La classe $\mathcal{B}$ des arbres binaires est définissable récursivement par l'équation (symbolique) qui traduit qu'un arbre binaire est soit l'arbre formé d'un seul sommet externe ("feuille"), soit l'arbre formé d'une racine et de deux arbres qui sont les sous-arbres gauche et droite de la racine. Soit $B(z) = \sum_{n\geq 0} b_n z^n$ la série génératrice des arbres binaires où b_n est le nombre d'arbres binaires de taille n (la taille d'un arbre binaire est mesurée par le nombre de ses sommets internes); on a immediatement

$$B(z) = 1 + z\,(B(z))^2\,,$$

d'où

$$B(z) = \frac{1-\sqrt{1-4z}}{2z} \quad \text{et} \quad b_n = \frac{1}{n+1}\binom{2n}{n}.$$

Les nombres b_n sont les nombres de *Catalan*.

De manière analogue, la famille $\mathcal{P}$ des arbres planaires est définie par d'où on tire pour la série génératrice $P(z) = \sum_{n\geq 1} t_n z^n$ des arbres planaires

$$P(z) = \frac{z}{1-P(z)} \;\text{ et donc }\; P(z) = \frac{1-\sqrt{1-4z}}{2} \;\text{ et }\; t_n = \frac{1}{n}\binom{2n-2}{n-1}.$$

Dans les deux cas, la singularité mentionnée est $z = \frac{1}{4}$, et

$$B(z) \sim 2 - 2\sqrt{1-4z}, \;\text{ donc }\; b_n \sim \frac{1}{\sqrt{\pi}} 4^n n^{-3/2}$$

respectivement

$$P(z) \sim \frac{1}{2} - \frac{1}{2}\sqrt{1-4z}, \;\text{ donc }\; t_n \sim \frac{1}{\sqrt{\pi}} 4^{n-1} n^{-3/2}.$$

2. Le nombre moyen de registres pour evaluer un arbre binaire de taille n

Avant de commencer notre traitement, citons les références sur ce sujet: [4], [5], [8], [11], [17], [19]. On s'abstiendra de présenter ici une description de la problématique informatique. A cet effet on renvoie à la thèse de P.Flajolet [4].

La fonction “nombre de registres” *reg* est donnée recursivement par

$$\mathrm{reg}(\square) = 0$$

$$\mathrm{reg}\left(\quad\right) = \begin{cases} \max\{\mathrm{reg}(t_1), \mathrm{reg}(t_2)\} & \text{si } \mathrm{reg}(t_1) \neq \mathrm{reg}(t_2) \\ 1 + \mathrm{reg}(t_1) & \text{sinon .} \end{cases}$$

Le problème est l'étude de R_n, la valeur moyenne de la fonction *reg*, relativement à l'ensemble des arbres binaires de taille n. On a

$$R_n = \frac{1}{b_n} \cdot [z^n] E(z),$$

avec

$$E(z) = \sum_{n \geq 0} z^n \sum_{t \text{ de taille } n} \mathrm{reg}(t).$$

Soit $\mathcal{R}_p$ la famille des arbres t avec $\mathrm{reg}(t) = p$. De la définition de *reg* on déduit facilement :

Soit $\mathcal{S}_p$ la famille des arbres t avec $\mathrm{reg}(t) \geq p$, c'est à dire

$$\mathcal{S}_p = \mathcal{R}_p + \mathcal{R}_{p+1} + \dots \; .$$

On peut écrire

$$\sum_{j < p-1} \mathcal{R}_j = \mathcal{B} \setminus \mathcal{S}_{p-1}.$$

On obtient une récurrence pour les familles $\mathcal{S}_p$ d'une manière tout à fait semblable:

Les fonctions $S_p(z)$ sont définies par

$$S_p(z) := \sum_{n \geq 0} z^n \sum_{t \text{ de taille } n,\ \mathrm{reg}(t) \geq p} 1.$$

Il est clair que

$$E(z) = \sum_{p \geq 1} S_p(z).$$

Les fonctions $S_p(z)$ sont décrites par les récurrences immédiates

$$S_0 = B; \quad S_p = z S_{p-1}^2 + 2z S_p \left(B - S_{p-1}\right).$$

Il est commode de définir $\varepsilon := \sqrt{1-4z}$ et $z = u/(1+u)^2$. D'où

$$S_p = \frac{zS_{p-1}^2}{1-2zB+2zS_{p-1}},$$

c'est à dire

$$\frac{1}{S_p} = \frac{\varepsilon}{zS_{p-1}^2} + \frac{2}{S_{p-1}}.$$

Ce qui conne finalement

$$\begin{aligned} 1+\frac{\varepsilon}{z}\frac{1}{S_p} &= 1+\frac{\varepsilon^2}{z^2S_{p-1}^2}+\frac{2\varepsilon}{z}\frac{1}{S_{p-1}} = \left(1+\frac{\varepsilon}{z}\cdot\frac{1}{S_{p-1}}\right)^2 \\ &= \cdots = \left(1+\frac{\varepsilon}{z}\cdot\frac{1}{S_0}\right)^{2^p}. \end{aligned}$$

On en déduit l'équation

$$S_p = \frac{\varepsilon}{z}\left(\left(1+\frac{\varepsilon}{zB}\right)^{2^p}-1\right)^{-1}.$$

On voit tout de suite que

$$\frac{\varepsilon}{z} = \frac{1-u^2}{u}, \quad \text{et} \quad 1+\frac{\varepsilon}{zB} = \frac{1}{u},$$

donc

$$S_p = \frac{1-u^2}{u}\frac{1}{u^{-2^p}-1} = \frac{1-u^2}{u}\frac{u^{2^p}}{1-u^{2^p}}.$$

La fonction $E(z)$ a donc pour expression

$$\begin{aligned} E(z) &= \frac{1-u^2}{u}\sum_{p\geq 1}\frac{u^{2^p}}{1-u^{2^p}} \\ &= \frac{1-u^2}{u}\sum_{p\geq 1}\sum_{\lambda\geq 1}u^{\lambda 2^p} \\ &= \frac{1-u^2}{u}\sum_{n\geq 2}v_2(n)u^n, \end{aligned}$$

où $v_2(n)$ est l'exposant de 2 dans la décomposition de n en facteurs premiers:

$$v_2(n) = \sum_{n=\lambda 2^p,\ \lambda,p\geq 1} 1.$$

Si on écrit, d'une manière unique, $n = 2^m(1+2i)$, $(m\geq 1,\ i\geq 0)$, on a

$$v_2(n) = \sum_{p=1}^{m} 1 = m.$$

On appelle souvent la quantité $v_2(n)$ "valuation dyadique", pour des raisons évidentes. Comme il était annoncé dans l'introduction, on va démontrer ultérieurement que l'on a dans un voisinage de $z = \frac{1}{4}$ (c'est à dire $u = 1$):

$$E(z) \sim 2\varepsilon \log_2 \varepsilon + A\varepsilon + 4 \sum_{k \neq 0} c_k \varepsilon^{1-\chi_k} + \dots,$$

avec

$$A = 3 - 2\log_2 2\pi + \frac{2\gamma}{\log 2}, \quad (\gamma \text{ est la constante d'Euler})$$

et

$$\chi_k = \frac{2k\pi i}{\log 2} \quad \text{et} \quad c_k = \frac{1}{\log 2}\Gamma(\chi_k)\zeta(\chi_k)$$

(Γ est la fonction gamma d'Euler, ζ est la fonction zeta de Riemann).

Mentionnons le développement suivant ([10])

$$[z^n]\log(1-z)(1-z)^\alpha \sim \frac{-n^{-\alpha-1}\log n}{\Gamma(-\alpha)} + n^{-\alpha-1}\frac{\Gamma(-\alpha)}{(\Gamma(-\alpha))^2},$$

ce qui nous donne en particulier

$$[z^n]\,\varepsilon\log\varepsilon \sim \frac{1}{2}4^n n^{-3/2}\left(\frac{\log n}{2\sqrt{\pi}} + \frac{\gamma + 2\log 2 - 2}{2\sqrt{\pi}}\right).$$

On utilise également le résultat

$$[z^n]\,\varepsilon^{1-\chi_k} \sim 4^n n^{(\chi_k-3)/2}\Big/\Gamma\left(\frac{\chi_k - 1}{2}\right).$$

Observons alors qu'on peut simplifier l'expression

$$c_k\Big/\Gamma\left(\frac{\chi_k-1}{2}\right) = \frac{\zeta(\chi_k)\Gamma\left(\frac{\chi_k}{2}\right)(\chi_k - 1)}{4\sqrt{\pi}\log 2};$$

pour cela on utilise la formule de duplication de Legendre-Gauss (voir p.e. [2])

$$\Gamma\left(\frac{z}{2}\right)\Gamma\left(\frac{z+1}{2}\right) = \sqrt{\pi}2^{1-z}\Gamma(z).$$

Enfin, si l'on collecte tous les ingrédients, on trouve

$$\begin{aligned}[z^n]\,E(z) \sim{}& \frac{4^n n^{-3/2}}{2\sqrt{\pi}\log 2}(\log n + \gamma + 2\log 2 - 2) \\ &- A\frac{4^n}{2\sqrt{\pi}}n^{-3/2} + \sum_{k\neq 0}\frac{4^n n^{(\chi_k-3)/2}\zeta(\chi_k)\Gamma(\chi_k/2)(\chi_k - 1)}{\sqrt{\pi}\log 2}.\end{aligned}$$

Si on prend en considération que

$$b_n \sim \frac{4^n}{\sqrt{\pi}}n^{-3/2} \qquad \text{et} \qquad n^{\chi_k/2} = \varepsilon^{\log_4 n \cdot 2k\pi i},$$

on obtient

Theoreme 1. *Le nombre moyen R_n de registres nécessaires à l'évaluation d'un arbre de taille n verifie*

$$R_n = \log_4 n + D(\log_4 n) + o(1),$$

où $D(x)$ est une fonction continue périodique de période 1; cette fonction peut être développée en série de Fourier

$$D(x) = \sum_{k\in\mathbb{Z}} d_k e^{2k\pi i x},$$

avec

$$d_0 = \frac{1}{2} - \frac{\gamma}{2\log 2} - \frac{1}{\log 2} + \log_2 \pi$$

et

$$d_k = \frac{1}{\log 2}\zeta(\chi_k)\Gamma(\frac{\chi_k}{2})(\chi_k - 1), \quad k \neq 0, \quad \chi_k = \frac{2k\pi i}{\log 2}.$$

Le reste de ce chapitre est consacré à la découverte du développement de la fonction $E(z)$.

Pour étudier $E(z)$ pour $z \to \frac{1}{4}$, i.e. pour $u \to 1$, on pose $u = e^{-t}$ et on considère $E(z)$ pour $t \to 0$. On obtient facilement que

$$\frac{1-u^2}{u} \sim 2t + \dots .$$

Passons au deuxième facteur de $E(z)$, dénoté par $V(t)$; la transformation de Mellin de $V(t)$ est notée $V^*(s)$:

$$V(t) = \sum_{k\geq 1} v_2(k)e^{-kt}; \qquad V^*(s) = \int_0^\infty x^{s-1}V(x)dx$$

$$\begin{aligned} V^*(s) &= \int_0^\infty x^{s-1} \sum_{k\geq 1} v_2(k)e^{-kx}dx \\ &= \sum_{k\geq 1} v_2(k)k^{-s} \int_0^\infty x^{s-1}e^{-x}dx \end{aligned}$$

On a

$$\sum_{k\geq 1} v_2(k)k^{-s} = 2^{-s}\sum_{k\geq 1} v_2(2k)k^{-s} = 2^{-s}\sum_{k\geq 1}\big[1 + v_2(k)\big]k^{-s},$$

donc

$$V^*(s) = \frac{\Gamma(s)\zeta(s)}{2^s - 1}.$$

La formule d'inversion de Mellin nous donne

$$V(t) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} V^*(s)t^{-s}ds.$$

On peut déplacer le chemin d'intégration vers la gauche, en prenant en compte les résidues de l'intégrande $V^*(s)t^{-s}$. Il y a des pôles en $s = 1$ (provenant de la fonction ζ de Riemann) , en $s = 0, -1, -2, \dots$ (provenant de la fonction Γ) et en $s = \chi_k$, $k \in \mathbb{Z}$ (de $(2^s - 1)^{-1}$). On utilise des développements classiques et on trouve

$$\begin{aligned} V(t) \sim \frac{1}{t} + \frac{1}{2}\log_2 t - \frac{1}{2}\log_2 2\pi + \frac{1}{4} + \frac{\gamma}{2\log 2} \\ + \frac{1}{\log 2}\sum_{k\neq 0}\Gamma(\chi_k)\zeta(\chi_k)t^{-\chi_k} + \dots \end{aligned}$$

ce qui donne la formule annoncée pour $E(z)$, après avoir multiplié les deux derniers développements et en notant que

$$t \sim 2\varepsilon + \dots \ .$$

3. Un probleme similaire: La fusion paire-impaire

Dans l'analyse d'un algorithme de tri (fusion paire-impaire) on doit considérer la quantité

$$B_n = \sum_{k\geq 1}\frac{\binom{2n}{n-k}}{\binom{2n}{n}}\Big(2F(k) + k\Big),$$

représentant le nombre d'échanges dans la fusion paire-impaire de $2n$ éléments. Dans cette égalité,

$$F(k) = \sum_{0\leq j<k} f(j)$$

et $f(j)$ est le nombre des chiffres "1" dans la représentation de j en code de Gray (voir [6], [21]). On trouve ci-dessous les premiers nombres dans cette représentation:

0	0	1100	8
1	1	1101	9
11	2	1111	10
10	3	1110	11
110	4	1010	12
111	5	1011	13
101	6	1001	14
100	7	1000	15
		11000	16

Les nombres

$$\vartheta(j) := f(j) - f(j-1)$$

sont ègaux à ± 1: si k est de la forme $k = 2^m(4i+1)$, $\vartheta(k) = 1$; si k est de la forme $k = 2^m(4i+3)$, $\vartheta(k) = -1$. On peut constater une grande similarité entre $\vartheta(k)$ et $v_2(k)$ du chapitre 2. La série génératrice de Dirichlet de la suite $\vartheta(k)$ est

$$\begin{aligned}
\sum_{k\geq 1} \vartheta(k) k^{-s} &= \sum_{m\geq 0}\sum_{i\geq 0} 2^{-ms}(4i+1)^{-s} - \sum_{m\geq 0}\sum_{i\geq 0} 2^{-ms}(4i+3)^{-s} \\
&= \frac{1}{1-2^{-s}} \cdot \frac{1}{4^s}\left(\zeta(s,\frac{1}{4}) - \zeta(s,\frac{3}{4})\right),
\end{aligned}$$

où la fonction $\zeta(s,a)$ de Hurwitz est définie par

$$\zeta(s,a) = \sum_{n\geq 0}(n+a)^{-s} \qquad \text{pour } \Re(s) > 1.$$

On peut écrire

$$B_n = \frac{n}{2} + \frac{2}{\binom{2n}{n}} C_n,$$

avec

$$C_n = \sum_{k\geq 1} F(k)\binom{2n}{n-k}.$$

On veut trouver une fonction analogue à $E(z)$ (chapitre 2). Donc on écrit

$$\begin{aligned}
C_n &= \sum_{k\geq 1} F(k)\left[u^{n-k}\right](1+u)^{2n} \\
&= [u^n]\sum_{k\geq 1} F(k) u^k (1+u)^{2n} \\
&= [u^n]\sum_{k\geq 1}\sum_{0\leq j<k} f(j)\ u^k(1+u)^{2n} \\
&= [u^n]\sum_{j\geq 0} f(j)u^j\frac{u}{1-u}(1+u)^{2n} \\
&= [u^n]\sum_{j\geq 0}\sum_{0\leq i\leq j}\vartheta(i)u^j\frac{u}{1-u}(1+u)^{2n} \\
&= [u^n]\sum_{i\geq 0}\vartheta(i)u^i\frac{u}{(1-u)^2}(1+u)^{2n}.
\end{aligned}$$

En général, si on a une série $\Psi(z)$, et si on pose $z = u/(1+u)^2$, on a d'après le theorème de Cauchy

$$\begin{aligned}
[z^n]\,\Psi(z) &= \frac{1}{2\pi i}\oint \frac{dz}{z^{n+1}}\Psi(z) \\
&= \frac{1}{2\pi i}\oint \frac{du(1-u)}{(1+u)^3}\frac{(1+u)^{2n+2}}{u^{n+1}}\Psi\Big(\frac{u}{(1+u)^2}\Big) \\
&= [u^n]\,(1-u)(1+u)^{2n-1}\Psi\Big(\frac{u}{(1+u)^2}\Big).
\end{aligned}$$

C'est pourquoi on peut écrire

$$C_n = [u^n]\, \frac{u(1+u)}{(1-u)^3} \sum_{i\geq 1} \vartheta(i) u^i.$$

La fonction dans le membre droit de cette égalité est l'analogue de $E(z)$ du chapitre 2. On trouve facilement, que

$$\frac{u(1+u)}{(1-u)^3} \sim \frac{2}{t^3} + \dots\ .$$

Soit

$$V(t) = \sum_{k\geq 1} \vartheta(k) e^{-kt},$$

d'où

$$V^*(s) = \Gamma(s) \cdot \frac{1}{2^s} \cdot \frac{1}{2^s - 1} \left(\zeta(s, \frac{1}{4}) - \zeta(s, \frac{3}{4}) \right).$$

Par conséquant

$$V(t) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s) \cdot \frac{1}{2^s} \cdot \frac{1}{2^s - 1} \left(\zeta(s, \frac{1}{4}) - \zeta(s, \frac{3}{4}) \right) t^{-s} dt.$$

La différence des fonctions de Hurwitz est une fonction entière; l'intégrande a des pôles en $s = \chi_k$ et $s = 0, -1, -2, \dots$. On a (voir [22])

$$\zeta(s, \frac{1}{4}) - \zeta(s, \frac{3}{4}) \sim \frac{1}{2} + s \left(2 \log \Gamma(\frac{1}{4}) - \frac{1}{2} \log 2 - \log \pi \right) + \dots$$

pour $s \to 0$. De même,

$$\begin{aligned} \Gamma(s) &\sim \frac{1}{s} - \gamma + \dots \\ \frac{1}{2^s - 1} &\sim \frac{1}{\log 2} \cdot \frac{1}{s} - \frac{1}{2} + \dots \\ (2t)^{-s} &\sim 1 - s \cdot \log(2t) + \dots \end{aligned}$$

ce qui nous donne le résidu en $s = 0$

$$-\frac{5}{4} - \frac{\gamma}{2 \log 2} + 2 \log_2 \Gamma(\frac{1}{4}) - \log_2 \pi - \frac{1}{2} \log_2 t.$$

Le résidu en $s = \chi_k$ pour $k \neq 0$ est simplement

$$\frac{1}{\log 2} \Gamma(\chi_k) \left(\zeta(\chi_k, \frac{1}{4}) - \zeta(\chi_k, \frac{3}{4}) \right) t^{-\chi_k}.$$

On peut simplifier, puisque

$$\zeta(\chi_k, \frac{3}{4}) = -\zeta(\chi_k, \frac{1}{4}).$$

Donc, on obtient

$$C_n \sim [z^n]\, \frac{2}{t^3}\Big(-\frac{1}{2}\log_2 t - \frac{5}{4} - \frac{\gamma}{2\log 2} + 2\log_2 \Gamma(\frac{1}{4}) - \log_2 \pi \\ + \frac{2}{\log 2}\sum_{k\neq 0} \Gamma(\chi_k)\zeta(\chi_k, \frac{1}{4}) t^{-\chi_k}\Big),$$

où, avec $t \sim 2\varepsilon$

$$C_n \sim [z^n]\Big(-\frac{1}{8}\varepsilon^{-3}\log_2 \varepsilon - \varepsilon^3\Big(\frac{7}{16} + \frac{\gamma}{8\log 2} - \frac{1}{2}\log_2 \Gamma(\frac{1}{4}) + \frac{1}{4}\log_2 \pi\Big) \\ + \frac{1}{2\log 2}\sum_{k\neq 0} \Gamma(\chi_k)\zeta(\chi_k, \frac{1}{4})\varepsilon^{-3-\chi_k}\Big).$$

On trouve, comme au chapitre 2,

$$[z^n]\, \varepsilon^3 \log \varepsilon \sim \frac{4^n\sqrt{n}}{\sqrt{\pi}}\Big(-\log n + (2 - \gamma - 2\log 2)\Big),$$

$$[z^n]\, \varepsilon^{-3} \sim \frac{2\cdot 4^n \sqrt{n}}{\sqrt{\pi}},$$

$$[z^n]\, \varepsilon^{-3-\chi_k} \sim 4^n \frac{n^{(1+\chi_k)/2}}{\Gamma\left(\frac{3+\chi_k}{2}\right)}.$$

En regroupant les résultats précédents, on obtient

$$C_n \sim \frac{4^n\sqrt{n}}{\sqrt{\pi}}\log n \cdot \frac{1}{8\log 2} \\ + \frac{4^n\sqrt{n}}{\sqrt{\pi}}\left(-\frac{1}{4\log 2} - \frac{\gamma}{8\log 2} - \frac{5}{8} + \log_2 \Gamma(\frac{1}{4}) - \frac{1}{2}\log_2 \pi\right) \\ + \frac{1}{2}4^n\sqrt{n}\frac{1}{\log 2}\sum_{k\neq 0} \Gamma(\chi_k)\zeta(\chi_k, \frac{1}{4})\frac{n^{\chi_k/2}}{\Gamma\left(\frac{3+\chi_k}{2}\right)}.$$

On a par la formule de duplication

$$\Gamma(\chi_k)\Big/\Gamma\left(\frac{3+\chi_k}{2}\right) = \frac{1}{1+\chi_k}\frac{1}{\sqrt{\pi}}\Gamma\left(\frac{\chi_k}{2}\right).$$

Avec

$$\binom{2n}{n} \sim \frac{4^n n^{-1/2}}{\sqrt{\pi}},$$

on obtient

$$
\begin{aligned}
\frac{D_n}{\binom{2n}{n}} \cdot \frac{1}{n} \sim \frac{1}{8} \log_2 n \\
& - \frac{1}{4 \log 2} - \frac{\gamma}{8 \log 2} - \frac{5}{8} + \log_2 \Gamma\big(\frac{1}{4}\big) - \frac{1}{2} \log_2 \pi \\
& + \frac{1}{2 \log 2} \sum_{k \neq 0} \zeta\big(\chi_k, \frac{1}{4}\big) \frac{\Gamma(\chi_k/2)}{1+\chi_k} n^{\chi_k/2}.
\end{aligned}
$$

Finalement on a trouvé le

Theorème 2. *Le nombre moyen d'échanges dans la fusion paire-impaire de $2n$ éléments vérifie*

$$B_n \sim \frac{1}{4} n \log_2 n + nB\big(\log_4 n\big),$$

ou $B(x)$ est une fonction continue périodique de période 1; cette fonction peut être développée en série de Fourier

$$B(x) = \sum_{k \in \mathbb{Z}} b_k e^{2k\pi i x},$$

avec

$$b_0 = -\frac{1}{2 \log 2} - \frac{\gamma}{4 \log 2} - \frac{3}{4} + 2 \log_2 \Gamma\big(\frac{1}{4}\big) - \log_2 \pi$$

et pour $k \neq 0$

$$b_k = \frac{1}{\log 2} \zeta\big(\chi_k, \frac{1}{4}\big) \frac{\Gamma(\chi_k/2)}{1+\chi_k}.$$

4. La hauteur des feuilles dans les arbres planaires

On peut comprendre la problématique de ce chapitre très simplement en considérant l'example d'un arbre planaire:

Cet arbre a 4 feuilles de hauteur (de gauche à droite) $3,3,4,2$. Nous remarquons que la hauteur de l'arbre est donnée par la hauteur de la feuille la plus haute. Nous imaginons maintenant que les feuilles sont ordonnées par ordre décroissant de leurs hauteurs. Dans notre exemple, la suite des hauteurs de feuilles est ainsi $4,3,3,2$. Remarquons qu'il peut y avoir plusieurs feuilles de même hauteur. Pour faciliter les calculs, nous complétons cette suite par une infinité de zéros. Ainsi, nous pouvons parler de la d-ième feuille pour n'importe quel d, même si l'arbre a moins de d feuilles.

On considère maintenant $h_n^{(d)}$, la hauteur de la d-ième feuille en remarquant encore une fois que $h_n^{(1)}$ est simplement la hauteur moyenne d'un arbre de taille n.

Quelques lemmes techniques apparaissent déjà dans divers articles, voir [1], [12], [13], [14], [20].

Lemme 1. *Soit $F(z,y)$ la série génératrice telle que $[z^n y^m]F$ soit le nombre d'arbres de taille n avec m feuilles (l'arbre $\circ$ a zéro feuilles); on a alors pour tout $p \geq 0$*

$$F^p(z,y) = z^p + p \sum_{0 \leq \lambda < n} z^n y^{n-\lambda} \frac{1}{n} \binom{n}{\lambda} \binom{n-p-1}{\lambda-p}.$$

Idée de la demonstration. (Voir [12]) On introduit $G(z,y)$, la quantité analogue, où on considère que $\circ$ a une feuille; en traduisant l'équation symbolique pour les arbres planaires, on obtient

$$G(z,y) = z\Bigl(y + \sum_{r \geq 1} G^r(z,y)\Bigr) = \bigl(1 - z(1-y)\bigr) P\left(\frac{zy}{\bigl(1-z(1-y)\bigr)^2}\right).$$

On utilise le développement classique

$$P^m(z) = \sum_{k \geq 0} \left[\binom{2k+m-1}{k} - \binom{2k+m-1}{k-1}\right] z^{k+m}.$$

Après avoir utilisé deux fois le théorème binomial, on se sert de l' identité

$$\sum_{k \geq 0} \binom{r}{k} \binom{s+k}{p} (-1)^k = (-1)^r \binom{s}{p-r}$$

pour obtenir

$$G^p(z,y) = z^p y^p + p \sum_{p \leq \lambda \leq n} z^n y^\lambda \frac{1}{n} \binom{n}{\lambda} \binom{n-p-1}{\lambda-p}.$$

Enfin, on peut constater que

$$F(z,y) = G(z,y) - zy + z = G(zy, y^{-1}),$$

d'où la formule annoncée.

Lemme 2. *Soit $A_k(z)$ la série génératrice des arbres de hauteur $\leq k$ et, comme précédemment, $z = u/(1+u)^2$. On a*

$$A_k(z) = \frac{u}{1+u} \cdot \frac{1-u^k}{1-u^{k+1}}.$$

Démonstration. En traduisant l'équation symbolique pour la famille $\mathcal{P}$, on obtient

$$A_k = \frac{z}{1-A_{k-1}}, \qquad k \geq 1, \quad A_0 = 0;$$

on peut donc démontrer la formule très facilement par récurrence. $\square$

Lemme 3. *Soit $Q_k(z,y)$ la série génératrice, où $[z^n y^m]Q_k$ est le nombre d'arbres de taille n avec m nœuds au niveau k, qui est le niveau maximal. (C'est à dire, si $m \geq 1$, la hauteur d'arbre est exactement k.) On a pour $k \geq 1$ et $d \geq 1$*

$$[y^d]Q_k(z,y) = A_{k-1}^{d+1}(z) - A_{k-1}^{d}(z) + zA_{k-1}^{d-1}(z).$$

Démonstration. En traduisant comme d'habitude, on obtient

$$Q_1 = zy \qquad \text{et} \qquad Q_k = \frac{z}{1-Q_{k-1}} \quad \text{pour } k \geq 2.$$

Comme on a

$$[y^0]Q_k(z,y) = A_{k-1}(z),$$

il suffit de démontrer que

$$\begin{aligned} Q_k(z,y) &= A_{k-1}(z) + \sum_{d\geq 1} y^d \Big(A_{k-1}^{d+1}(z) - A_{k-1}^{d}(z) + zA_{k-1}^{d-1}(z)\Big) \\ &= \frac{A_{k-1}}{1-yA_{k-1}} - \frac{yA_{k-1}}{1-yA_{k-1}} + \frac{zy}{1-yA_{k-1}} \\ &= \frac{(1-y)A_{k-1} + zy}{1-yA_{k-1}}; \end{aligned}$$

ce qui est un simple exercice, en utilisant la récurrence des A_k.

Soit $B_{d,k}(z)$ la série génératrice, où $[z^n]B_{d,k}$ est le nombre d'arbres de taille n où d feuilles sont à hauteur $> k$ et toutes les autres sont à hauteur $\leq k$. On peut constater que $B_{0,k} + B_{1,k} + \cdots + B_{d-1,k}$ est la série génératrice d'arbres où la d-ième feuille (au sens défini dessus) est à hauteur $\leq k$. Donc la quantité $h_n(d)$ est donnée par

$$h_n^{(d)} = \frac{[z^n]\sum_{k\geq 0}\Big(P(z) - B_{0,k}(z) - B_{1,k}(z) - \cdots - B_{d-1,k}(z)\Big)}{[z^n]\,P(z)}.$$

Lemme 4. *La hauteur moyenne d'un arbre planaire de taille n vérifie*

$$\frac{[z^n]\sum_{k\geq 0}\Big(P(z)-B_{0,k}(z)\Big)}{[z^n]\,P(z)}\sim\sqrt{\pi n}-\frac{1}{2}+\dots\ .$$

Idée de la demonstration. On trouve que

$$\begin{aligned}\sum_{k\geq 0}\Big(P(z)-B_{0,k}(z)\Big)&=\sum_{k\geq 0}\Big(P(z)-A_k(z)\Big)\\&=\frac{1-u}{1+u}\sum_{k\geq 1}\frac{u^k}{1-u^k}=\frac{1-u}{1+u}\sum_{k\geq 1}d(k)u^k\end{aligned}$$

où $d(k)$ est le nombre de diviseurs de k. On constate que

$$\sum_{k\geq 1}d(k)k^{-s}=\zeta^2(s),$$

et on peut utiliser la transformation de Mellin comme au chapitre 2.

Comme dans les chapitres 2 et 3 on va démontrer aue

$$\sum_{k\geq 0}B_{s+1,k}(z)\sim A_s-B_s t+\dots,$$

où A_s et B_s sont des constantes et $u=e^{-t}$. Comme

$$P(z)=\frac{u}{1+u}\sim\frac{1}{2}-\frac{1}{4}t+\dots,$$

on peut conclure que

$$[z^n]\sum_{k\geq 0}B_{s+1,k}(z)\Big/t_n\sim 4B_s.$$

Lemme 5. *On a pour* $s\geq 0$

$$\sum_{n\geq p}z^n\binom{n}{s}\binom{n-p}{s}=\sum_{j=0}^{s}\frac{z^{s+p+j}}{(1-z)^{s+1+j}}\frac{(p+s)!(s+j)!}{j!(s-j)!(p+j)!s!}.$$

Démonstration. Avec $D=d/dz$ et $(n)_d=n(n-1)\dots(n-d+1)$ on a

$$\begin{aligned}\sum_{n\geq p}z^n\binom{n}{s}\binom{n-p}{s}&=\frac{1}{(s!)^2}\sum_{n\geq 0}(n)_s z^s D^s z^{n+p}\\&=\frac{1}{(s!)^2}z^s D^s z^{p+s}\sum_{n\geq 0}z^n\\&=\frac{1}{s!}z^s D^s\frac{z^{p+s}}{(1-z)^{s+1}}\\&=\frac{1}{s!}z^s\sum_{j=0}^{s}\binom{s}{j}\big[D^{s-j}z^{p+s}\big]\big[D^j(1-z)^{-s-1}\big].\quad\square\end{aligned}$$

Lemme 6. *On a pour $s \geq 0$ et $k \geq 1$*

$$B_{s+1,k}(z) = \sum_{j=0}^{s} \frac{(s+j)!}{j!(s-j)!(s+1)!} \left(\frac{u}{1+u+u^2}\right)^{s+1+j}$$
$$\times \frac{(1-u)^2}{1+u} \frac{u^{k-1}}{(1-u^{k-1})(1-u^k)} \sum_{p\geq 1} \frac{p(p+s)!}{(p+j)!} A_{k-1}^p.$$

Démonstration. On remplace chaque feuille au niveau maximal k (marquée par y en $Q_k(z,y)$) par un arbre arbitraire dont les feuilles sont marquées par y. Alors

$$\begin{aligned} B_{s+1,k} &= \left[y^{s+1}\right] Q_k\Big(z, \frac{1}{z}F(z,y)\Big) \\ &= \left[y^{s+1}\right] \sum_{p\geq 1} \left[\frac{1}{z}F(z,y)\right]^p \left[A_{k-1}^{p+1} - A_{k-1}^p + zA_{k-1}^{p-1}\right] \\ &= \sum_{p\geq 1} z^{-p} p \sum_{n\geq 1} z^n \frac{1}{n} \binom{n}{s+1} \binom{n-p-1}{s} A_{k-1}^p \left[A_{k-1} - 1 + zA_{k-1}^{-1}\right]. \end{aligned}$$

La formule annoncée est obtenue en remarquant que

$$A_{k-1} - 1 + zA_{k-1}^{-1} = \frac{u^{k-1}}{(1-u^{k-1})(1-u^k)} \frac{(1-u)^2}{1+u},$$
$$\frac{z}{1-z} = \frac{u}{1+u+u^2}$$

et en utilisant le Lemme 5.

Lemme 7. *La série génératrice $B_{s+1,k}(z)$ est donnée par*

$$B_{s+1,k}(z) = \sum_{j=0}^{s} \sum_{i=0}^{s-j} \frac{(s+j)!(i+1)}{j!(s-j-i)!(j+i+1)!} \times$$
$$\times \left(\frac{u}{1+u+u^2}\right)^{s+1+j} (1-u)^2 u^{i+1} u^{k-1} \frac{(1-u^{k-1})^i}{(1-u^{k+1})^{i+2}}.$$

Démonstration. On utilise la formule (déduite de la formule de *Vandermonde*)

$$\frac{p(p+s)!}{(s-j)!(s+1)!(p+j)!} = \sum_{i=0}^{s-j} (i+1) \binom{p}{i+1} \frac{1}{(s-j-i)!(j+i+1)!}$$

et aussi les formules

$$\frac{A_k}{z} = \frac{1}{1-A_{k-1}}$$

et

$$\sum_{p\geq 1} \binom{p}{i+1} A_{k-1}^p = \frac{A_{k-1}^{i+1}}{\left(1-A_{k-1}\right)^{i+2}}$$

Nous sommes prêts à attaquer la détermination des développements locaux en $\sum_{k\geq 0} B_{s+1,k}$. On constate facilement que

$$\frac{u}{1+u+u^2} \sim \frac{1}{3}; \qquad (1-u)^2 \sim t^2(1-t); \qquad u^{i+1} \sim 1-(i+1)t.$$

Lemme 8. *On a le développement local*

$$\sum_{k\geq 0} u^{k-1}\frac{(1-u^{k-1})^i}{(1-u^{k+1})^{i+2}}$$
$$\sim t^{-2}(1+2t)\Big(\sum_{l=0}^{i}\binom{i}{l}(-2)^l(1+lt)\Big(\zeta(l+2)-\frac{l}{2}\zeta(l+1)t\Big)-\frac{1}{2}t\Big).$$

Démonstration. On peut écrire

$$\begin{aligned}\sum_{k\geq 0} u^{k-1}\frac{(1-u^{k-1})^i}{(1-u^{k+1})^{i+2}} &= u^{-2}\sum_{k\geq 1}\frac{u^k}{(1-u^k)^2}\left(1+\frac{(1-u^{-2})u^k}{1-u^k}\right)^i\\ &= u^{-2}\sum_{l=0}^{i}\binom{i}{l}\left(1-u^{-2}\right)^l\sum_{k\geq 1}\frac{u^{k(l+1)}}{(1-u^k)^{l+2}}.\end{aligned}$$

On voit immédiatement que

$$u^{-2}\sim 1+2t \qquad \text{et} \qquad (1-u^{-2})^l\sim(-2)^l t^l(1+lt).$$

On applique encore la transformation de Mellin:

$$\begin{aligned}\mathcal{M}\sum_{k\geq 1}\frac{e^{-tk(l+1)}}{(1-e^{tk})^{l+2}} &= \sum_{k\geq 1}k^{-s}\mathcal{M}\frac{e^{-t(l+1)}}{(1-e^{t})^{l+2}}\\ &= \zeta(s)\Gamma(s)\sum_{\lambda\geq 0}\binom{\lambda}{l+1}\lambda^{-s}\\ &= \zeta(s)\Gamma(s)\sum_{\lambda\geq 0}\frac{1}{(l+1)!}\sum_{i=0}^{l+1}s_{l+1,i}\lambda^i\lambda^{-s}\end{aligned}$$

(où les $s_{l+1,j}$ sont les nombres *Stirling* de première espèce)

$$=\frac{1}{(l+1)!}\zeta(s)\Gamma(s)\sum_{i=0}^{l+1}s_{l+1,i}\zeta(s-i).$$

Comme au chapitre 2 on utilise la formule d'inversion de Mellin; on change le chemin d'intégration en prenant en compte les résidues. Dans notre problème, il suffit de considérer les termes où $i=s+1$ et $i=s$. On peut constater que

$$s_{l+1,l+1}=1 \qquad \text{et} \qquad s_{l+1,l}=-\binom{l+1}{2}.$$

Lemme 9. *La somme des $B_{s+1,k}(z)$ vérifie le développement local*

$$\sum_{k\geq 0} B_{s+1,k}(z) \sim const - \frac{1}{2}t\sum_{j=0}^{s}\sum_{i=0}^{s-j}\frac{(s+j)!(i+1)}{j!(s-j-i)!(j+i+1)!}3^{-s-1-j}.$$

Idée de la Démonstration.

$$\sum_{k\geq 0} B_{s+1,k}(z) \sim \sum_{j=0}^{s}\sum_{i=0}^{s-j}\frac{(s+j)!(i+1)}{j!(s-j-i)!(j+i+1)!}3^{-s-1-j}t^2(1-t)\big(1-(i+1)t\big)\times$$
$$\times\, t^{-2}(1+2t)\left(\sum_{l=0}^{i}\binom{i}{l}(-2)^l(1+lt)\Big(\zeta(l+2)-\frac{l}{2}\zeta(l+1)t\Big)-\frac{1}{2}t\right)$$

On veut obtenir le coefficient de t dans cette expression. C'est une chose très simple; après quelques manipulations de nombreuses simplifications apparaissent. $\square$

A partir des égalités précédentes, on tire

$$[z^n]\, B_{s+1,k}(z)\Big/t_n \sim 4\cdot\frac{1}{2}\sum_{j=0}^{s}\sum_{i=0}^{s-j}\frac{(s+j)!(i+1)}{j!(s-j-i)!(j+i+1)!}3^{-s-1-j}.$$

Pour la quantité $h_n^{(d)}$ on doit sommer:

$$\sum_{s=0}^{d-2}2\sum_{j=0}^{s}\sum_{i=0}^{s-j}\frac{(s+j)!(i+1)}{j!(s-j-i)!(j+i+1)!}3^{-s-1-j}.$$

On peut écrire pour cette quantité

$$\sum_{s=0}^{d-2}\frac{2}{(s+1)3^{s+1}}\,[x^s]\,\frac{1}{(1-x)^2}\left(\frac{1+x}{1-\frac{x}{3}}\right)^{s+1}.$$

Donc nous avons le

Theorème 3. *La hauteur moyenne de la d-ième feuille d'un arbre planaire de taille n, a le développement asymptotique (pour d fixe et $n\to\infty$)*

$$h_n^{(d)} = \sqrt{\pi n}-\frac{1}{2}-\sum_{s=0}^{d-2}\frac{2}{(s+1)3^{s+1}}\,[x^s]\,\frac{1}{(1-x)^2}\left(\frac{1+x}{1-\frac{x}{3}}\right)^{s+1} + \ o(1).$$

5. Quelques remarques sur la technique d'extraction de coefficients

Si on suit la stratégie mentionnée dans le titre, les problèmes de dénombrement d'arbres demandent souvent d'évaluer des sommes du type

$$\sum_{k\ge 1}\omega(k)\frac{\binom{2n}{n+k+1}-2\binom{2n}{n+k}+\binom{2n}{n+k-1}}{\binom{2n}{n}}.$$

où $\omega(k)$ peut être, par exemple, $v_2(k)$ ou $\vartheta(k)$ ou $d(k)$ comme dans les chapitres 2, 3 et 4.

On a besoin d'approximations de certains coefficients binomiaux dans une région centrale (p.e. $|k|\le\sqrt{n}\log n$). On obtient, par exemple par la formule de *Stirling*

$$\frac{\binom{2n}{n+k}}{\binom{2n}{n}}\sim e^{-k^2/n}\cdot\left(1+\frac{k^2}{2n^2}-\frac{k^4}{6n^3}+\dots\right)$$

(voir [9] pour une formule générale). On voit souvent apparaitre les différences r-ièmes de coefficients binomiaux

$$\xi=[u^n]\,u^k(1-u)^r(1+u)^{2n}\Big/\binom{2n}{n}=\sum_{\lambda=0}^{r}(-1)^\lambda\binom{r}{\lambda}\binom{2n}{n+k+r}\Big/\binom{2n}{n}.$$

On rappelle l'opérateur des différences Δ:

$$\Delta^r\binom{2n}{n+k}=\sum_{\lambda=0}^{r}\binom{r}{\lambda}(-1)^{r-\lambda}\binom{2n}{n+k+\lambda}$$

Nous considérons $\Delta^r e^{-k^2/n}$. Pour cela, on a besoin des polynomes d'*Hermite*. Voici les propriétées necessaires [3]:

$$e^{2zt-t^2}=\sum_{n\ge 0}H_n(z)\frac{t^n}{n!}$$

$$H_0=1,\quad H_1=2z,\quad H_{n+1}=2zH_n-2nH_{n-1}$$

$$H_n(z)=\sum_{k\ge 0}\frac{n!}{k!(n-2k)!}(-1)^k(2z)^{n-2k}$$

$$H_n(-z)=(-1)^nH_n(z)$$

Donc

$$\begin{aligned}\Delta^r e^{-k^2/n}&=\sum_{\lambda=0}^{r}(-1)^{r-\lambda}\binom{r}{\lambda}e^{-(k+\lambda)^2/n}\\&=e^{-k^2/n}\sum_{\lambda=0}^{r}(-1)^{r-\lambda}\binom{r}{\lambda}e^{-\frac{2k\lambda}{n}-\frac{\lambda^2}{n}}\\&=e^{-k^2/n}\sum_{\lambda=0}^{r}(-1)^{r-\lambda}\binom{r}{\lambda}\sum_{m\ge 0}H_m\left(-\frac{k}{\sqrt{n}}\right)\left(\frac{\lambda}{\sqrt{n}}\right)^m\frac{1}{m!}.\end{aligned}$$

On sait (avec les nombres de *Stirling* de second espèce) que

$$\sum_{\lambda=0}^{r}(-1)^{r-\lambda}\binom{r}{\lambda}\lambda^m = r!S_{m,r}.$$

Donc

$$(-1)^r\Delta^r e^{-k^2/n} \sim e^{-k^2/n}\left[H_r\left(\frac{k}{\sqrt{n}}\right)\frac{1}{n^{r/2}} - H_{r+1}\left(\frac{k}{\sqrt{n}}\right)\frac{1}{n^{(r+1)/2}}\frac{r}{2}+\dots\right].$$

On trouve donc

$$\xi \sim e^{-k^2/n}H_r\left(\frac{k}{\sqrt{n}}\right)\frac{1}{n^{r/2}},$$

résultat annoncée en [5].

Pour obtenir plusieurs termes, on doit considérer $\Delta^r k^t e^{-k^2/n}$.

On peut utiliser la formule générale (de type *Leibniz*):

$$\Delta^r f(k)g(k) = \sum_{\lambda=0}^{r}\binom{r}{\lambda}\left(\Delta^{r-\lambda}f(k+\lambda)\right)\left(\Delta^\lambda g(k)\right)$$

avec $f(k) = k^t$ et $g(k) = e^{-k^2/n}$.

Si on a (pour un certain polynome p)

$$[u^n]\, u^k p(u)(1+u)^{2n},$$

on peut développer $p(u)$ en puissances de $1-u$ et utiliser la machinerie comme nous l'avons indiqué.

Dans de nombreux autres cas, on rencontre les quantités

$$[u^n]\, u^k \frac{1}{1-u}(1+u)^{2n};$$

c'est à dire qu'on doit considérer

$$\sum_{j\ge k} e^{-j^2/n} \sim \sqrt{n}\,\mathrm{erfc}\left(\frac{k}{\sqrt{n}}\right) + \frac{1}{2}e^{-k^2/n}+\dots,$$

avec la formule de sommation d'*Euler* et la fonction d'erreur

$$\mathrm{erfc}(z) := \int_z^\infty e^{-t^2}dt.$$

Pour obtenir plusieurs termes, on doit considérer

$$I_m(z) = \int_z^\infty t^m e^{-t^2}dt;$$

on peut évaluer cette intégrale avec une intégration par parties:

$$I_m(z) = \frac{m-1}{2} I_{m-2}(z) + \frac{1}{2} z^{m-1} e^{-z^2},$$
$$I_1(z) = \frac{1}{2} e^{-z^2}, \qquad I_0(z) = \operatorname{erfc}(z).$$

Enfin un problème plus général de même nature apparait dans l'evaluation des sommes du type

$$[u^n]\, u^k \frac{1}{(1-u)^a} (1+u)^{2n}$$

où a est un nombre entier ≥ 1.

Bibliographie

1. de Bruijn, Knuth, Rice, *The average height of planted plane trees*, Graph Theory and Computing (R.C. Read, eds.), Academic Press, New York, 1972, pp. 15–22.
2. Dieudonné, *Calcul Infinitesimal*, Hermann, Paris, 1968.
3. Erdélyi, Magnus, Oberhettinger, Tricomi, *Higher Transcendental Functions*, Mc-Graw-Hill, 1955.
4. Flajolet, *Analyse d'algorithmes d'arbres et de fichiers*, Cahiers du BURO **34–35** (1981), 1–209.
5. Flajolet, Raoult, Vuillemin, *The number of registers required for evauluating arithmetic expressions*, Theor. Comput. Sci. **9** (1979), 99–125.
6. Flajolet, Ramshaw, *A note on Gray Code and odd-even merge*, SIAM J. on Computing **9** (1980), 142–158.
7. Flajolet, Odlyzko, *The average height of binary trees and other simple trees*, J. Comput. Syst. Sci. **25** (1982), 171–213.
8. Flajolet, Prodinger, *Register Allocation for unary-binary Trees*, SIAM J. on Computing **15** (1986), 629–640.
9. Greene, Knuth, *Mathematics for the Analysis of Algorithms*, Birkhäuser, 1981.
10. Jungen, *Sur les séries de Taylor n'ayant que des singularités algébro-logarithmiques sur leur cercle de convergence*, Commentarii Math. Helvetici **3** (1931), 266–306.
11. Kemp, *The average number of registers to evaluate a binary tree optimally*, Acta. Inf. **11** (1979), 363–372.
12. Kemp, *On the average oscillation of a stack*, Combinatorica **2** (1982), 157–176.
13. Kemp, *The average height of planted plane trees with m leaves*, Journal of Combinatorial Theory Ser. B **B 34** (1983), 191–208.
14. Kemp, *On the number of deepest nodes in ordered trees*, Discrete Mathematics **81** (1990), 247–258.
15. Knuth, *The Art of Computer Programming, Vol.1*, Addison-Wesley, 1968.
16. Knuth, *The Art of Computer Programming, Vol.3*, Addison-Wesley, 1968.
17. Meir, Moon and Pounder, *On the order of random channel networks*, SIAM J. on Alg. Disc. Meth. **1** (1980), 25–33.
18. Prodinger, *The average height of the second highest leaf of a planted plane trees*, Europ. J. on Combinatorics **5** (1984), 351–357.
19. Prodinger, *Die Bestimmung gewisser Parameter von binären Bäumen mit Hilfe analytischer Methoden*, Lecture Notes in Mathematics (E.Hlawka, Ed.) **1114** (1985), 118–133.
20. Prodinger, *The Average Height of the d-th Highest Leaf of a Planted Plane Tree*, Networks **16** (1985), 67–75.
21. Sedgewick, *Data movement in odd-even merge*, SIAM J. on Computing **7** (1978), 239–272.
22. Whittaker, Watson, *A course in modern analyis*, Cambridge University Press, 1930.