L'analyse de la hauteur des arbres

Philippe Flajolet
Résumé par *Dominique Gouyou-Beauchamps*

Séminaire de combinatoire énumérative et analytique de l'IHP - 2010-11

Résumé

Quantifier la distribution de la hauteur d'un arbre est un problème de base partagé par les probabilistes et les combinatoriciens. D'un point de vue probabiliste, d'intéressantes connexions ont été établies avec le mouvement Brownien, les processus de branchement, et le modèle dit "de l'arbre continu" (CRT). L'objectif de cet exposé est une présentation synthétique de méthodes asymptotiques fondées sur l'analyse complexe et la combinatoire analytique : ces méthodes analytiques ont en effet permis d'obtenir une caractérisation très complète de la hauteur dans les principales familles d'arbres combinatoires (arbres de Catalan, binaires ou généraux ; arbres d'Otter). Lois limites locales ou centrales, convergence de moments et bornes de grandes déviations en résultent. On croisera au passage l'ensemble de Mandelbrot, les transformations de fonctions elliptiques théta, ainsi que la théorie élémentaire de l'itération analytique. [Présentation fondée notamment sur des travaux commun avec Broutin, Gao, Odlyzko, et Richmond.]

1 Introduction

La question qu'on se pose est : pour différents modèles d'arbres combinatoires comme les arbres généraux ou arbres de Catalan, les arbres binaires, les variétés simples d'arbres et les arbres non planaires, quelle est la hauteur d'un arbre aléatoire ? On peut aussi se poser cette question pour le diamètre ou la largeur. Parmi les résultats présentés ici, beaucoup se trouvent dans [8].

2 Les arbres "généraux" de Catalan

2.1 Arbres généraux et nombres de Catalan

Les arbres généraux sont planaires et tous les degrés sont permis. Leur famille $\mathcal{G}$ est décrite par l'équation symbolique : $\mathcal{G} = \mathcal{Z} \times \mathrm{SEQ}(\mathcal{G})$, c'est à dire qu'un arbre général est formé par une racine $\mathcal{Z}$ à laquelle est attachée une suite d'arbre généraux $\mathrm{SEQ}(\mathcal{G})$. Sa fonction génératrice $G(z) := \sum G_n z^n$, où la variable z compte les sommets, vérifie donc l'équation $G(z) = \frac{z}{1-G(z)}$, ce qui donne la formule $G(z) = \frac{1}{2}(1 - \sqrt{1-4z})$. Les coefficients de $G(z)$ sont les nombres de Catalan $G_{n+1} = \frac{1}{n+1}\binom{2n}{n}$ qui, asymptotiquement, sont égaux à $G_{n+1} \sim \frac{4^n}{\sqrt{\pi\, n}}$

2.2 De Bruijn, Knuth et Rice 1972

Soit $\mathcal{G}^{[h]}$ la famille des arbres généraux de hauteur au plus h. Elle est décrite par les équations symboliques : $\mathcal{G}^{[h+1]} = \mathcal{Z} \times \mathrm{SEQ}(\mathcal{G}^{[h]})$ et $\mathcal{G}^{[0]} = \mathcal{Z}$. Sa fonction génératrice $G^{[h]}(z)$, vérifie donc les équations $G^{[h+1]}(z) = \frac{z}{1-G^{[h]}(z)}$ et $G^{[0]}(z) = z$, ce qui donne la fraction continue tronquée à la profondeur h :

$$G^{[h]}(z) = \left.\cfrac{z}{1 - \cfrac{z}{1 - \cfrac{z}{\ddots\, \overline{1-z}}}}\right\} h \text{ étages} \tag{1}$$

La série $G^{[h]}(z)$ peut s'exprimer à l'aide des polynômes de Fibonacci $F_h(z)$, qui vérifient la récurrence linéaire $F_{h+1} = F_h - z\, F_{h-1}$, sous la forme $G^{[h]}(z) = z\frac{F_{h+1}}{F_{h+2}}$. L'équation caractéristique associée à la récurrence est $\rho^2 = \rho - z$, équation qui a deux solutions $\rho, \bar{\rho} = \frac{1}{2}(1 \mp \sqrt{1-4z})$. Ainsi $F_h = \frac{\rho^h - \bar{\rho}^h}{\rho - \bar{\rho}}$. On peut tout exprimer en fonction de la seule racine $\rho \equiv G(z)$, ce qui permet d'appliquer la formule d'inversion de Lagrance :

Théorème 1 ([6])

$$G^{[h-1]} = z\frac{\rho^h - \bar{\rho}^h}{\rho^{h+1} - \bar{\rho}^{h+1}}\ ; \quad G_{n+1} - G_{n+1}^{[h-1]} = \sum_{j\geq 1} \Delta^2 \binom{2n}{n-jh}$$

où l'opérateur Δ *est défini par* $\Delta^2 f(x) = f(x+1) - 2\,f(x) + f(x-1)$.

Ainsi le nombre d'arbres généraux de taille $n+1$ et de hauteur bornée peut s'exprimer comme la somme d'un échantillonage de la ligne $2n$ du triangle de Pascal.

Si on connaît l'adage disant que "toutes les récurrences linéaires du second ordre sont de même nature et sont équivalentes à un produit de sinus et de cosinus", on peut établir que :

$$F_h\left(\frac{1}{4\cos^2\theta}\right) = \frac{1}{(2\cos\theta)^{h-1}}\frac{\sin h\theta}{\sin\theta}.$$

Et nous pouvons alors constater que les polynômes de Fibonacci sont fortement liés à ceux de Chebyshev, que les racines de $F_h(z) = 0$ sont $z = \frac{1}{4\cos^2\theta}$ où $\sin h\theta = 0$ et donc que nous pouvons obtenir le développement en fractions rationnelles de $G^{[h]}(z)$ afin d'obtenir une forme trigonométrique du nombre d'arbres de hauteur bornée :

Théorème 2 ([6])

$$G_{n+1}^{[h-2]} = \frac{4^n}{h}\sum_{1\le j\le h/2}\sin^2\frac{j\pi}{h}\cos^{2n}\frac{j\pi}{h}.$$

On peut noter que le théorème 1 était connu par Lord Kelvin (1824-1907) et par Delannoy (1833-1915) [2] et le théorème 2 par Lagrange [14].

2.3 Lois explicites et limites, centrales ou locales

Il est facile d'obtenir les distributions limites en utilisant soit la forme binomiale de la série génératrice soit la forme trigonométrique puisque, si $h = x\sqrt{n}$, d'une part l'approximation de Stirling implique que $\binom{2n}{n-kh} / \binom{2n}{n} \sim e^{-k^2x^2}$ et d'autre part $\cos^{2n}\frac{j\pi}{h} \sim e^{-j^2\pi^2/x^2}$:

Théorème 3 (Loi de la limite locale [8])

$$\mathbb{P}_{\mathcal{G}_{n+1}}\left(H = \lfloor x\sqrt{n}\rfloor\right) \sim \frac{1}{\sqrt{n}}\Theta'(x)\,; \qquad \Theta(x) \simeq \left\{\begin{array}{ll} \sum e^{-k^2x^2}\cdots & \textit{forme binomiale} \\ \sum e^{-k^2\pi^2/x^2}\cdots & \textit{forme trigonométrique}\end{array}\right.$$

Théorème 4 (Théorème central limite [8])

$$\mathbb{P}_{\mathcal{G}_{n+1}}\left(H \le \lfloor x\sqrt{n}\rfloor\right) \to \Theta(x) \quad \textit{où} \qquad \Theta(x) := \sum_{j\ge 1} e^{-j^2x^2}(4j^2x^2 - 2).$$

Pour obtenir les moments de la distribution des hauteurs, on constate que l'espérance $\mathbb{E}_{\mathcal{G}_{n+1}}(H)$ est de la forme $\sum_m d(m)e^{-m^2x^2}$ où $d(m)$ est le nombre de diviseurs de m. On a alors besoin de la quantité $S_r(t) = \sum_h h^r\Theta'(ht)$ quand $t = \frac{1}{\sqrt{n}}$ tend vers 0. Or $S_r(t)$ est une somme de Riemann qu'on peut approximer par une intégrale qui est une transformée de Mellin déguisée à laquelle on peut appliquer un traitement de sommes harmoniques :

Théorème 5 (Moments de la hauteur [6])

$$\mathbb{E}_{\mathcal{G}_{n+1}}(H) = \sqrt{\pi\, n} - \frac{3}{2} + O\left(\frac{1}{\sqrt{n}}\right) \qquad et \qquad \mathbb{E}_{\mathcal{G}_{n+1}}(H^r) = r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}.$$

2.4 Transformations Thêta, fractions continues

En comparant les formes binomiales et trigonométriques du nombre d'arbres de hauteur bornée, on obtient une identité des fonctions Thêta bien connue en théorie des fonctions elliptiques :

$$\frac{1}{\sqrt{\pi\, x}} \sum_{k=-\infty}^{+\infty} e^{-k^2 x^2} = \sum_{k=-\infty}^{+\infty} e^{-k^2\pi^2/x^2}.$$

Notons qu'on peut faire du "Reverse-engineering" à partir de la hauteur des arbres de Catalan. On part de la fonction $f(z) = (1+z)^{2n}$. La multisection de la séries $f(z)$ donne $\sum_h f_{n,h} = \frac{1}{h}\sum_{\omega^h=1} f(\omega)$. On peut alors faire une analyse asymptotique, quand $h = x\sqrt{n}$, des deux formes équivalentes de la fonction (binomiale et trigonométrique). Pour les connexions avec les fonctions Thêta, le mouvement Brownien et l'équation fonctionnelle de la fonction zeta de Riemann, on pourra se reporter à Pólya [17] et Biane, Pitman et Yor [3].

Si on revient à la fraction continue (1), on peut obtenir la fonction génératrice des arbres où u_j marque les sommets au niveau j comme $\cfrac{z\,u_0}{1-\cfrac{z\,u_1}{1-\cfrac{z\,u_2}{\ddots}}}$

, ou ce qui revient au même :

Théorème 6 (Chemins de Dyck et niveau des descentes [9][21][13][18]) *La fonction génératrice des chemins de Dyck où u_j marque les descentes du niveau j au niveau $j-1$ est donnée par*

$$\cfrac{1}{1-\cfrac{z\,u_1}{1-\cfrac{z\,u_2}{1-\cfrac{z\,u_3}{\ddots}}}}.$$

3 Arbres binaires

Les arbres binaires de Catalan sont planaires et seuls les degrés 0 et 2 sont permis. Leur famille $\mathcal{B}$ est décrite par l'équation symbolique : $\mathcal{B} =$

$\mathcal{Z} + \mathcal{B} \times \mathcal{B}$. Sa fonction génératrice $B(z) := \sum B_n z^n$, où la variable z compte les sommets, vérifie donc l'équation $B(z) = z + B^2(z)$ et elle est donnée par la formule $B(z) = \frac{1}{2}(1 - \sqrt{1-4z})$.

La famille $\mathcal{B}^{[h]}$ des arbres binaires de hauteur inférieure ou égale à h a pour série géneratrice $B^{[h]}(z)$, qui vérifie :

$$B^{[0]}(z) = z\,; \qquad B^{[h+1]}(z) = z + \left(B^{[h]}(z)\right)^2.$$

Nous avons des polynômes qui sont définis par une récurrence quadratique. Notons que le degré de ces polynômes double à chaque itération (deg($B^{[h]} = 2^h$).

Par rapport aux arbres généraux de Catalan, la famille $\mathcal{B}$ a aussi une série génératrice algébrique, mais sa série génératrice des arbres de hauteur bornée est un polynôme donné de façon implicite et dont le degré est une exponentielle. On ne connait pas de forme close pour les coefficients de cette dernière et leur asymptotique s'obtient par analyse de singularité.

Sur l'axe des réels, on a :

- pour z entre 0 et 1/4, $B(z) - B^{[h]}(z)$ est dominé par $\sum_{n>h+1} B_n z^n$, ce qui implique une convergence géométrique.
- pour $z > 1/4$, on a une explosion doublement exponentielle.
- à la singularité $z = 1/4$, on ne peut rien dire.

3.1 Iteration de fonctions géneratrices à un point fixe

Dans le plan complexe, on constate que l'on a la récurrence $u_0 = z$ et $u_{h+1} = z + u_h^2$, qui n'est autre que celle qui a rendu les travaux de Mandelbrot célébre et qui détermine de belles figures fractales.

On itére donc la fonction $f(y) = z + y^2$; le point fixe est $\frac{1}{2}(1 - \sqrt{1-4z})$ et le multiplicateur $f'(\xi) = 2\xi = 1 - \sqrt{1-4z}$. En prouvant la convergence de $B(z)$ sur le cercle $|z| = \frac{1}{4}$ et par des arguments de continuité, on obtient :

Lemme 1 *La convergence locale est garantie à l'intérieur de la cardioïde* $|1 - \sqrt{1-4z}| < 1$. *La convergence partant de* $u_0 = z$ *est garantie autour de tous les points* $|z| = \frac{1}{4}$ *si* $z \neq \frac{1}{4}$ *et elle est géométrique.*

3.2 Analyse de singularité

Notons $e_h := B(z) - B^{[h]}(z)$ la série génératrice des arbres de hauteur supérieure à h. Pour avoir l'asymptotique des coefficients $e_{n,h}$ de e_h, on utilise la formule des coefficients de Cauchy :

$$e_{n,h} = \frac{1}{2i\pi} \int_\gamma e_h(z) \frac{dz}{z^{n+1}},$$

avec un contour γ bien choisi. Dans le cas des arbres d'Otter (arbres binaires non planaires), Broutin et Flajolet [4] [5] ont calculé des estimations de $e_h(z)$ pour les valeurs de z à la fois à l'interieur et à l'extérieur du disque de convergence $|z| < \rho$ en utilisant un "tube" autour du cercle $|z| = \rho$ et un "sablier" ancré à ρ.

Dans le cas des arbres binaires planaires, on constate qu'à la singularité $z = 1/4$, $e_{h+1} = e_h(1 - e_h)$. Un argument de convexité implique une convergence vers 0, mais ne dit rien sur sa vitesse puisque $e_{h+1} \sim e_h$!

L'astuce (suggérée par De Bruijn) est de prendre les inverses :

$$\frac{1}{e_{h+1}} = \frac{1}{e_h} \cdot \frac{1}{1 - e_h} = \frac{1}{e_h} \cdot (1 + e_h + e_h^2 + \cdots) = \frac{1}{e_h} + 1 + e_h + e_h^2 + \cdots .$$

On peut alors échanger les bornes inférieures et les bornes supérieures :

$$\frac{1}{e_h} \sim h + \log h + C(e_0) + \cdots \quad \text{et} \qquad e_h \sim \frac{1}{h} - \frac{\log h}{h^2} - \frac{C}{h^2} + \cdots .$$

Il est possible d'établir une relation avec les processus de branchement. Si on considère un événement $\mathcal{E}$ qui a pour fonction génératrice $E(z)$, la probabilité de E pour le proccessus de branchement critique est $2E(1/4)$. On peut remarquer que le proccessus de branchement critique est équivalent au modèle de Boltzmann critique et que

$$\mathbb{P}^{B.P.}(\text{un arbre } \tau) = \frac{1}{2^{2|\tau|+1}}.$$

Corollaire 1 (Processus de branchement *binaire* critique)

$$\mathbb{P}(\mathit{Hauteur} \geq h) \sim \frac{2}{h} \quad \mathit{et} \qquad \mathbb{P}(\mathit{Hauteur} = h) \sim \frac{2}{h^2}.$$

Si maintenant on se place près de la singularité 1/4, dans le sablier, les "écarts" $e_h = y - u_h = \{\text{les arbres de hauteur} > h\}$ vérifient :

$$\left.\begin{array}{rcl} y & = & z + y^2 \\ u_{h+1} & = & z + u_h^2 \end{array}\right\} \qquad \Longrightarrow \qquad e_{h+1} = 2y\left(1 - \frac{e_h}{2y}\right) e_h.$$

En normalisant avec $e_h - (2y)^h f_h$, on obtient la récurrence :

$$f_{h+1} = f_h(1 - (2y)^{h+1} f_h).$$

De nouveau en prenant les inverses, on a :

$$\frac{1}{f_{h+1}} = \frac{1}{f_h} + (2y)^{h+1} + (2y)^{2h+2} f_h + \cdots$$

Lemme 2 (Approximation principale : hauteur$> h$)

$$B - B^{[h]} \approx \epsilon \frac{(1-\epsilon)^h}{1-(1-\epsilon)^h} \quad \textit{avec} \qquad \epsilon := \sqrt{1-4z}.$$

On peut résumer cette partie en disant que du point de vue de la hauteur, les arbres binaires se comportent comme les arbres généraux de Catalan (convergence géométrique pour $z \neq 1/4$, convergence harmonique pour $z = 1/4$). Ces résultats sont encore valables dans un sablier avec des termes d'erreur uniformes.

3.3 Limites locales et centrales

On note $\Theta(x)$ la fonction $\sum_{j\geq 1} e^{-j^2x^2}(4j^2x^2-2)$. On utilise la formule de Cauchy pour les coefficients et un contour de Hankel près de la singularité dans un sablier.

Théorème 7 (Loi de la limite locale [10])

$$\mathbb{P}_{\mathcal{B}_n}\left(H = \lfloor 2x\sqrt{n}\rfloor\right) \sim \frac{1}{\sqrt{n}}\Theta'(x).$$

Théorème 8 (Théorème central limite [10])

$$\mathbb{P}_{\mathcal{B}_n}\left(H \leq \lfloor 2x\sqrt{n}\rfloor\right) \to \Theta(x).$$

On peut donc dire approximativement que $H[\mathcal{B}_n] \approx 2H[\mathcal{G}_n]$.

4 Autres sujets

4.1 Les variétés simples d'arbres

Pour les variétés simples d'arbres, on n'autorise qu'un ensemble fini de sommets. Il y a universalité de la singularité en racine carrée. Il faut utiliser une analyse de perturbation de singularité. On peut se référer au papier de Flajolet, Gao, Odlyzko et Richmond [10]. Pour les arbres de Cayley, le papier de référence est celui de Renyi et Szekeres [19].

4.2 Les arbres non-planaires binaires

Il s'agit des arbres de Otter [16]. Broutin et Flajolet [4] [5] ont récemment étudié la distribution de leur hauteur.

4.3 Vitesse de convergence, larges déviations

Les méthodes précédentes donnent une vitesse de convergence en $\frac{\log n}{\sqrt{n}}$. Par exemple, la hauteur moyenne des arbres binaires est (cf. un papier en préparation de Broutin et Flajolet) :

$$\mathbb{E}_{\mathcal{B}_n}[H] \sim 2\sqrt{\pi\, n} + c \log n + c' + \frac{c'' \log n}{\sqrt{n}} + \cdots$$

La probabilité d'une petite ou d'une grande hauteur est exponentiellement faible :

Théorème 9 (Flajolet, Gao, Odlyzko et Richmond [10]) *Il existe $\delta > 0$ tels que le nombre d'arbres binaires de taille n et de hauteur h, pour $1 \leq h \leq n$ vérifie :*

$$B_n - B_n^{[h]} = O\left(B_n n^{3/2} e^{-h^2/(4n)}\right) \quad \textit{et} \quad B_n^{[h]} = O\left(B_n n^{3/2} e^{-\delta n/h^2}\right).$$

Théorème 10 (Flajolet, Gao, Odlyzko et Richmond [10])

$$B_n^{[h]} - B_n^{[h-1]} \sim \frac{4\epsilon^2 A(\epsilon)}{(1-\epsilon^2)\sqrt{\pi(1+\epsilon)\, n}} \left((1-\epsilon)^{1-\epsilon}(1+\epsilon)^{1+\epsilon}\right)^{-h/2\epsilon} 4^n$$

uniformément pour tout h tel que $h/n = 2\epsilon/(1+\epsilon)$ avec $\epsilon \in [\delta', 1-\delta']$ où δ' est une constante positive qui peut être arbitrairement petite et $A(\epsilon)$ est une fonction continue et positive pour $\epsilon \in [\delta', 1-\delta']$.

4.4 Structures équilibrées

Théorème 11 (Flajolet et Odlyzko [11]) *Les coefficients des polynômes $p_h(z)$ qui vérifient*
$p_{h+1}(z) = P(z, p_h(z))$ pour un polynôme P non-linéaire à coefficients positifs suivent localement une loi Gaussienne.

$p_h(1)$ croît en double exponentielle et vérifie la formule exacte $p_h(1) = \lfloor \alpha^{2^h} \rfloor$. On utilise alors une méthode de perturbation et une méthode de col.

Application : si on regarde la distribution des tailles des arbres binaires de hauteur h ($p_{h+1}(z) = z + p_h^2(z)$ et $p_0(z) = z$) et si on regarde la distribution des tailles des arbres équilibrés binaires-ternaires de hauteur h ($p_{h+1}(z) = p_h^2(z) + p_h^3(z)$ et $p_0(z) = z$), on constate qu'elles suivent la même loi, c'est-à-dire celle des polynômes vérifiant $p_{h+1}(z) = p_h^2(z)$ (i.e. $p_h(z) = (1+z)^{2^h}$).

4.5 Diamètre et largeur

L'étude du diamètre des arbres de Cayley non-enracinés a été faite par Szekeres en 1982 [20]. Celle du diamètre des arbres binaires non-planaires non-enracinés (arbres de Otter) a été faite par Broutin et Flajolet en 2010 [5].

La distribution fait intervenir une fonction Thêta. Le rapport entre le nombre d'arbres non-planaires à un seul centre et le nombre d'arbres à deux centres est en accord avec le même rapport pour le modèle CRT d'Aldous [1] ("continuum random tree") dans lequel les arbres sont planaires.

Théorème 12 *Le rapport l'espérance du diamètre et l'espérance de la hauteur des arbres de Otter vérifie :*

$$\lim_{n\to\infty} \frac{\mathbb{E}_n(D)}{\mathbb{E}_n(H)} = \frac{4}{3}.$$

Ce rapport est identique à celui trouvé par Szekeres [20] pour les arbres étiquetés. Aldous [1] a montré de façon semi-heuristique qu'il est universel pour toutes les familles d'arbres qui ont un ordre sur les fils d'un sommet. On peut aussi consulter [12] et [15] pour l'extension du modéle CRT d'Aldous aux arbres non-planaires.

On calcule la largeur en étudiant la hauteur du mouvement Brownien (cf. le travail de Chassaing, Market et Yor [7] qui fait autorité). Pour les arbres de Cayley et les variétés simples d'arbres :

$$\mathbb{E}_n(W) = \sqrt{\frac{\pi\, n}{2}} + O\left(n^{1/4}\sqrt{\log n}\right) \quad , \quad \mathbb{P}_n(\sqrt{2}W \leq x) \to 1 - \Theta(x).$$

Pour atteindre la largeur par une méthode d'analyse, on pense naturellement à la méthode des matrices de transfert.

5 Message subliminal ou conclusion

Asymptotiquement, la seule différence entre les diverses formules de hauteur (voire largeur) est, in fine, une constante de normalisation calculable à parir de la définition des arbres.

Références

[1] D. Aldous. The continuum random tree. II. An overview. In Stochastic analysis (Durham 1990), volume 167 of *London Math. Soc. Lecture Note Ser.*, Cambridge Univ. Press, Cambridge, 1991, pp.23-70.

[2] J.-M. Autebert, A.-M. Décaillot and S.R. Schwer. H.-A. Delannoy et les œuvres posthumes d'Edouard Lucas. Gazette des Mathématiciens **95** (janvier 2003).

[3] P. Biane and J. Pitman and M. Yor. Probality laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. Bull. Amer. Math. Soc. (N.S.) **38** (2001), 435-465.

[4] N. Broutin and Ph. Flajolet. The height of random binary unlabelled trees. In Fifth Colloquium on Mathematics and Computer Science : Algorithms, Trees, Combinatorics and Probabilities, DMTCS Proceedings, vol. AI, Blaubeuren, 2008, pp. 121-134, http ://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/article/viewArticle/dmAI0106.

[5] N. Broutin and Ph. Flajolet. The distribution of height and diameter in random non-plane binary trees. arXiv :1009.1515, 2010.

[6] N.G. De Bruijn, D.E. Knuth and S.O. Rice. The average height of planted plane trees. In Graph Theory and Computing (R.C. Read, éd.), Academic Press, 1972, pp. 15-22.

[7] P. Chassaing, J.-F. Marckert and M. Yor. The height and width of simple trees. In *Mathematics and Computer Science : Algorithms, Trees, Combinatorics and Probabilities*, Versailles, (D. Gardy and A. Mokkadem, eds), Birkhauser, 2000, pp. 17-30.

[8] Ph. Flajolet and R. Sedgewick. Analytic Combinatorics. Cambridge University Press, 2009.

[9] Ph. Flajolet. Combinatorial Aspects of Continued Fractions. Discrete Mathematics **32** 1980, 125-161. Reprinted in the 35th Special Anniversary Issue of Discrete Mathematics, Volume 306, Issue 10–11, Pages 992-1021 (2006).

[10] Ph. Flajolet, Z. Gao, A. Odlyzko and B. Richmond. The distribution of heights of binary trees and other simple trees. Combinatorics, Probability, and Computing **2** (1993), 145-156.

[11] Ph. Flajolet and A. Odlyzko. Limit distributions for coefficients of iterates of polynomials with applications to combinatorial enumeration. Math. Proc. Cambridge Phil. Soc. **96** (1984), 237-253.

[12] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. arXiv :1003.3632v2, 2010.

[13] D.M. Jackson. Somme results on "Product-weighted led codes". J. Combinatorial Theory (Ser. A) **25** (1978), 181-187.

[14] J.L. Lagrange. Recherches sur les suites récurrentes dont les termes varient de plusieurs manières, ou sur l'Intégration des équations linéaires aux différences finies et partielles ; et sur l'usage de ces équations dans la théorie des hasards, in Œuvres de Lagrange (également in Nouveaux mémoires de l'Académie royale des sciences et belles-lettres de Berlin, année 1775). Vol. 4, pp. 149-251, Paris, 1869.

[15] J.-F. Marckert and G. Miermont. The CRT is the scalling limit of unordered binary trees. Random Structure and Algorithms **96** (2010). To appear. arXiv :0902.4570.

[16] R. Otter. The number of trees. Ann. of Math. (2) **49** (1948), 583-599.

[17] G. Pólya. Elementarer Beweis einer Thetaformel. Sitzungsberichten der Preuß. Akad. des Wissenschaften, **3** (1927), 157-161.

[18] R.C. Read. The chord intersection problem. In *Second International Conference on Combinatorial Mathematics*, Annals of New-York Ac. of Sc., vol. 319, 1979, pp. 444-454.

[19] A. Rényi and G. Szekeres. On the height of trees. Journal of the Australian Mathematical Society **7** (1967), 497-507.

[20] G. Szekeres. Distribution of labelled trees by diameter. In *Combinatorial Mathematics X*, Proc. 10th Australian Conference on Combinatorial Mathematics, Springer Lecture notes in Mathematics, 1982, pp. 392-397.

[21] J. Touchard. Sur un problème de configurations et sur les fractions continues. Canad. J. Math **4** (1952), 2-25.