



U.F.R. S.T.M.I.A.
Ecole Doctorale I.A.E.+M.
Université Henri Poincaré Nancy I
D.F.D. Mathématiques.

THÈSE

présentée pour l'obtention du

Doctorat de l'Université Henri Poincaré Nancy I

(Spécialité Mathématiques Appliquées)

par

Jean-François Marckert

Marches Aléatoires, Arbres

et

Optimalité d'Algorithmes

soutenue publiquement le 7 décembre 1999 devant le jury composé de :

Philippe CHASSAING	Directeur de thèse	Université Nancy I
Philippe FLAJOLET	Rapporteur, Président du Jury	INRIA Rocquencourt
Guy LOUCHARD	Rapporteur	Université Libre de Bruxelles
Bernard ROYNETTE	Examineur	Université Nancy I
René SCHOTT	Examineur	INRIA Nancy
Pierre VALLOIS	Examineur	Université Nancy I
Marc YOR	Examineur	Université Paris VI

Institut Elie Cartan Nancy

A ma famille,

Jean-Michel, Thérèse, Virginie, Emmanuelle et Laurence

Avant Propos

Je tiens d'abord à remercier Philippe Chassaing avec qui j'ai beaucoup travaillé durant ces trois années et sous la direction duquel ce travail a été réalisé. Sa grande disponibilité, son soutien amical, son enthousiasme ainsi que ses critiques m'ont permis de progresser et de mener à bien cette thèse.

Je remercie Philippe Flajolet et Guy Louchard pour l'intérêt qu'ils ont porté à ce travail et l'honneur qu'ils m'ont fait d'accepter d'être rapporteur.

Je remercie René Schott d'avoir accepté d'examiner cette thèse et pour la sympathie qu'il m'a toujours témoignée.

Je sais également gré à Bernard Roynette et Pierre Vallois d'avoir accepté de faire partie de ce jury. J'apprécie à leurs justes valeurs les conseils amicaux et les critiques qu'ils m'ont prodigués tout au long de ces trois années.

Ce fut un réel honneur et plaisir pour moi de travailler avec Marc Yor que je remercie également de participer à ce jury.

Enfin, j'aimerais insister sur la très bonne ambiance de travail et d'amitié que j'ai rencontré au sein de l'équipe de probabilité de Nancy. En particulier, j'aimerais remercier Mădălina Deaconu, Jean-Sébastien Giet et Etienne Tanré pour les nombreuses discussions qui ont égayées ces trois années.

Table des matières

INTRODUCTION	1
Références	10
CHAPITRE I	
An average case quasi optimal search algorithm for the zeros of a random walk	11
1 Introduction	13
2 The “double-jump” algorithm	15
2.1 A set of moderate variation	15
2.2 Description of the “double-jump” algorithm	16
3 Quasi-optimality of $dj(n)$	16
3.1 Sketch of proof of Theorems 1.2 and 1.3	16
3.2 Proof of Theorem 3.1	18
3.3 Proof of Theorem 3.2	20
3.4 Proof of Theorem 3.3	23
4 Other case	25
References	30
CHAPITRE II	
A Stochastically Quasi-Optimal Search Algorithm for the Maximum of the Simple Random Walk	31
1 Introduction	33
2 Optimality criteria	35
3 Main steps of proofs of Theorems 1.1 and 1.2	37
4 Proof of Theorem 1.3	40

5	A counting principle	48
6	Proof of Theorem 3.3	55
	References	62

CHAPITRE III

	Stochastic properties of two algorithms searching the zeros of simple random walks, and applications	67
--	---	-----------

1	Introduction	69
1.1	The simple and double jump algorithms	69
1.2	Trees and Brownian excursions	72
2	Quasi-Optimality of $d_j(n)$ for the stochastic order	73
2.1	Sketch of the proofs of Theorems 1.3 and 1.4	73
2.2	A counting principle	74
2.3	Proof of Theorem 2.4	79
3	Relation between the search of zeros, trees and excursions	80
3.1	The couple height-width	81
3.1.1	Excursions	82
3.1.2	Expression of $(H_B^{(n)}, W_B^{(n)})$ in term of Bernoulli excursions . . .	84
3.2	Harmonic mean of depth of nodes in size n general trees	90
3.2.1	Table of convergences	92
	References	93

CHAPITRE IV

	Parking functions, empirical processes and the width of rooted labeled trees	95
--	---	-----------

1	Introduction	97
1.1	Rooted labeled trees	97
1.2	Results	98
1.3	Motivations	99

2	Rooted labeled trees, parking functions and empirical processes	99
2.1	The breadth first search random walk	99
2.2	Rooted labeled trees and parking functions	101
2.3	Short and loose historic of empirical processes	103
2.4	Parking functions and empirical processes	104
2.5	Generation of a random rooted labeled tree	106
3	Bounds for moments	107
3.1	Proof of Theorem 3.1	107
3.2	Proof of Theorem 2.1	109
4	Concluding remarks	111
	References	113
	Appendice	116
 CHAPITRE V		
	La montée de barre optimale en présence d'une échéance aléatoire	119
1	Introduction	121
1.1	Cadre général	121
1.2	Quelques exemples	123
2	Le cas T de loi exponentielle	124
2.1	Une équation de Bellman	125
2.2	"Unicité" de la solution de l'équation de Bellman	126
2.3	Lien avec un système différentiel "exotique"	128
2.4	Preuve de la Proposition 2.7	130
2.5	Preuve de la Proposition 2.8	131
2.6	Une "sous" solution	132
2.7	Existence d'une solution z	132
2.8	Approximation de z et de Ψ	133
	Références	135

ANNEXES	137
1 Processus browniens	139
1.1 Marches aléatoires et excursions discrètes	139
1.2 Mouvement brownien	140
1.3 Méandre brownien	144
1.4 Pont brownien	144
1.5 Excursion brownienne renormalisée	145
1.6 Décompositions	145
2 Arbres - Processus de Galton-Watson	147
2.1 Processus de Galton-Watson	147
2.2 Arbres de Galton-Watson	149
2.3 Hauteur et largeur des arbres	152
2.4 Parcours d'arbres	153
2.5 Largeur de bande (bandwidth)	156
2.6 Arbres binaires de recherche	157
2.7 Arbres et combinatoire	158
Références	160

Introduction

En informatique comme dans beaucoup d'autres domaines, une tâche peut souvent être exécutée de diverses manières. Chacun des procédés, des *algorithmes*, permettant d'obtenir le résultat désiré, possède ses caractéristiques propres et le choix de l'un ou l'autre n'est en général pas anodin. On compare deux algorithmes en mesurant leurs efficacités respectives en terme de temps, de place mémoire utilisée ou encore en nombre d'utilisations d'une sous tâche élémentaire caractéristique de cette efficacité. Par exemple, il existe de nombreux algorithmes permettant de ranger par ordre croissant une liste de n nombres. La plupart de ces algorithmes ont pour point commun d'utiliser et de reposer sur la tâche élémentaire suivante : la comparaison de deux nombres. Ainsi, dans ce cas, on identifie le *coût* de chaque algorithme de tri avec le nombre de comparaisons qu'il effectue (voir [7], [10], [15]).

L'analyse du coût des algorithmes informatiques est une science qui a l'âge des premiers ordinateurs. On peut être enclin à penser que l'intérêt de cette science diminue au fur et à mesure que les ordinateurs deviennent plus puissants. C'est en partie vrai pour les petites applications. Mais, en dehors du fait qu'elle mène à de jolis problèmes combinatoires ou probabilistes (on pourra voir à ce sujet les livres de Knuth [10] et de Flajolet & Sedgewick [15]), l'importance de cette science reste considérable en pratique et risque de l'être encore très longtemps. Pour illustrer l'importance du choix d'un algorithme, on peut donner l'exemple spectaculaire suivant (voir Ciarlet [6]) : le calcul du déterminant d'une matrice $n \times n$ par la méthode de Gauss nécessite de l'ordre de n^3 opérations. Un ordinateur modeste peut donc en quelques secondes, calculer le déterminant d'une matrice 100×100 . Ce même calcul par la méthode de Cramer demande, quant à lui, de l'ordre de $(n+2)!$ opérations, c'est-à-dire pour le cas qui nous intéresse environ 10^{162} opérations et on peut douter du fait qu'un jour, il existe un ordinateur suffisamment puissant pour venir à bout de ces calculs. Dans les applications gourmandes en calculs, C.A.O., météorologie, résolutions de grands systèmes linéaires, approximation des solutions des équations différentielles..., l'efficacité des méthodes et algorithmes employés est donc primordiale et fait l'objet de beaucoup de recherche comme en témoigne la vaste littérature à ce sujet.

Pour comparer les algorithmes, on munit l'ensemble des tâches d'une mesure de taille. Par exemple, dans le cas du tri, le nombre d'éléments à ranger sera la taille du problème, dans le cas des matrices, ce sera la taille de la matrice. On compare alors les algorithmes lors de leurs exécutions sur des tâches de même taille. Un problème se pose alors : en général, le coût d'un algorithme n'est pas une fonction de la taille du problème, c'est-à-dire qu'il peut prendre des valeurs très différentes pour un taille donnée.

Par exemple, considérons le calcul du produit de deux entiers de longueur n à la main par la méthode habituelle ; la *complexité* des nombres employés prend une part déterminante dans la vitesse de calcul (pour un calculateur normal!). Il est plus rapide

de calculer 10000×10000 que 67897×79876 . Ce phénomène apparaît à peu près sur tous les types de problèmes et à peu près pour tous les algorithmes.

Un autre problème qui apparaît également est que le nombre d'algorithmes à notre disposition est une fonction de la taille. Reprenons le cas du problème du tri de n nombres. Si $n = 2$ il n'y a clairement qu'un seul algorithme, ce qui n'est plus le cas si n est plus grand. Chercher le meilleur algorithme pour un problème donné, ne pourra donc se faire qu'à taille fixée et devra tenir compte des fluctuations du coût sur les différents problèmes. Ainsi, divers types d'optimalité sont considérés par les informaticiens et les deux types les plus fréquemment étudiés sont les suivants (la taille du problème est fixée) :

- L'optimalité *dans le pire des cas* : un algorithme est dit optimal dans le pire des cas si son coût maximal est inférieur au coût maximal de tous les autres algorithmes. Certains auteurs Cormen et al. [7] ou encore Aho et al. [1] s'intéressent tout particulièrement à cette optimalité. Utiliser un algorithme optimal dans le pire des cas permet d'être sûr d'avoir le résultat escompté avec un coût maximal connu à l'avance et garantit donc une certaine sécurité. On peut reprocher à ce critère de n'être pas très révélateur de l'efficacité d'un algorithme : souvent le coût maximal est considérable par rapport au coût moyen et il est atteint sur des données très peu probables.

- L'optimalité *en moyenne*. On suppose que les tâches (de taille fixée) se présentent selon une loi de probabilité μ . Un algorithme est dit optimal en moyenne s'il a réalisé le coût moyen minimal par rapport à μ . Par exemple, pour le problème du tri de n nombres, on considère que les $n!$ permutations sont équiprobables. Les résultats d'optimalité en moyenne sont relativement rares et difficiles à obtenir ; les méthodes utilisées sont, la plupart du temps, particulières au problème rencontré et donc non applicables à d'autres cas. Le Chapitre II dresse l'historique (probablement non exhaustif) des résultats obtenus pour ce type d'optimalité.

Afin de mieux connaître les algorithmes, on s'intéresse depuis peu à la loi ou à la limite en loi de leur coût. On peut citer par exemple les travaux de G. Louchard [12] qui fut l'un des premiers à utiliser les outils probabilistes afin d'analyser les algorithmes. Connaître la loi du coût d'un algorithme donne évidemment toutes les informations souhaitables. Elle permet bien sûr de calculer le coût moyen, de connaître le coût maximal ou minimal, mais surtout elle permet d'analyser le comportement de l'algorithme lorsque l'on doit l'exécuter un grand nombre de fois. En pratique, il s'avère souvent que l'on ne parvient à obtenir qu'un comportement limite de ces lois lorsque la taille n des problèmes tend vers l'infini et après renormalisation.

Le travail présenté ici porte sur l'étude de certains algorithmes et sur leur optimalité.

La première partie traite de résultats obtenus lors de l'étude des algorithmes de recherche de valeurs attachées à des marches aléatoires simples symétriques.

La deuxième partie est l'étude d'un problème d'optimisation de ressources provenant de l'informatique.

Avant de décrire succinctement le contenu des chapitres, il me semble nécessaire de

signaler que des annexes placées en fin de thèse donnent certains éclaircissements sur les objets mathématiques utilisés; en particulier, on y trouvera des éléments sur les marches aléatoires, sur les processus browniens et sur les arbres.

Première partie :

Elle est composée de quatre chapitres. La plupart des questions auxquelles je me suis intéressé dans cette première partie ont pour origine commune le problème résolu par A.M. Odlyzko [13] en 1995. Dans son article *Search for the maximum of a random walk*, Odlyzko s'intéresse au problème suivant :

Un oracle a tiré au sort de façon équiprobable une marche aléatoire simple symétrique $\omega = (S_k(\omega))_{k=0, \dots, n}$, $S_0(\omega) = 0$ parmi les 2^n marches aléatoires de longueur n . On note

$$M_n(\omega) = \max\{S_k(\omega) | 0 \leq k \leq n\},$$

le maximum de la marche sur $\{0, \dots, n\}$. On doit payer 1 à l'oracle pour connaître la valeur de $S_l(\omega)$ pour une position l de notre choix. On se propose de trouver $M_n(\omega)$ en dépensant le moins possible. La somme nécessaire pour trouver la valeur M_n dépend donc de la marche aléatoire bien sûr mais aussi de la stratégie employée (on emploiera indifféremment les mots *stratégie* et *algorithme* pour désigner la suite d'opérations que l'on exécute pour trouver le maximum). La question que l'on se pose est donc la suivante : Existe-t-il une stratégie qui permet de minimiser le coût de la découverte de M_n et si oui, laquelle ?

Parmi les différents résultats obtenus par Odlyzko, le principal est celui de l'analyse en moyenne du coût de ces algorithmes. Ainsi, il prouve que pour trouver le maximum M_n , on devra payer au moins $c_0\sqrt{n}(1+o(1))$ en moyenne et ceci quelque soit la stratégie employée, c_0 étant une constante décrite dans le Chapitre II. Il donne ensuite un algorithme que l'on appellera $Od(n)$ (décrit dans le Chapitre II, Section 5) dont le coût est également $c_0\sqrt{n}(1+o(1))$. L'algorithme $Od(n)$ est dit *quasi optimal* en moyenne, *quasi* signifiant que le coût de $Od(n)$ et la borne inférieure sont équivalents. Pour donner une borne inférieure au coût des algorithmes, il étudie les points $(m_i^{(a)}, S_{m_i^{(a)}})_{i \in I}$ sondés pendant la recherche du maximum par un algorithme a quelconque. Ces points sondés doivent garantir que le maximum trouvé est le vrai maximum ce qui mène à la condition nécessaire (*) suivante vérifiée par ces points : (on suppose ici que les $(m_i^{(a)})_i$ sont rangés par ordre croissant)

$$S_{m_{i+1}^{(a)}} + S_{m_i^{(a)}} + m_{i+1}^{(a)} - m_i^{(a)} \leq 2M_n.$$

Odlyzko associe alors à chaque entier k de $\{m_i^{(a)} + 1, \dots, m_{i+1}^{(a)}\}$ un poids $W^{(a)}(k) = 1/(m_{i+1}^{(a)} - m_i^{(a)})$ exprimant la *densité* de points sondés par a "autour" de k ; de cette façon, le coût de l'algorithme a est égal à la somme des $W^{(a)}(k)$. Ainsi, en utilisant (*),

Odlyzko prouve que le nombre moyen de points sondés possède une borne inférieure :

$$v_n = (1 + o(1))\mathbb{E}\left(\frac{1}{2} \sum_{k=1}^n \frac{1}{M_n - S_k + 1}\right).$$

Il fournit alors un algorithme $Od(n)$ dont le coût est asymptotiquement le même. Ces arguments de densité sont repris dans le Chapitre II et systématisés.

Chapitre I

Après la recherche du maximum, il est de coutume de s'intéresser à la recherche des racines. Le Chapitre I consiste donc en l'analyse en moyenne des stratégies de recherche de tous les zéros de ω , c'est-à-dire des indices k tels que $S_k(\omega) = 0$ pour $k \in \{0, \dots, n\}$. Certaines techniques mises en oeuvre par Odlyzko, sont employées ici, en particulier la notion de *poids* d'un point. Les principales différences proviennent du fait que le temps passé par la marche aux alentours du maximum est très faible relativement à celui passé autour de zéro. On construit deux algorithmes que l'on nomme $sj(n)$ et $dj(n)$; $sj(n)$ (*the single jump algorithm*) est l'algorithme le plus simple que l'on puisse construire. Il consiste à sonder la marche de gauche à droite et à sonder le premier endroit où un zéro est possible. Par exemple, si l'on sait que $S_{20} = 6$, le prochain point sondé sera 26, puisque la marche ne peut pas s'annuler entre l'indice 20 et l'indice 26. Une description de $sj(n)$ est donnée au début du Chapitre I. Malgré sa simplicité cet algorithme et ses propriétés sont fondamentaux dans le Chapitre III.

L'algorithme $dj(n)$ (*the double jump algorithm*) ne sera pas décrit ici. Comme son nom l'indique, il consiste (moralement...) à faire des sauts de longueur deux fois plus grande qu'avec le simple saut et revenir en arrière s'il existe un risque d'avoir manqué un zéro. En utilisant la notion de densité de points sondés d'Odlyzko, on donne une borne inférieure au coût moyen de tout algorithme et on montre que le coût de $dj(n)$ atteint cette borne, en tout cas pour le premier terme du développement asymptotique. Ce coût minimal moyen est donné par la quantité :

$$w_n = (1 + o(1))\mathbb{E}\left(\frac{1}{2} \sum_{k=1}^n \frac{1}{|S_k| + 1}\right)$$

et on montre alors que

$$\frac{\mathbb{E}(W_n)}{\sqrt{n} \log n} \xrightarrow{n \rightarrow +\infty} \frac{1}{\sqrt{2\pi}}.$$

Suite à un exposé de ce résultat, Ph. Flajolet a suggéré de nous intéresser aux lois limites de ces deux algorithmes ainsi qu'à celui d'Odlyzko, ces lois ne semblant pas classiques *a priori* ni même après simulation. Cette recherche s'est révélée vite fructueuse, la loi limite de l'algorithme d'Odlyzko n'étant pas connue et celle des algorithmes $sj(n)$ et $dj(n)$ nous amenant à étudier les bijections existant entre les excursions de Bernoulli,

les arbres binaires et les arbres généraux ; ainsi, on a pu obtenir la loi limite du couple hauteur-largeur dans les arbres binaires de taille n , la loi limite de la moyenne harmonique de la hauteur des feuilles de l'arbre général de taille n ...

Les deux chapitres suivants ont été l'objet de deux articles écrits avec Philippe Chassaing et Marc Yor. La première motivation était, à l'origine, de trouver la loi limite des deux algorithmes $Od(n)$ et $dj(n)$, mais comme on va le voir, ces objectifs ont été débordés assez largement.

Chapitre II

Ce deuxième chapitre traite donc de l'étude des propriétés stochastiques du coût de l'algorithme $Od(n)$. En premier lieu, on affine légèrement le raisonnement d'Odlyzko pour prouver plus que l'optimalité en moyenne de son algorithme. Notons $T_{Od(n)}$ le coût de $Od(n)$ et V_n la v.a.

$$V_n(\omega) = \frac{1}{2} \sum_{k=1}^n \frac{1}{M_n(\omega) - S_k(\omega) + 1}.$$

On montre alors que

$$\|V_n - T_{Od(n)}\|_1 = o(\sqrt{n})$$

et que

$$\frac{V_n}{\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{loi} I$$

où I est une fonctionnelle du mouvement brownien standard \mathcal{B} et de son maximum $M_1^{\mathcal{B}}$ sur l'intervalle $[0, 1]$:

$$I = \frac{1}{2} \int_0^1 \frac{dt}{M_1^{\mathcal{B}} - \mathcal{B}_t}.$$

Cette convergence n'est pas une conséquence du principe de Donsker (ni du théorème de Komlos-Major-Tusnady), la fonctionnelle mise en jeu étant trop irrégulière. On utilise alors le plongement de Skohorod qui consiste à construire le mouvement brownien \mathcal{B} et une marche aléatoire ω sur le même espace de probabilité ; on prouve alors la convergence de V_n/\sqrt{n} vers I "à la main". La loi de la limite I que l'on obtient se révèle difficile à appréhender. La Section 4 du Chapitre II fait le tour des propriétés de I que l'on a obtenues. On peut remarquer que ces propriétés découlent toutes de la transformée de Laplace de $1/I^2$ qui identifie I entièrement :

$$\mathbb{E} \left[\exp \left(- \frac{t}{I^2} \right) \right] = \frac{2 \log(\cosh \sqrt{2t})}{\sinh^2 \sqrt{2t}}.$$

En particulier, bien que l'on ne réussisse pas à inverser explicitement cette transformée ni à exprimer celle de I , on parvient à calculer tous les moments (entiers) de I en fonction de la fonction ζ de Riemann. La densité de I est également exprimée comme produit de convolution de deux densités de types Θ (s'exprimant en fonction de la fonction *theta* de Jacobi), densités connues comme étant celle de certaines v.a. liées à des processus browniens (on pourra voir par exemple l'article de Biane & Yor [5]).

L'autre point de cet article que je voudrais souligner est la preuve de la quasi-optimalité pour l'ordre stochastique du coût de $Od(n)$. C'est à notre connaissance, la première fois que ce type de résultat est établi. Une v.a. X est dite inférieure stochastiquement à une v.a. Y si pour tout $x \in \mathbb{R}$, on a :

$$\mathbb{P}(X \geq x) \leq \mathbb{P}(Y \geq x).$$

Ainsi, les théorèmes 1.1 et 1.2 du chapitre II prouvent que la limite en loi de $T_{Od(n)}$ (normalisé par \sqrt{n}) est celle de I et que le coût de tout autre algorithme (normalisé par \sqrt{n}) est supérieur stochastiquement à I .

L'optimalité pour l'ordre stochastique du coût d'un algorithme A parmi tous les algorithmes donne une garantie que ne donne pas sa seule optimalité en moyenne. En effet, dans le contexte des files d'attente, on voit souvent l'ordinateur comme un serveur, les tâches à traiter comme des clients et le coût d'un l'algorithme comme le temps de service. Comme le montre la discussion de la partie 2 du Chapitre II, l'optimalité pour l'ordre stochastique de A garanti que le temps moyen de séjour dans le système par un client est minimal avec A parmi tous les autres algorithmes ; ceci n'est pas garanti par la seule optimalité en moyenne.

Chapitre III

Ce chapitre contient l'analyse de la limite en loi des algorithmes $sj(n)$ et $dj(n)$ et la preuve du caractère quasi-optimal pour l'ordre stochastique de $T_{dj(n)}$, le coût de $dj(n)$. Notons W_n la v.a. définie par

$$W_n(\omega) = \frac{1}{2} \sum_{k=1}^n \frac{1}{|S_k(\omega)| + 1},$$

On montre que :

$$\frac{W_n}{\sqrt{n} \log n} \xrightarrow{\text{loi}} |N|$$

où N est une normale réduite centrée.

Le deuxième point à souligner est l'analyse de l'algorithme $sj(n)$ et en particulier l'étude de son coût pour traverser une excursion de Bernoulli. On sait en effet que les excursions de Bernoulli de longueur $2n$ sont en bijection avec les arbres binaires ayant $n-1$ noeuds internes et avec les arbres que l'on appellera généraux conditionnés à avoir

n noeuds. Certaines de ces bijections ont des propriétés tout à fait remarquables. Par exemple, il existe une bijection entre les excursions de Bernoulli et les arbres généraux qui transforme le maximum de l'excursion en la hauteur de l'arbre associé (à une unité près). Lors de l'étude de l'algorithme $sj(n)$, on a remarqué que le nombre de coups de sondes nécessaires à la traversée d'une excursion de longueur $2n$ était égal à la hauteur de l'arbre binaire associé à l'excursion lors de son parcours en largeur. Cette loi est connue comme étant une loi Θ (voir Flajolet & Odlyzko [9]).

D'autre part, comme pour les problèmes précédents, on prouve que le coût de la traversée d'une excursion de Bernoulli $\tilde{\omega} = (\tilde{S}_k(\tilde{\omega}))_{k=0, \dots, 2n}$ est approximativement :

$$\tilde{V}_{2n}(\tilde{\omega}) = \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k(\tilde{\omega})}.$$

Ces considérations mènent à un calcul élémentaire de la limite en loi du couple (hauteur, largeur) dans un arbre binaire ayant n feuilles. On donne une preuve élémentaire de

$$\frac{1}{\sqrt{2n}}(H_B^{(n)}, W_B^{(n)}) \xrightarrow{\text{loi}} \left(\int_0^1 \frac{1}{e_s} ds, M_1^{(e)} \right)$$

où $H_B^{(n)}$ (resp. $W_B^{(n)}$) désigne la hauteur (resp. la largeur) d'un arbre binaire ayant n feuilles. Ce résultat est également un corollaire direct d'un théorème récent d'Aldous [4, th. 3] (beaucoup plus général) dont la preuve repose sur trois papiers particulièrement techniques : Aldous [3], Jeulin [11] et Drmota & Gittenberger [8].

L'utilisation de ce résultat et de la bijection entre arbres généraux (à n noeuds) avec les excursions de Bernoulli (à $2n$ pas) permet d'obtenir la loi de la moyenne harmonique des hauteurs des noeuds d'un arbre général conditionné par sa population totale.

Chapitre IV

Ce chapitre porte sur l'étude des moments de la largeur des arbres non ordonnés, étiquetés, conditionnés à avoir n noeuds (les arbres de Cayley de taille n). C'est un travail commun avec Philippe Chassaing.

La largeur d'un arbre T est le nombre maximal de noeuds sur un des niveaux de l'arbre ; autrement dit, si on note $Z_0, Z_1, \dots, Z_{H(T)}$ le nombre de noeuds à la distance $0, 1, \dots, H(T)$ de la racine ($H(T)$ étant la hauteur de T), la largeur $W(T)$ de l'arbre T est définie par :

$$W(T) = \max\{Z_0, Z_1, \dots, Z_{H(T)}\}.$$

À propos de la largeur des arbres de Cayley de taille n on peut citer principalement deux résultats :

- Odlyzko & Wilf [14] ont prouvé qu'il existait deux constantes positives C_1 et C_2 (ne dépendant pas de n) telles que :

$$C_1 \sqrt{n} \leq \mathbb{E}(W_n) \leq C_2 \sqrt{n \log n}.$$

- En ce qui concerne la loi limite de la largeur, Takács [16] a prouvé que :

$$\mathcal{L}\left(\frac{W_n}{\sqrt{n}}\right) \xrightarrow{\text{loi}} m$$

où m est le maximum de l'excursion brownienne normalisée.

Notre travail répond à la question de Odlyzko & Wilf quant au terme principal du développement asymptotique de W_n . En effet, on prouve que :

Pour tout $p \geq 1$, il existe une constante C_p telle que

$$|E(n^{-p/2}W_n^p) - E(m^p)| \leq C_p n^{-1/4} \sqrt{\log n}$$

Ainsi, la convergence prouvée par Takács est également une convergence des moments.

Pour obtenir ce résultat, on rappelle tout d'abord la correspondance exacte existant entre les arbres étiquetés et le problème de parking. On exhibe ensuite un lien fort entre le problème de parking et le processus empirique, cher aux statisticiens. Il apparaît alors par un argument de couplage (en fait le théorème de Komlos, Major et Tusnady) que l'on peut prouver la convergence des moments de la largeur de l'arbre vers ceux de l'excursion brownienne renormalisée.

En appendice de ce chapitre, on trouvera une preuve (légèrement différente de celle proposée par Takács) que les moments de la largeur (normalisée) des arbres binaires de taille n convergent vers ceux de l'excursion brownienne normalisée. Les premiers arguments utilisés (similaires à ceux du cas des arbres de Cayley), en particulier, l'utilisation du parcours en largeur, montrent que ce problème n'est pas si éloigné qu'il y paraît des chapitres précédents.

Deuxième partie :

Chapitre V

La deuxième partie est composée d'un seul chapitre relativement étranger aux précédents et malheureusement, il n'a pu être totalement achevé, un léger (??!) problème nous entravant jusqu'à aujourd'hui. C'est un travail commun avec Philippe Chassaing.

Le problème auquel on s'intéresse ici est issu d'un problème d'optimisation de ressources issu de l'informatique et plus particulièrement de l'intelligence artificielle.

On a vu au début de l'introduction, des algorithmes (tri, calculs) qui permettent de donner une réponse *exacte* à un certain type de questions. Par opposition, pour nombre de problèmes, la qualité de la réponse dépend d'un paramètre, par exemple du temps. On peut par exemple penser à la digitalisation d'une image ou même plus simplement au calcul numérique d'une intégrale, à l'approximation décimale de π ... Les algorithmes effectuant ces tâches se regroupent grossièrement en deux classes : les algorithmes

interruptibles et les autres. Les algorithmes interruptibles comme leurs noms l'indiquent ont la propriété agréable de fournir un résultat même s'ils sont interrompus en cours d'exécution (par exemple, la recherche d'une racine par dichotomie); la plupart des algorithmes ne possèdent pas cette propriété. En général, un algorithme interrompu ne donne pas de résultat valable. (Tri, calcul d'une intégrale par la méthode de Newton, des trapèzes, etc... C'est également le cas pour la digitalisation d'une image avec les algorithmes classiques. Si on l'interrompt, on obtient une image tronquée et non pas une image de plus faible précision.)

Supposons que l'on ait toute une classe d'algorithmes $(A(t))_{t \in I}$ de durée t non interruptibles permettant de traiter une certaine tâche. Le traitement de la tâche par l'algorithme $A(t)$ donne un résultat auquel on associe une valeur $f(t)$, f étant une fonction croissante de t . On se pose le problème suivant :

On dispose d'un temps T aléatoire (et inconnu) pour traiter une tâche. On va procéder de la manière suivante : on va tenter d'exécuter un algorithme $A(t_1)$; si on n'est pas interrompu, on tentera d'exécuter un algorithme $A(t_2)$ ($t_2 > t_1$), etc... La valeur d'un algorithme interrompu avant terme est $f(0)$. Au temps T , on disposera du résultat de l'algorithme le plus long que l'on aie pu exécuter entièrement et donc, de sa valeur.

Le but recherché est de maximiser cette valeur *moyenne* obtenue en choisissant judicieusement la suite $(t_i)_{i \geq 1}$ des durées d'algorithmes à exécuter. Une fois fixée ces durées, on disposera en quelque sorte d'un algorithme interruptible composé de plus petits algorithmes non interruptibles.

Dans ce chapitre, on formalise ces notions de *valeurs*, et de temps T aléatoire; on montre, au travers de quelques exemples, les difficultés mathématiques qui apparaissent lorsque l'on cherche à trouver une suite $(t_i)_{i \geq 1}$ qui maximise la valeur moyenne obtenue.

En particulier, lorsque T suit une loi exponentielle de moyenne 1, cette recherche de la suite $(t_i)_{i \geq 1}$ mène à une équation de Bellman et également à l'équation fonctionnelle non standard suivante :

$$z(x) = 1 + x + e^{-z(z(x))} (2 - e^{-z(z(z(x)))}).$$

Pour prouver l'existence (et calculer) d'une suite (t_i) optimisante, on est amené à prouver que cette équation fonctionnelle possède une solution sur tout \mathbb{R}^+ ; pour l'instant, on n'est parvenu à montrer l'existence d'une telle solution que sur $[0.4, +\infty[$.

Références

- [1] A.V. Aho, J.E. Hopcroft, J.D. Ullman, (1983) *Data Structures and Algorithms*, Addison-Wesley .
- [2] D. Aldous, (1991) *The continuum random tree. II : An overview*. Stochastic analysis, Proc. Symp., Durham/UK 1990, Lond. Math. Soc. Lect. Note Ser. 167, 23-70.
- [3] D. Aldous, (1993) *The continuum random tree. III*. Ann. Probab. 21, No.1, 248-289.
- [4] D. Aldous, (1998) *Brownian excursion conditioned on its local time*, Elect. Comm. in Probab., **3** , 79-90.
- [5] P. Biane, M. Yor, (1987) *Valeurs principales associées aux temps locaux browniens*, Bull. Sci. Maths **111**, 23-101.
- [6] P.G. Ciarlet, (1988) *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson.
- [7] T.H. Cormen, E. Leiserson, R.L. Rivest, (1990) *Introduction to Algorithm*, MIT press.
- [8] M. Drmota, B. Gittenberger, (1997) *On the profile of random trees.*, Random Structures Algorithms 10, no. 4, 421-451.
- [9] P. Flajolet, A.M. Odlyzko, (1982) *The average height of binary trees and other simple trees*, J. Comput. Syst. Sci. 25, 171-213.
- [10] D.E. Knuth, (1997) *The art of computer programming. Vol. 3 : sorting and searching*, 2nd ed., Addison-Wesley.
- [11] Th. Jeulin, (1980) *Semi-martingales et grossissement d'une filtration*. Lecture Notes in Mathematics, 833, Springer-Verlag.
- [12] G. Louchard, (1986) *Brownian motion and algorithm complexity*. [J] BIT 26, 17-34.
- [13] A.M. Odlyzko, (1995) *Search for the maximum of a random walk*, Random Structures and Algorithms, Vol. 6, p. 275-295.
- [14] A.M. Odlyzko, H.S. Wilf, (1987) *Bandwidths and profiles of trees*, J. Comb. Theory, Ser. B 42, 348-370.
- [15] R. Sedgewick, P. Flajolet, (1996) *An Introduction to the Analysis of Algorithm*, Addison-Wesley.
- [16] L. Takács, (1993) *Limit distributions for queues and random rooted trees*, J.Ap. Math. Stoch. Ana, **6** , N.3, p.189 - 216.

CHAPITRE I

An average case quasi optimal search algorithm for the zeros of a random walk¹

Jean-François Marckert

Laboratoire de Mathématiques,

Institut Elie Cartan,

BP 239

54 506 Vandoeuvre les Nancy Cedex

Odlyzko [8] proves that on average $c\sqrt{n} + o(\sqrt{n})$ probes are necessary and sufficient to find the maximum of a simple symmetric random walk with n steps (in which $c = \sqrt{8/\pi} \log 2$). In this paper we prove that $\frac{1}{\sqrt{2\pi}}\sqrt{n} \log n + o(\sqrt{n} \log n)$ probes are necessary and sufficient to find all the zeros of a simple symmetric random walk with n steps.

1 Introduction

Consider a length n simple random walk $\omega = (S_k(\omega))_{k \in \{0, \dots, n\}}$ defined by $S_0 = 0$ and $S_k = X_1 + \dots + X_k$, where the X_i are independent random variables such that

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}.$$

We assume the sample path ω unknown. The knowledge of one of the $S_l(\omega)$ for an l of our choice costs 1. The question is: how much do we pay to find each k such that $S_k = 0$?

The cost of an algorithm is naturally identified with the number of probes it uses. Let us describe a fairly natural algorithm for the search of all the zeros, say, the simple jump algorithm, $sj(n)$:

At step 1, we first probe position 2. Then we cover the walk from left to right. Assume that at step T , we probed position m . At step $T+1$:

- If $S_m = 0$ we probe position $m+2$.
- Else we probe position $m + |S_m|$,

and this until $m + |S_m| > n$. Figure 1.1 shows a sample path where we have blackened the points probed by $sj(n)$. Due to the bounded variation of the sample path, $sj(n)$ does not miss any zero.

The cost of the simple-jump algorithm is

Theorem 1.1

$$\mathbb{E}(T_{sj(n)}) \sim \sqrt{\frac{2}{\pi}} \sqrt{n} \log n.$$

Denote by \mathcal{A}_n the finite set of algorithms that determine all indices k such that $S_k(\omega) = 0$. For an algorithm a in \mathcal{A}_n , denote by $T_a(\omega)$ the number of probes used by a to find all the zeros of ω . The mains result of this paper are

¹AMS 1991 *subject classifications*. Primary: 68Q25; Secondary: 68P10, 90B40
Key words and phrases: analysis of algorithms, searching, random walk, average case

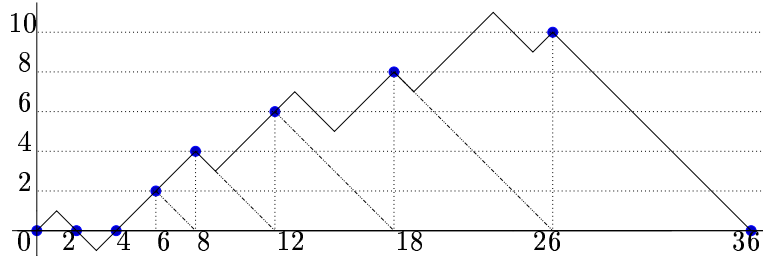


Figure 1.1

Theorem 1.2 For any algorithm a in \mathcal{A}_n , we have:

$$\mathbb{E}(T_a) \geq \frac{1}{\sqrt{2\pi}}\sqrt{n} \log n + o(\sqrt{n} \log n).$$

Theorem 1.3 There exists an algorithm $dj(n)$ in \mathcal{A}_n such that

$$\mathbb{E}(T_{dj(n)}) = \frac{1}{\sqrt{2\pi}}\sqrt{n} \log n + o(\sqrt{n} \log n).$$

This algorithm is described in Section 2. Note that the rather **natural** algorithm $sj(n)$ is actually twice slower.

In this paper, we are essentially interested by the average cost analysis of algorithms that determine all the zeros of ω . There is a general concern for fast search of roots of functions (see Novak & Ritter [7] for a survey). Though the problem of finding the zeros of a random walk does not arise directly in any application, we believe that the fairly complete answer we give in this rather simple setting will help to understand similar search problems. For instance, it is interesting to compare our result to the result obtained in [8]: Odlyzko proved that the search of the maximum M_n needs $c\sqrt{n}$ probes (on the average). Knowing that the mean number of zeros is $\sqrt{2n/\pi}(1 + o(1))$, while the mean number of maxima is $O(1)$, the “best” algorithm needs only, on the average, $O(\log n)$ probes for each zero, but $O(\sqrt{n})$ probes at least for one maximum. From a more probabilistic point of view, the behavior of these algorithms gives informations about the sample paths, particularly about the distribution of their zeros and about their local time near the maximum and at 0.

As regards the worst case complexity, it is nearly the same as in the case of the search of the maximum: let us describe the worst paths, say ω : $S_0(\omega) = 0$, $S_k(\omega) = \pm 1$ for k odd, $S_k(\omega) = 0$ for k even. Clearly, any algorithm needs at least $\lceil n/2 \rceil$ probes to find every zero of ω . Consider now the algorithm which consists in probing each even position k . This algorithm uses $\lceil n/2 \rceil$ probes to find all the zeros of any path. (The simple jump algorithm uses at most $\lceil n/2 \rceil$ probes on any path too.)

For related results, one can see Odlyzko [8] who studied the search of the maximum of random walks with n steps, in the symmetric case. He proved that, on average, $c\sqrt{n} + o(\sqrt{n})$ probes are necessary and sufficient to find the maximum, where $c = \sqrt{8/\pi} \log 2$ (Hwang

[5], Chassaing [2]). In the asymmetric case (where $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = -1) = p \neq \frac{1}{2}$), Chassaing [2] proved that only $\log n / (|\log(p) - \log(1-p)|)$ probes are necessary and sufficient on average.

The paper is organized as follows: Section 2 contains the description of $dj(n)$. Section 3 contains the proofs of Theorems 1.2 and 1.3 (the computation of the mean cost of $sj(n)$ being similar to the one of $dj(n)$, the proof of Theorem 1.1 will be omitted). Section 4 is devoted to the study of the following problem: μ being a fixed real number, how many probes are needed to find out all indices k such that $S_k = \lfloor \mu\sqrt{n} \rfloor$?

2 The “double-jump” algorithm

2.1 A set of moderate variation

We prove the existence of a large subset of sample paths satisfying a property of moderate variation. Following Odlyzko [8], we consider the set $\Omega_\alpha(n)$ of sample paths ω such that for any positive k and m , $m+k \leq n$,

$$|S_{m+k}(\omega) - S_m(\omega)| \leq \alpha\sqrt{k \log n}.$$

We have

Lemma 2.1 (Odlyzko[8]) *For each positive number β there exists a constant c such that for any n*

$$\mathbb{P}(\Omega_c(n)) \geq 1 - n^{-\beta}.$$

For the sequel of the paper, we choose a real number $\beta > \frac{1}{2}$.

Proof: We have

$$\begin{aligned} \mathbb{P}(\exists k, \exists m \setminus |S_{m+k} - S_m| \geq c\sqrt{k \log n}) &\leq \sum_{m=0}^{n-1} \sum_{k=1}^{n-m} \mathbb{P}(|S_{m+k} - S_m| \geq c\sqrt{k \log n}) \\ &\leq 2 \frac{n(n+1)}{2} \exp\left(-\frac{c^2 \log n}{2}\right) \\ &\leq 2n^{2-\frac{c^2}{2}}, \end{aligned} \tag{2.1}$$

where the second inequality follows from Chernov bounds (Bollobás [1] p.12.): for any $\alpha \geq 0$ and $k \in \mathbb{N}$,

$$\mathbb{P}(|S_k| > \alpha) \leq 2 \exp\left(-\frac{\alpha^2}{2k}\right). \tag{2.2}$$

Then we choose $c = \sqrt{2(\beta + 3)}$ in (2.1). \square

In the sequel, we will often note Ω_c instead of $\Omega_c(n)$.

2.2 Description of the “double-jump” algorithm

Let us describe the algorithm $dj(n)$ mentioned in the introduction. At step 1, we probe position 2. Then we cover the sample path from left to right. Assume that at step T we probed position m . Two cases arise:

- a) $|S_m| \leq \log^{1+\varepsilon} n$
- b) $|S_m| > \log^{1+\varepsilon} n$ (ε being any fixed real number in the interval $]0,1[$).

At step $T+1$:

In Case a), we do a “simple jump”: we probe position $m+2$ or $m+|S_m|$ depending on whether $S_m = 0$ or not.

In Case b), we do a “double jump”: we probe position $m+K$ where

$$K = \lfloor 2|S_m| - c\sqrt{2 \log n |S_m|} \rfloor, \quad (2.3)$$

and we compute

$$B_m = \frac{S_m + S_{m+K} - K}{2}.$$

Without loss of generality, assume that $S_m > 0$.

If $B_m > 0$, ω does not cancel on the interval $[m, m+K]$, since the minimum of the sample path on the interval $[m, m+K]$ is not smaller than B_m (see Figure 2.1). We have found every zero at the left of $m+K$, and the algorithm is ready for the next step.

On the other hand, if $B_m \leq 0$, ω is in $\mathfrak{C}\Omega_c$ (the complement of Ω_c) and may cancel. Then, we use any algorithm that finds the zeros, for instance we probe every position starting from m . These probes contribute for at most $O(n^{-\beta+1})$ to $\mathbb{E}(T_{dj(n)})$ (due to Lemma 2.1). Figure 2.2 explains the choice of K in (2.3): if ω is in Ω_c , the sample path remains inside the parabola and, in particular, the probed point $(m+K, S_{m+K})$. Note that for small values of

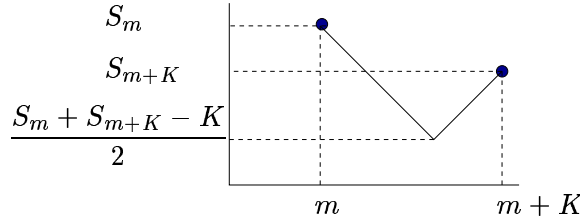


Figure 2.1

$|S_m|$, we might have $K \leq |S_m|$ or even K negative; a double-jump is an improvement on the simple-jump only when $|S_m| \gg \log n$.

3 Quasi-optimality of $dj(n)$

3.1 Sketch of proof of Theorems 1.2 and 1.3

- There exists a lower bound for the mean cost of any algorithm.

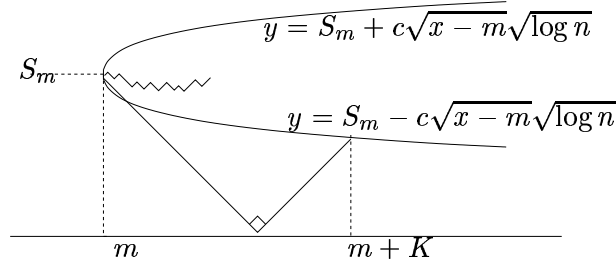


Figure 2.2

We assume that an algorithm a has succeeded after exactly T_a probes and has probed the locations $m_1 < m_2 < \dots < m_{T_a}$. This order is not necessarily the chronologic order of the probes (algorithms don't necessarily probe from left to right). We set $k_i = m_{i+1} - m_i$ for $i \in \{1, \dots, T_a - 1\}$ and $k_0 = m_1$. As in Odlyzko [8], the main idea is to introduce the concept of density of probed points. We define a density function Wt_a associated with a by:

$$Wt_a(j) = \frac{1}{k_i}$$

for any j in $]m_i, m_i + k_i]$. The “weight” $WT_a(\omega)$ of a sample path ω associated with a is defined by

$$WT_a(\omega) = \sum_{j=1}^n Wt_a(j).$$

Since $\sum_{m_i+1}^{m_{i+1}} 1/k_i = 1$, $WT_a(\omega)$ is exactly the number of probes used by a on ω , thus:

$$WT_a(\omega) = T_a(\omega) \quad \text{and} \quad \mathbb{E}(WT_a) = \mathbb{E}(T_a). \quad (3.4)$$

In the description of $d_j(n)$ (Section 2) we have seen that $S_{m_i} + S_{m_i+k_i} - k_i$ has to be positive. This is also true for any algorithm a , otherwise some zeros may be missed. This explains why one can compare $1/k_i$ with $1/(S_{m_i} + S_{m_i+k_i})$ and thus $\mathbb{E}(WT_a)$ (using the moderate variation of the paths) with U_n , where U_n is the r.v. defined by:

$$U_n = \frac{1}{2} \mathbb{E} \left(\sum_{k=1}^n \frac{1}{|S_k| + 1} \right).$$

Theorems 1.2 and 1.3 appear to be the corollaries of the three following theorems (proved in the three following subsections).

Theorem 3.1

$$U_n \sim \frac{1}{\sqrt{2\pi}} \sqrt{n} \log n.$$

U_n is a “quasi lower bound” to the cost of every algorithm:

Theorem 3.2 For any a in \mathcal{A}_n ,

$$\mathbb{E}(T_a) \geq (1 + o(1))U_n. \quad (3.5)$$

The proof of Theorem 3.1 (resp. 3.2) is given in subsection 3.2 (resp. 3.3).

Now, the second point explains why $dj(n)$ is quasi-optimal in the average case.

- The cost of $dj(n)$ is equivalent to U_n :

Theorem 3.3

$$\mathbb{E}(T_{dj(n)}) = \mathbb{E}(WT_{dj(n)}) \sim U_n.$$

To prove this theorem, we use again the sequence k_i (k_i being in this case the length of the i^{th} jump done with $dj(n)$), and define

$$L_1 = \{i | k_i > \log^{1+\varepsilon} n\} \quad (3.6)$$

and

$$L_2 = \{i | k_i \leq \log^{1+\varepsilon} n\}. \quad (3.7)$$

$\#L_1$ (resp. $\#L_2$) is the number of double (resp. simple) jumps, if n is large enough. $\#L_1$ and $\#L_2$ are related with $T_{dj(n)}$ by:

$$\mathbb{E}(WT_{dj(n)}) = \mathbb{E}(T_{dj(n)}) = \mathbb{E}(\#L_1 + \#L_2).$$

If i is in L_1 , k_i is a double jump defined at (2.3). One has $k_i \sim 2S_{m_i}$ and using the moderate variation properties of the paths of Ω_c , we show that

$$\mathbb{E}(\#L_1) = \mathbb{E}\left(\sum_{i \in L_1} \sum_{j=m_i+1}^{m_{i+1}} \frac{1}{k_i}\right) \simeq \mathbb{E}\left(\sum_{i \in L_1} \sum_{j=m_i+1}^{m_{i+1}} \frac{1}{2S_j}\right) \sim U_n.$$

In a second step, we prove that $\mathbb{E}(\#L_2)$ is negligible (cf. Lemma 3.13).

The proof of Theorem 3.3 is the object of subsection 3.4.

3.2 Proof of Theorem 3.1

Theorem 3.1 is a corollary of the two following lemmas. Let α be a fixed real number in $]0,1[$. We decompose U_n into two parts.

$$\frac{1}{2}\mathbb{E}\left(\sum_{k=1}^n \frac{1}{|S_k|+1}\right) = \sum_{j=0}^n \sum_{k=1}^n \frac{\mathbb{P}(|S_k|=j)}{2(j+1)} = A(n, \alpha) + B(n, \alpha) \quad (3.8)$$

where

$$A(n, \alpha) = \sum_{j=[\sqrt{n}\log^\alpha n]+1}^n \sum_{k=1}^n \frac{\mathbb{P}(|S_k|=j)}{2(j+1)} \quad \text{and} \quad B(n, \alpha) = \sum_{j=0}^{[\sqrt{n}\log^\alpha n]} \sum_{k=1}^n \frac{\mathbb{P}(|S_k|=j)}{2(j+1)}.$$

Lemma 3.4

$$A(n, \alpha) \leq \frac{\sqrt{n}}{2 \log^\alpha n}.$$

Proof:

$$\begin{aligned} A(n, \alpha) &\leq \frac{1}{2\sqrt{n} \log^\alpha n} \sum_{k=1}^n \sum_{j=[\sqrt{n} \log^\alpha n]+1}^n \mathbb{P}(|S_k| = j) \\ &\leq \frac{1}{2\sqrt{n} \log^\alpha n} \sum_{k=1}^n \mathbb{P}(|S_k| \geq [\sqrt{n} \log^\alpha n]). \quad \square \end{aligned}$$

Lemma 3.5

$$B(n, \alpha) \sim \sqrt{\frac{n}{2\pi}} \log n.$$

Proof: As in Spitzer [10], we define

$$G_n(0, j) = \sum_{k=1}^n (\mathbb{P}(S_k = j) - \mathbb{P}(S_k = 0))$$

$G_n(0, j)$ satisfies:

$$|G_n(0, j)| \leq |j|. \quad (3.9)$$

This result follows a standard formula in potential theory:

$$\mathbb{P}(S_k = j) = \frac{1}{2\pi} \int_0^{2\pi} \cos^k(\theta) \cos(j\theta) d\theta.$$

Thus,

$$\begin{aligned} |G_n(0, j)| &\leq \frac{1}{2\pi} \int_0^{2\pi} (1 - \cos(j\theta)) \frac{1 - \cos^{n+1} \theta}{1 - \cos \theta} d\theta \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} \frac{1 - \cos(j\theta)}{1 - \cos \theta} d\theta \\ &\leq |j|. \end{aligned}$$

We decompose $B(n, \alpha)$ into two parts:

$$B(n, \alpha) = f(n, \alpha) + g(n, \alpha)$$

where

$$f(n, \alpha) = \left(\sum_{k=1}^n \mathbb{P}(S_k = 0) \right) \left(\sum_{j=1}^{[\sqrt{n} \log^\alpha n]} \frac{1}{j+1} \right)$$

and

$$g(n, \alpha) = \sum_{j=1}^{\lfloor \sqrt{n} \log^\alpha n \rfloor} \frac{G_n(0, j)}{j+1}.$$

Since $0 < \alpha < 1$, we have immediately

$$f(n, \alpha) \sim \sqrt{\frac{n}{2\pi}} \log n,$$

and using (3.9), we obtain

$$|g(n, \alpha)| \leq \sum_{j=1}^{\lfloor \sqrt{n} \log^\alpha n \rfloor} \frac{j}{j+1} \leq \sqrt{n} \log^\alpha n. \quad \square$$

3.3 Proof of Theorem 3.2

We assume that algorithm a has succeeded after exactly T_a probes. The set $\mathcal{P}r_a(\omega) = \{(m_1, S_{m_1}(\omega)), \dots, (m_{T_a}, S_{m_{T_a}}(\omega))\}$ of probed points makes up a proof of all the zeros of ω , that is, it locates uniquely all the zeros. We will assume that $m_0 = 0 < m_1 < \dots < m_{T_a}$. Recall that we do not assume that the algorithm a has probed these positions in this order, but that they have been ordered afterwards (if needed).

Denote by I_i the interval $]m_i, m_{i+1}]$, for $0 \leq i \leq T_a - 1$. Assume that all the zeros of ω are in $\mathcal{P}r_a(\omega)$; this condition will be removed later. Then, we know that the sample path has a constant sign between two probes. Without loss of generality, assume that ω is positive on $]m_i, m_{i+1}]$; on $]m_i, m_{i+1}]$, only four possible situations arise. See figures 3.1.

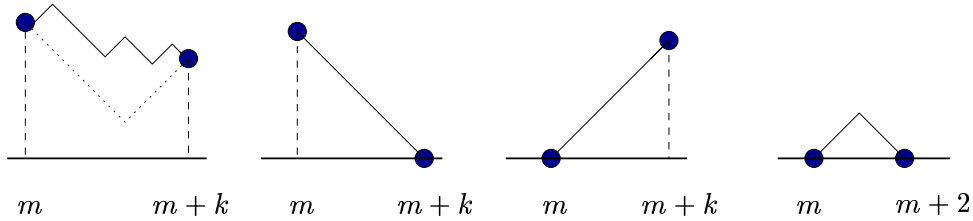


Figure 3.1

In these four cases, we have for any i

$$k_i - |S_{m_i}| - |S_{m_i+k_i}| \leq 2. \quad (3.10)$$

In order to bound the cardinality of $\mathcal{P}r_a$ in the three following lemmas, the following identity will be useful:

$$\sum_{i \in D} \sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} = \#D.$$

For ω in Ω_c , we define

$$\begin{aligned} A(\omega) &= \{i | k_i \geq \log^{1+\varepsilon} n\} \\ B(\omega) &= \{i | k_i \leq \log^{1+\varepsilon} n \text{ and } k_i - |S_{m_i}| - |S_{m_i+k_i}| + \log^{1+\varepsilon} n \leq 0\} \\ C(\omega) &= \{k_i \leq \log^{1+\varepsilon} n \text{ and } k_i - |S_{m_i}| - |S_{m_i+k_i}| + \log^{1+\varepsilon} n \geq 0\} \end{aligned}$$

Lemma 3.6 For n large enough,

$$\sum_{i \in A} \sum_{h=m_i+1}^{m_i+k_i} \frac{1}{2(|S_h|+1)} \leq (1 + 4c \log^{-\varepsilon/2} n) \#A$$

Proof: for $i \in A$ and $h \in]m_i, m_{i+1}]$,

$$\begin{aligned} \frac{1}{2(|S_h|+1)} &\leq \frac{1}{k_i + |S_h| - |S_{m_i}| + |S_h| - |S_{m_i+k_i}|} \\ &\leq \frac{1}{k_i} \frac{1}{1 - \frac{2c \log^{1/2} n}{k_i^{1/2}}} \\ &\leq \frac{1}{k_i} (1 + 4c \log^{-\varepsilon/2} n) \end{aligned}$$

the last inequality is true for $\omega \in \Omega_c$ and for $n \geq \exp((4c)^{-\varepsilon/2})$. \square

Lemma 3.7 For n large enough,

$$\sum_{i \in B} \sum_{h=m_i+1}^{m_i+k_i} \frac{1}{2(|S_h|+1)} \leq \#B.$$

Proof: for $i \in B$ and $h \in]m_i, m_{i+1}]$,

$$\frac{1}{2(|S_h|+1)} \leq \frac{1}{k_i} \frac{1}{1 - \frac{2c \log^{1/2} n}{\sqrt{k_i}} + \frac{\log^{1+\varepsilon} n}{k_i}} \leq \frac{1}{k_i},$$

if $n \geq \exp((2c)^{-\varepsilon/2})$. \square

Lemma 3.8

$$\mathbb{E} \left(\sum_{i \in C} \sum_{h=m_i+1}^{m_i+k_i} \frac{1}{2(|S_h|+1)} \right) = O(\sqrt{n} \log \log n).$$

Proof: for $i \in C$ and $h \in]m_i, m_{i+1}]$,

$$|S_h| \leq |S_{m_i}| + k_i \text{ and } |S_h| \leq |S_{m_i+k_i}| + k_i$$

So, $|S_h| \leq \frac{1}{2}(|S_{m_i}| + |S_{m_i+k_i}| + 2k_i) \leq 3k_i + \log^{1+\varepsilon} n \leq 2 \log^{1+\varepsilon} n$. Now,

$$\mathbb{E}\left(\sum_{i \in C} \sum_{h=m_i+1}^{m_i+k_i} \frac{1}{2(|S_h|+1)}\right) \leq \mathbb{E}\left(\sum_{h=1}^n \frac{\mathbb{I}_{\{h \mid |S_h| \leq 2 \log^{1+\varepsilon} n\}}}{2(|S_h|+1)}\right)$$

To end the proof of Lemma 3.8, we use:

$$\frac{1}{2} \mathbb{E}\left(\sum_{h=1}^n \frac{\mathbb{I}_{\{h \mid |S_h| \leq 2 \log^{1+\varepsilon} n\}}}{|S_h|+1}\right) = O(\sqrt{n} \log \log n). \quad (3.11)$$

To prove (3.11), we use that for any i , $C_h^{\lfloor \frac{h}{2} \rfloor} \geq C_h^i$. Thus

$$\mathbb{E}\left(\frac{\mathbb{I}_{\{h \mid |S_h| \leq \log^{1+\varepsilon} n\}}}{|S_h|+1}\right) = \sum_{i=0, i \equiv h[2]}^{\lfloor \log^{1+\varepsilon} n \rfloor} \frac{2C_h^{\frac{h}{2} + \frac{i}{2}} 2^{-h}}{i+1} \leq \frac{C_4}{\sqrt{h}} \log \log n. \quad \square$$

Proof of Theorem 3.2: Using Lemmas 3.6, 3.7, 3.8, we obtain:

$$\begin{aligned} U_n &= \mathbb{E}\left(\sum_{h=1}^n \frac{\mathbb{I}_{\omega \in \Omega_c}}{2(|S_h|+1)}\right) + \mathbb{E}\left(\sum_{h=1}^n \frac{\mathbb{I}_{\omega \notin \Omega_c}}{2(|S_h|+1)}\right) \\ &= \mathbb{E}\left(\sum_{i \in A \cup B \cup C} \sum_{h=m_i+1}^{m_i+k_i} \frac{1}{2(|S_h|+1)}\right) + \mathbb{E}\left(\sum_{h=1}^n \frac{\mathbb{I}_{\omega \notin \Omega_c}}{2(|S_h|+1)}\right) \\ &\leq (1 + 4c \log^{-\varepsilon/2} n) \mathbb{E}(\#A + \#B) + O(\sqrt{n} \log \log n) + O(n^{-\beta+1}). \end{aligned}$$

Using Theorem 3.1, we deduce that

$$\mathbb{E}(T_a) = \mathbb{E}(\#\mathcal{P}r_a) = \mathbb{E}(\#A + \#B + \#C) + \mathbb{E}(\#\mathcal{P}r_a \mathbb{I}_{\complement \Omega_c}) \geq \mathbb{E}(\#A + \#B) \geq (1 + o(1))U_n. \quad \square$$

Remarks

We have shown that for an algorithm a , if the set $\mathcal{P}r_a$ contains the zeros, then it satisfies $\mathbb{E}(\#\mathcal{P}r_a) \geq (1 + o(1))U_n$. What about $\mathbb{E}(\#\mathcal{P}r_a)$ if $\mathcal{P}r_a$ does not contain all the zeros? The set $\mathcal{P}r_a$ make up a proof of all the zeros. Denote by $\mathcal{P}r_a^+$ the union of $\mathcal{P}r_a$ and the set of zeros. Since the average number of zeros of a length n symmetric random walk is $\sim \sqrt{2n/\pi}$, we have

$$\mathbb{E}(\#\mathcal{P}r_a^+) = \mathbb{E}(\#\mathcal{P}r_a) + O(\sqrt{n}).$$

$\mathcal{P}r_a^+$ locates and contains the zeros. So,

$$\mathbb{E}(\#\mathcal{P}r_a^+) \geq (1 + o(1))U_n,$$

and thus

$$\mathbb{E}(\#\mathcal{P}r_a) \geq (1 + o(1))U_n,$$

too.

3.4 Proof of Theorem 3.3

Write

$$\mathbb{E}(WT_{dj(n)}(\omega)) = A_1(n) + A_2(n) + A_3(n)$$

where

$$\begin{aligned} A_1(n) &= \mathbb{E}\left(\sum_{i \in L_1} \sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} \mathbb{I}_{\Omega_c}(\omega)\right), \\ A_2(n) &= \mathbb{E}\left(\sum_{i \in L_2} \sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} \mathbb{I}_{\Omega_c}(\omega)\right), \\ A_3(n) &= \mathbb{E}\left(\sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} \mathbb{I}_{\mathbb{C}\Omega_c}(\omega)\right) \end{aligned}$$

and L_1 and L_2 are defined in (3.6) and (3.7).

Proposition 3.9

$$A_1(n) = (1 + o(1))U_n.$$

Proof: Proposition 3.9 follows immediately from Lemmas 3.10 and 3.11 and Theorem 3.1. \square

Lemma 3.10 *There exists a constant C_0 such that for any n large enough,*

$$\mathbb{E}\left(\left|\sum_{i \in L_1, \omega \in \Omega_c} \sum_{h=m_i+1}^{m_i+k_i} \left(\frac{1}{k_i} - \frac{1}{2(|S_h|+1)}\right)\right|\right) \leq C_0 \log^{-\varepsilon/2} n \sqrt{n} \log n.$$

Proof: Consider i an element of L_1 ; for ω in Ω_c , we have :

$$k_i = 2|S_{m_i}| - c\sqrt{2|S_{m_i}| \log n} = 2|S_{m_i}|(1 - \gamma(i))$$

where

$$0 < \gamma(i) = c\sqrt{\frac{\log n}{2|S_{m_i}|}} < \frac{c}{\sqrt{2}} \log^{-\varepsilon/2} n.$$

S_{m_i} and $S_{m_i+k_i}$ have the same sign; for $h \in]m_i + 1, m_{i+1}]$, we have

$$\frac{1}{k_i} - \frac{1}{2(|S_h|+1)} = \frac{\eta_{i,h}}{2(|S_h|+1)} \tag{3.12}$$

and

$$|\eta_{i,h}| = \left| \frac{2|S_h| - 2|S_{m_i}| + 2 + 2|S_{m_i}|\gamma(i)}{k_i} \right| \leq \frac{2c\sqrt{k_i \log n}}{k_i} + \frac{2}{k_i} + \frac{\gamma(i)}{1 - \gamma(i)}.$$

Using that $\frac{x}{1-x}$ is bounded on $[0, \frac{c}{\sqrt{2}} \log^{-\varepsilon/2} n]$ by $C_1 \log^{-\varepsilon/2} n$ for n large enough, we obtain,

$$|\eta_{i,h}| \leq C_2 \log^{-\varepsilon/2} n,$$

and thus

$$\left| \sum_{i \in L_1} \sum_{h=m_i+1}^{m_i+k_i} \left(\frac{1}{k_i} - \frac{1}{2(|S_h|+1)} \right) \mathbb{I}_{\Omega_c} \right| \leq C_2 \log^{-\varepsilon/2} n \sum_{i \in L_1} \sum_{h=m_i+1}^{m_i+k_i} \frac{\mathbb{I}_{\Omega_c}}{2(|S_h|+1)}$$

and we conclude using Theorem 3.1. \square

Lemma 3.11

$$\mathbb{E} \left(\sum_{i \in L_1} \sum_{h=m_i+1}^{m_i+k_i} \frac{\mathbb{I}_{\Omega_c}}{2(|S_h|+1)} \right) \sim U_n.$$

Proof: Since $\{h \mid h \in]m_i, m_{i+1}], i \in L_2\} \subset \{h \mid |S_h| < 2 \log^{1+\varepsilon} n\}$, we have:

$$\begin{aligned} \left| \mathbb{E} \left(\sum_{h=1}^n \frac{1}{|S_h|+1} - \sum_{i \in L_1} \sum_{h=m_i+1}^{m_i+k_i} \frac{\mathbb{I}_{\Omega_c}}{|S_h|+1} \right) \right| &= \left| \mathbb{E} \left(\sum_{i \in L_2} \sum_{h=m_i+1}^{m_i+k_i} \frac{\mathbb{I}_{\Omega_c}}{|S_h|+1} + \sum_{h=1}^n \frac{\mathbb{I}_{\mathfrak{C}\Omega_c}}{|S_h|+1} \right) \right| \\ &\leq \mathbb{E} \left(\sum_{h=1}^n \frac{\mathbb{I}_{\{|S_h| \leq (2 \log n)^{1+\varepsilon}\}}}{|S_h|+1} \right) + \mathbb{E} \left(\sum_{h=1}^n \frac{\mathbb{I}_{\mathfrak{C}\Omega_c}}{|S_h|+1} \right). \end{aligned}$$

Moreover, we have

$$\mathbb{E} \left(\sum_{h=1}^n \frac{\mathbb{I}_{\mathfrak{C}\Omega_c}}{2(|S_h|+1)} \right) \leq n \mathbb{P}(\omega \in \mathfrak{C}\Omega_c) \leq n^{-\beta+1},$$

and using (3.11), we obtain that

$$\mathbb{E} \left(\sum_{i \in L_1} \sum_{h=m_i+1}^{m_i+k_i} \frac{\mathbb{I}_{\Omega_c}}{2(|S_h|+1)} \right) = \mathbb{E} \left(\sum_{h=1}^n \frac{1}{2(|S_h|+1)} \right) + O(\sqrt{n} \log \log n) + O(n^{-\beta+1}). \quad \square$$

Proposition 3.12

$$A_2(n) = O(\sqrt{n} \log^{\frac{1+\varepsilon}{2}} n).$$

Proof: Since $\sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} = 1$, we have $\mathbb{E}(A_2) \leq \mathbb{E}(\#L_2)$. We will prove that

$$\mathbb{E}(\#L_2) \leq 6\sqrt{\frac{2n}{\pi}} \log^{\frac{1+\varepsilon}{2}} n. \quad (3.13)$$

Proof: Assume that i is in L_2 and $m_i < h \leq m_{i+1}$. We have $|S_{m_i}| \leq \log^{1+\varepsilon} n$. So, $|S_h| \leq |S_{m_i}| + k_i \leq 2 \log^{1+\varepsilon} n$.

$$\begin{aligned} \mathbb{E}(\#\{h|h \in [m_i, m_{i+1}], i \in L_2\}) &\leq \mathbb{E}(\#\{h/|S_h| \leq 2 \log^{1+\varepsilon} n\}) \\ &\leq 4(\log^{1+\varepsilon} n) \mathbb{E}(\#\{h/S_h = 0\}) \\ &\leq 4\sqrt{\frac{2n}{\pi}} \log^{1+\varepsilon} n \end{aligned}$$

So, $\mathbb{E}(\#L_2) = \mathbb{E}(\#\{i \in L_2 | k_i \leq \log^{\frac{1+\varepsilon}{2}} n\}) + \mathbb{E}(\#\{i \in L_2 | \log^{\frac{1+\varepsilon}{2}} n \leq k_i \leq \log^{1+\varepsilon} n\})$.

If $\log^{\frac{1+\varepsilon}{2}} n \leq k_i \leq \log^{1+\varepsilon} n$, any l in $[m_i, m_i + \log^{\frac{1+\varepsilon}{2}} n]$ is also in $[m_i, m_{i+1}]$. It follows that with each element i of $\{i \in L_2 | \log^{\frac{1+\varepsilon}{2}} n \leq k_i \leq \log^{1+\varepsilon} n\}$ we can associate at least $\log^{\frac{1+\varepsilon}{2}} n$ elements of $\{h/|S_h| \leq 2 \log^{1+\varepsilon} n\}$. So

$$\begin{aligned} \mathbb{E}(\#\{i \in L_2 | \log^{\frac{1+\varepsilon}{2}} n \leq k_i \leq \log^{1+\varepsilon} n\}) &\leq 4\sqrt{\frac{2n}{\pi}} \log^{1+\varepsilon} n \times \log^{-\frac{1+\varepsilon}{2}} n \\ &\leq 4\sqrt{\frac{2n}{\pi}} \log^{\frac{1+\varepsilon}{2}} n \end{aligned} \quad (3.14)$$

On the other hand

$$\mathbb{E}(\#\{i \in L_2 | k_i < \log^{\frac{1+\varepsilon}{2}} n\}) \leq \mathbb{E}(\#\{m/|S_m| < \log^{\frac{1+\varepsilon}{2}} n\}) \leq 2\sqrt{\frac{2n}{\pi}} \log^{\frac{1+\varepsilon}{2}} n \quad (3.15)$$

Using (3.14) and (3.15), relation (3.13) is proved. \square

Lemma 3.13

$$A_3(n) = o(\sqrt{n}).$$

Proof: We have $A_3(n) = O(n^{-\beta+1})$ and $\beta > \frac{1}{2}$. \square

Using Propositions 3.9, 3.12 and Lemma 3.13, we obtain Theorem 3.3. \square

4 Other case

A natural extension of the problem of finding the zeros is the problem of finding the indices m such that $S_m = \alpha$. One can see that for α constant, the cost of the best algorithms searching α are equivalent (when n goes to $+\infty$) to the cost of the best algorithms searching the zeros. The behavior of the cost of the “best algorithms” really changes when α is of the order of \sqrt{n} . We study here the search of all the indices m such that $S_m = \alpha_n = \lfloor \mu\sqrt{n} \rfloor$, μ being a fixed positive real number. Following the notations given in the introduction, we denote by $\mathcal{A}_n^{\alpha_n}$ the set of algorithms doing this job. We have

Theorem 4.1

$$\min_{a \in A_n^{\alpha_n}} \mathbb{E}(T_a) \sim C_\mu \sqrt{n} \log n$$

in which

$$C_\mu = \frac{1}{4} \int_0^1 \sqrt{\frac{2}{\pi x}} \exp\left(-\frac{\mu^2}{2x}\right) dx = \frac{\mu}{4\sqrt{\pi}} \Gamma\left(-\frac{1}{2}, \frac{\mu^2}{2}\right) \quad \text{if } \mu \neq 0$$

where $\Gamma(x, a) = \int_a^{+\infty} t^{x-1} \exp(-t) dt$ is the incomplete Euler gamma function.

The quasi optimal algorithm $dj^{\alpha_n}(n)$ is an obvious generalization of the double jump algorithm $dj(n)$. (If we take $\mu = 0$, we obtain Theorem 1.2)

$dj^{\alpha_n}(n)$ algorithm: We consider that at time 0, we have probed 0. Assume that at time T , we have probed m and that we know S_m . Two cases can occur:

- 1) $|S_m - \alpha_n| \leq \log^{1+\varepsilon} n$
- 2) $|S_m - \alpha_n| > \log^{1+\varepsilon} n$

In the first case, at step $T+1$, we probe S_{m+2} or $S_{m+|\alpha_n - S_m|}$ depending on $S_m = \alpha_n$ or not. In the second case, at step $T+1$, we probe S_{m+K} where $K = \lfloor 2|\alpha_n - S_m| - c\sqrt{2}|\alpha_n - S_m| \log n \rfloor$ and for the same reason as in the Section 2, if $|S_{m+K} - S_m| > c\sqrt{K} \log n$ we probe every position of the sample path. Note that for $dj^{\alpha_n}(n)$, the first jump is a double jump. Let us compute the cost of this algorithm.

Proof of Theorem 4.1. Applying (3.11), Propositions 3.9, 3.12 and Lemmas 3.10, 3.11, 3.13 and relation (3.13), we obtain

$$\mathbb{E}(T_{dj^{\alpha_n}(n)}) = (1 + o(1))U_n^{(\alpha_n)}$$

in which

$$U_n^{(\alpha_n)} = \mathbb{E}\left(\sum_{m=1}^n \frac{1}{2(|\alpha_n - S_m| + 1)}\right) = \frac{1}{2} \sum_{m=1}^n \sum_{i=-n}^n \frac{\mathbb{P}(S_m = i)}{|\alpha_n - i| + 1}$$

Lemma 4.2 *Let $0 \leq \eta < 1$, if n is sufficiently large, $\forall \gamma > 0, \forall k \geq n^{5/6}$ $\forall i \in [(1 - \eta)\alpha_n, (1 + \eta)\alpha_n], k \equiv i[2]$, we have*

$$(1 - \gamma) \sqrt{\frac{2}{\pi k}} \exp\left(-\frac{i^2}{2k}\right) \leq \mathbb{P}(S_k = i) \leq (1 + \gamma) \sqrt{\frac{2}{\pi k}} \exp\left(-\frac{i^2}{2k}\right)$$

Proof: see Feller [3] p. 83-84 for instance.

Proof of Theorem 4.1: We have to compute $U_n^{(\alpha_n)}$. We have immediately the two following equations:

$$\begin{aligned} \frac{1}{2} \sum_{m=1}^n \sum_{i=(1+\eta)\alpha_n}^n \frac{\mathbb{P}(S_m = i)}{|\alpha_n - i| + 1} &\leq \frac{1}{2\mu\eta} \sqrt{n} \\ \frac{1}{2} \sum_{m=1}^n \sum_{i=-n}^{(1-\eta)\alpha_n} \frac{\mathbb{P}(S_m = i)}{|\alpha_n - i| + 1} &\leq \frac{1}{2\mu\eta} \sqrt{n}. \end{aligned}$$

Due to monotonicity of $\mathbb{P}(S_k = i)$ and by Lemma 4.2, we have:

$$\sum_{k=1}^{n^{5/6}} \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{\mathbb{P}(S_k = i)}{|\alpha_n - i| + 1} \leq (1 + \gamma) \sqrt{\frac{2}{\pi}} \frac{n^{5/12} \log n}{\eta \mu}, \quad (4.16)$$

because

$$\sum_{k=1}^{n^{5/6}} \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{\mathbb{P}(S_k = i)}{|\alpha_n - i| + 1} \leq \sum_{k=1}^{n^{5/6}} \mathbb{P}(S_{n^{5/6}} = [(1-\eta)\alpha_n]) \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{1}{|\alpha_n - i| + 1}.$$

The remaining term $F_{\alpha_n, \eta}(n) = \sum_{k=n^{5/6}}^n \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{\mathbb{P}(S_k = i)}{2(|\alpha_n - i| + 1)}$ satisfies

Proposition 4.3 For $0 \leq \eta < 1$, we have

$$\limsup_{n \rightarrow +\infty} \frac{F_{\alpha_n, \eta}(n)}{\sqrt{n} \log n} \leq (1 + \gamma) C_{\mu(1-\eta)} \quad \text{and} \quad \liminf_{n \rightarrow +\infty} \frac{F_{\alpha_n, \eta}(n)}{\sqrt{n} \log n} \geq (1 - \gamma) C_{\mu(1+\eta)}.$$

Proof: The proof of this two inequations are analogous. Using Lemma 4.2, we have

$$\begin{aligned} F_{\alpha_n, \eta}(n) &\leq (1 + \gamma) \sum_{k=n^{5/6}}^n \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{\exp(-\frac{i^2}{2k}) \mathbb{I}_{\{i \equiv k[2]\}}}{\sqrt{2\pi k} (|\alpha_n - i| + 1)} \\ &\leq (1 + \gamma) \sum_{k=n^{5/6}}^n \sum_{i=(1-\eta)\alpha_n}^{(1+\eta)\alpha_n} \frac{\exp(-\frac{(1-\eta)^2 \alpha_n^2}{2k}) \mathbb{I}_{\{i \equiv k[2]\}}}{\sqrt{2\pi k} (|\alpha_n - i| + 1)} \\ &\leq (1 + \gamma) \frac{1}{4} (\log n + C_1) \sqrt{n} \sum_{k=n^{5/6}}^n \sqrt{\frac{2}{\pi \frac{k}{n}}} \exp(-\frac{(1-\eta)^2 \mu^2 n}{2} \frac{n}{k}) \end{aligned}$$

So by Riemann integration,

$$\limsup_{n \rightarrow +\infty} \frac{F_{\alpha_n, \eta}(n)}{\sqrt{n} \log n} \leq (1 + \gamma) C_{\mu(1-\eta)}. \quad \square$$

To end the proof of Theorem 4.1, we note that Proposition 4.3 is true for any μ and η and that $a \mapsto C_a$ is continuous. \square

The optimality of these algorithms follows from the results of Section 3.3. With similar notations, we found

$$\begin{aligned} \mathbb{E} \left(\sum_{h=1}^n \frac{1}{2(|S_h - \alpha_n| + 1)} \right) &\leq \mathbb{E} \left((1 + 4 \log^{-\varepsilon/2} n) (\#A + \#B) \right. \\ &\quad \left. + \sum_{j \in C} \sum_{h \in I_j} \frac{\mathbb{I}_{|S_h - \alpha_n| \leq \log n \frac{1+\varepsilon}{2}}}{2(|S_h - \alpha_n| + 1)} \right) + O(n^{-\beta+1}) \end{aligned}$$

so $\mathbb{E}(T_a) \geq (1 + o(1)) U_n^{(\alpha_n)}$ for every algorithm a of $\mathcal{A}_n^{(\alpha_n)}$ \square

Remark :

Set $\Lambda_{\alpha_n} = \inf\{k | S_k = \alpha_n\}$, the first hitting time of α_n by ω .

$$\mathbb{E}(T_{dj^{\alpha_n}(n)}) = V_1(\alpha_n) + V_2(\alpha_n)$$

where

$$V_1(\alpha_n) = \mathbb{E}\left(\sum_{j=1}^{\Lambda_{\alpha_n} \wedge n} W t_{dj^{\alpha_n}(n)}(j)\right) \quad \text{and} \quad V_2(\alpha_n) = \mathbb{E}\left(\sum_{j=\Lambda_{\alpha_n} \wedge n+1}^n W t_{dj^{\alpha_n}(n)}(j)\right).$$

$V_1(\alpha_n)$ is the number of probes used by $dj^{\alpha_n}(n)$ to find the first index m such that $S_m = \alpha_n$ (the first passage in α_n); one may interpret $V_2(\alpha_n)$ as the number of probes required to find the zeros of a length $(n - \Lambda_{\alpha_n} \wedge n)$ random walk $\tilde{\omega}$, such that $\tilde{S}_0(\omega) = 0$.

Proposition 4.4

$$V_2(\alpha_n) \sim \mathbb{E}(T_{dj^{\alpha_n}(n)}).$$

and thus,

$$V_1(\alpha_n) = o(\mathbb{E}(T_{dj^{\alpha_n}(n)})).$$

That is, the number of probes required to find the first solution is negligible.

Proof: Using Theorem (1.3), we obtain

$$\begin{aligned} \mathbb{E}\left(\sum_{j=\Lambda_{\alpha_n} \wedge n+1}^n W t_{dj^{\alpha_n}(n)}(j)\right) &\sim \mathbb{E}\left(\frac{1}{\sqrt{2\pi}} \sqrt{n - \Lambda_{\alpha_n} \wedge n} \log(n - \Lambda_{\alpha_n} \wedge n)\right) \\ &= \sqrt{\frac{n}{2\pi}} \mathbb{E}\left(\sqrt{\frac{n - \Lambda_{\alpha_n} \wedge n}{n}} \log\left(\frac{n - \Lambda_{\alpha_n} \wedge n}{n}\right)\right) \\ &+ \sqrt{\frac{n}{2\pi}} \log n \mathbb{E}\left(\sqrt{\frac{n - \Lambda_{\alpha_n} \wedge n}{n}}\right) \end{aligned}$$

We have

$$\frac{n - (\Lambda_{\alpha_n} \wedge n)}{n} \xrightarrow{\mathcal{L}} 1 - (\Lambda_{\mu}^{\mathcal{B}} \wedge 1) \quad (4.17)$$

in which $\Lambda_{\mu}^{\mathcal{B}}$ is the hitting time of μ by a standard Brownian motion \mathcal{B} (see the proof below). The functions $x \rightarrow \sqrt{1-x}$ and $x \rightarrow \sqrt{1-x} \log(1-x)$ are sufficiently “nice” (see Revuz-Yor [9]), so we have

$$\lim_{n \rightarrow +\infty} \mathbb{E}\left(\sqrt{\frac{n - \Lambda_{\alpha_n} \wedge n}{n}} \log\left(\frac{n - \Lambda_{\alpha_n} \wedge n}{n}\right)\right) = \int_0^1 \sqrt{1-x} \log(1-x) \frac{\mu \exp(-\frac{\mu^2}{2x})}{\sqrt{x^3}} dx$$

and

$$\lim_{n \rightarrow +\infty} \mathbb{E}\left(\sqrt{\frac{n - \Lambda_{\alpha_n} \wedge n}{n}}\right) = \int_0^1 \sqrt{1-x} \frac{\mu \exp(-\frac{\mu^2}{2x})}{\sqrt{x^3}} dx.$$

The first integral converge and the second one satisfies

$$\int_0^1 \sqrt{1-x} \frac{\mu \exp(-\frac{\mu^2}{2x})}{\sqrt{x^3}} dx = \frac{\mu}{2} \Gamma(-\frac{1}{2}, \frac{\mu^2}{2}) = \sqrt{2\pi} C_\mu$$

due to formula 9 p. 319 of Gradshteyn [4]. So, as expected,

$$\mathbb{E}\left(\sum_{j=\Lambda_{\alpha_n} \wedge n+1}^n Wt_{dj^{\alpha_n}(n)}(j)\right) \sim C_\mu \sqrt{n} \log n. \quad \square$$

Proof of equation (4.17): We define the random variables $\{\tilde{T}_i\}_{i=1, \dots, \alpha_n}$ by:

$$\tilde{T}_1 = \inf\{k | S_k = 1\}, \tilde{T}_2 = \inf\{k | S_k = 2\} - \tilde{T}_1, \dots, \tilde{T}_{\alpha_n} = \inf\{k | S_k = \alpha_n\} - (\tilde{T}_1 + \dots + \tilde{T}_{\alpha_n-1}).$$

We have $\tilde{T}_a = \Lambda_a - \Lambda_{a-1}$ and $\Lambda_{\alpha_n} = \sum_{k=1}^{\alpha_n} \tilde{T}_k$. The random variables \tilde{T}_i are i.i.d. and have (see Feller [3] p. 272) the following Laplace transforms

$$\Phi_{\frac{\tilde{T}_1}{n}}(s) = \mathbb{E}(e^{-s \frac{\tilde{T}_1}{n}}) = e^{\frac{s}{n}} (1 - \sqrt{1 - e^{-\frac{2s}{n}}})$$

So, the Laplace transform of Λ_{α_n} satisfies

$$\lim_{n \rightarrow +\infty} \Phi_{\Lambda_{\alpha_n}}(s) = \lim_{n \rightarrow +\infty} \left(\Phi_{\frac{\tilde{T}_1}{n}}(s)\right)^{\alpha_n} = e^{-\mu \sqrt{2s}} = \Phi_{\mathcal{B}}(s)$$

for the last equation see Karatzas-Shreve [6] p.196. We conclude by Paul Lévy's continuity Theorem. \square

Acknowledgments. The author is grateful to Philippe Chassaing for the many helpful discussions, careful reading and stimulating criticisms about this work and to Hsien-Kuei Hwang for his useful remarks.

References

- [1] B. Bollobás, *Random Graphs* , Academic Press, p.12 (1985).
- [2] P. Chassaing, *How many questions are needed to find out the maximum of a Random Walk*, Preprint Elie Cartan
- [3] W. Feller, *An introduction to probability theory and its application*, Volume 1, Third edition, Wiley International Edition.
- [4] I.S. Gradshteyn, I. M. Ryzhik, *Table of Integral, Series and Products*, Academic Press p. 319 (9). (1980)
- [5] H.K. Hwang, *A constant arising from the analysis of algorithms for determining the maximum of a random walk*, R.S.A. Vol. 10 p. 333-335, (1997).
- [6] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag (1988).
- [7] E. Novak, K. Ritter, *Some complexity Results for Zero finding for Univariate Function*, Journal of Complexity 9,15-40 (1993).
- [8] A.M. Odlyzko, *Search for the maximum of a random walk*, R.S.A. Vol.6 p.275-295 (1995).
- [9] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, p. 9, 10.(1990)
- [10] F. Spitzer, *Principes of Random Walk* , D. Van Nostrand Compagny, p.61, p. 352 (1964)

CHAPITRE II

A Stochastically Quasi-Optimal Search Algorithm for the Maximum of the Simple Random Walk¹

P. Chassaing, J.F. Marckert, M. Yor

Odlyzko (1995) examines the efficiency of various strategies for searching in an unknown environment. He describes an average case quasi-optimal algorithm for the search of the maximum of a length- n symmetric random walk. In this paper, we prove that the same algorithm is also quasi-optimal for the stochastic order and we characterize the limit law of its cost.

1 Introduction

In a remarkable paper, Odlyzko (1995) introduces a new model that proves to be particularly suitable for the study of various searching strategies in an unknown environment (see the discussion in Odlyzko (1995) p.278-279, and also Chassaing (1999), Marckert (1997)).

The model is as follows: consider $\omega = (S_k(\omega))_{k=0, \dots, n}$ a simple symmetric random walk. That is, $S_0(\omega) = 0$ and for any $0 \leq k \leq n-1$, $S_{k+1}(\omega) = S_k(\omega) \pm 1$, the 2^n such sample paths being equiprobable. Set

$$M_n(\omega) = \max\{S_k(\omega) \mid 0 \leq k \leq n\}.$$

The knowledge of one of the $S_l(\omega)$'s for an integer l of our choice costs 1. The question is: How much will we have to pay in order to find M_n ? The cost of an algorithm is naturally identified with the number of probes it uses. For instance, the simplest and worst algorithm consists in probing each position l , for $1 \leq l \leq n$; its cost is n .

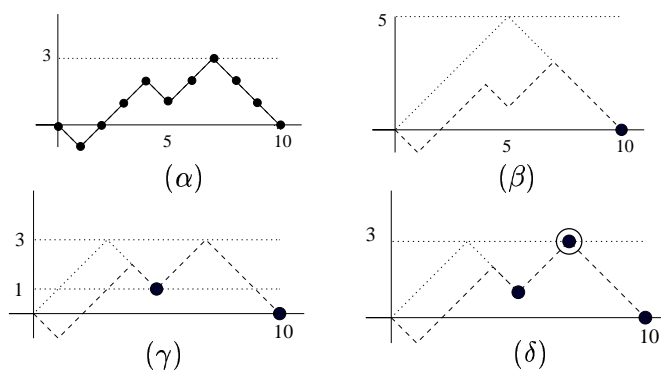


Figure 1.1

In (Figure (α)), $n = 10$ and $M_n = 3$. At the beginning, we know that $0 \leq M_{10} \leq 10$. The first probe shows that $S_{10} = 0$ and accordingly $0 \leq M_{10} \leq 5$ (Figure (β)). The second probe

¹AMS 1991 subject classifications. Primary: 68Q25, 60J65; Secondary: 60F17, 68P10, 90B40

Key words and phrases: analysis of algorithms, searching, random walk, stochastic order, Brownian motion, queues

indicates $S_5 = 1$; we can deduce that $1 \leq M_{10} \leq 3$ as one can see on the Figure (γ). Finally the third probe (Figure (δ)) gives us $M_{10} = 3$.

Denote by A_n the set of algorithms that find M_n . For an algorithm a in A_n , set $T_a(\omega)$ the number of probes that a requires to find the maximum of the path ω . (T_a is the *cost* of algorithm a). Odlyzko (1995) proves that

$$\min_{a \in A_n} E(T_a) = c_0 \sqrt{n} + o(\sqrt{n})$$

where c_0 is given by

$$c_0 = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \frac{dy}{y} \int_0^1 \frac{1}{\sqrt{w}} \exp\left(-\frac{y^2}{2w}\right) \operatorname{erf}\left(\frac{y}{\sqrt{2(1-w)}}\right) dw$$

and he describes an algorithm, which we shall denote $Od(n)$, such that

$$E(T_{Od(n)}) = c_0 \sqrt{n} + o(\sqrt{n}).$$

$Od(n)$ is called *average case quasi-optimal*. A description of $Od(n)$ is given in Section 5. Hwang (1997) and Chassaing (1999) independently computed

$$c_0 = \sqrt{\frac{8}{\pi}} \log 2.$$

In this paper, we prove that $Od(n)$ is quasi-optimal, not only through its expectation, but through its probability distribution, for the stochastic order.

In the last decades, the probabilistic analysis of algorithms has been going more and more frequently beyond the mere computation of the average cost; in fact one has often studied the variance or even the limit distribution (see for instance [24], [23], [43] or the *Algorithmica* issue [2] of 1998). Except for the study of comparison sorting (see [1] p.92, [33] p.193-195 or [45] Ch. 1, for historical notes), to our knowledge, the average case quasi-optimality has been proven only in a few cases (see [3], [12], [13], [17], [38], [39]).

In Section 2 we present a motivation to go farther than mere average case quasi-optimality, as we do in the next theorems.

Set

$$I = \frac{1}{2} \int_0^1 \frac{dt}{M_1^{\mathcal{B}} - \mathcal{B}_t},$$

where $(\mathcal{B}_t)_{t \geq 0}$ is a one-dimensionnal Brownian motion, and $M_1^{\mathcal{B}} = \max\{\mathcal{B}_t | 0 \leq t \leq 1\}$; also, let

$$\Phi_n(x) = \min_{a \in A_n} \Pr\left(\frac{T_a}{\sqrt{n}} \geq x\right).$$

Theorem 1.1 *For every $x \geq 0$,*

$$\lim_n \Pr\left(\frac{T_{Od(n)}}{\sqrt{n}} \geq x\right) = \Pr(I \geq x)$$

Theorem 1.2

$$\lim_n \Phi_n(x) = \Pr(I \geq x).$$

To be formal, in a context where A_n is the set of algorithms solving a given problem for an input of size n , we call a sequence of algorithms $a(n)$ ($a(n) \in A_n$) *stochastically quasi-optimal* if there exists a normalizing coefficient c_n (here $c_n = \sqrt{n}$) such that

$$\lim_n \Phi_n(x) = \lim_n \Pr\left(\frac{T_{a(n)}}{c_n} \geq x\right),$$

and such that the limit is the tail of a proper distribution or of a point mass at some positive constant. Concerning I , a rather unusual functional of the Brownian motion, the simplest description of its distribution (previously unknown, to our knowledge) is given by the Laplace transform of $\frac{1}{I^2}$:

Theorem 1.3 For every $t \geq 0$,

$$E\left[\exp\left(-\frac{t}{I^2}\right)\right] = \frac{2 \log(\cosh \sqrt{2t})}{\sinh^2 \sqrt{2t}}.$$

The paper is organized as follows: Section 3 contains the main steps of the proofs of Theorems 1.1 and 1.2. Section 4 contains the proof of Theorem 1.3 and its consequences: expressions of the density of probability of I in terms of Jacobi's theta function, and of the moments of I in terms of the Riemann zeta function. Sections 5 and 6 contain the key counting principle leading to the limit law of $T_{Od(n)}$, a description of the algorithm $Od(n)$ and some technicalities.

2 Optimality criteria

In the queuing context, a computer is usually seen as a server, the customers are, for instance, computations to be made by the computer (see Kleinrock (1976) or recently Louchard et al. (1997)), and the service time is the cost of the algorithm the computer uses. In this Section we discuss the choice between two algorithms a and b , solving the same problem, and with respective costs (service times) T_a or T_b . We focus on how this choice affects various parameters of the queuing system.

A first parameter is the total busy time until the departure of the n^{th} customer, that is, for $i \in \{a, b\}$

$$T_i^{(1)} + T_i^{(2)} + \dots + T_i^{(n)} = nE(T_i) + \sqrt{n\text{Var}(T_i)}Z_i^{(n)}$$

where $T_i^{(n)}$ denotes the service time of the n^{th} customer, when algorithm i is in use, and where $Z_i^{(n)}$ converges weakly to a standard Gaussian r.v.. According to this criterion, only

the average case analysis is really relevant, as the variance of the cost appears in the second term of the expansion.

Another possible parameter, of great concern to the customer, is the average sojourn time,

$$v_i = \lim_n \frac{1}{n} E[V_i^{(1)} + V_i^{(2)} + \dots + V_i^{(n)}]$$

where $V_i^{(n)}$ denotes the sojourn time (i.e. waiting time plus service time) of the n^{th} customer in the queuing system, when algorithm i is in use. Unfortunately v_i depends not only on the distribution of T_i but also, implicitly, on the distribution of the arrival point process. We give two examples of arrival point processes.

If customers arrive according to a Poisson process with intensity λ , the average sojourn time is given by one of the Pollaczek-Khintchine formulas:

$$v_i(\lambda) = E(T_i) + \frac{\lambda E(T_i^2)}{1 - \lambda E(T_i)} = \frac{E(T_i) + \lambda \text{Var}(T_i)}{1 - \lambda E(T_i)}$$

if $\lambda E(T_i) < 1$, $v_i(\lambda) = +\infty$ otherwise. The second moment of the distribution comes here into play, with the consequence that for a given value of λ , it is quite possible that $E(T_a) \leq E(T_b)$ but that the use of algorithm b in place of algorithm a results in an improvement of the average sojourn time for small λ 's, for instance if $\text{Var}(T_b)$ is a lot smaller than $\text{Var}(T_a)$. It is also possible that $\text{Var}(T_a) \leq \text{Var}(T_b)$ but that the use of algorithm a in place of algorithm b increases the average sojourn time for *any* λ , if $E(T_b)$ is smaller than $E(T_a)$. Actually, a will behave better than b , in the sense that $v_a(\lambda) \leq v_b(\lambda)$ for *any* λ , if and only if we have simultaneously $E(T_a) \leq E(T_b)$ and $E(T_a^2) \leq E(T_b^2)$.

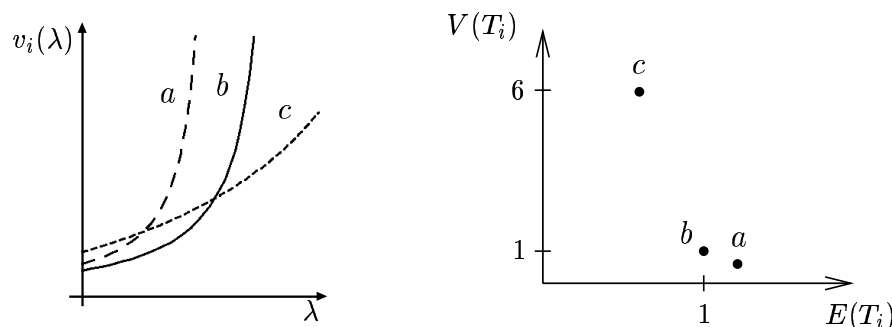


Figure 2.1 : *Repercussion of the service time on the sojourn time ($i \in \{a, b, c\}$).*

Now, if the customers arrive according to a renewal process with interarrival law μ and n^{th} interarrival time σ_n , $v_i(\mu)$ is given by Spitzer's identity:

$$v_i(\mu) = E(T_i) + \sum_{n \geq 1} \frac{1}{n} E[(s_n)_+]$$

(see Asmussen (1987) Ch. VIII, Spitzer (1957), or Kleinrock (1975) p.185) where s_n is the random walk with increment $T_i^{(n)} - \sigma_n$, starting from 0.

It would be highly valuable to the algorithm designer to be sure that the average sojourn time will be smaller for some algorithm a than for some algorithm b , *regardless of the arrival process*. Unfortunately, we cannot discard the possibility that for some μ we have simultaneously $E(T_a) \leq E(T_b)$, $E(T_a^2) \leq E(T_b^2)$ and nevertheless $v_a(\mu) \geq v_b(\mu)$. But we know that if T_b is *stochastically greater* than T_a ($T_b \geq_S T_a$), that is, if

$$\forall x \in \mathbf{R}, \Pr(T_b \geq x) \geq \Pr(T_a \geq x)$$

then we shall have (Stoyan (1983), Baccelli & Bremaud (1994)), for any probability measure μ :

$$v_b(\mu) \geq v_a(\mu).$$

So the existence of an algorithm $a(n) \in A_n$ such that for every x

$$\Pr\left(\frac{T_{a(n)}}{\sqrt{n}} \geq x\right) = \Phi_n(x)$$

would guarantee the optimality of $v_{a(n)}(\mu)$ for any μ , as well as the optimality of similar parameters such as the average waiting time or the average queue length. Unfortunately, it is doubtful that an optimal algorithm for the stochastic order will exist. Nevertheless Theorems 1.1 and 1.2 show that $Od(n)$ enjoys *asymptotically* the desired property. We shall say that $Od(n)$ is stochastically *quasi-optimal*.

Note that there exists a weaker sufficient condition for T_a to dominate T_b with respect to the average sojourn time: according to Stoyan (1983) Ch. 5.2 or Baccelli & Brémaud (1994) Ch. 4, v_a is smaller than v_b , independent of the arrival process, as soon as T_a is smaller than T_b for the *increasing convex order*, that is, as soon as:

$$\forall x, \int_x^{+\infty} \Pr(T_a \geq u) du \leq \int_x^{+\infty} \Pr(T_b \geq u) du.$$

3 Main steps of proofs of Theorems 1.1 and 1.2

Set

$$V_n = \frac{1}{2} \sum_{k=1}^n \frac{1}{M_n - S_k + 1}$$

The idea of the proofs of this theorems is, as in Odlyzko (1995), that on one hand $T_{Od(n)}$ is well approximated by V_n (Lemma 3.2), while on the other hand V_n is a lower bound for the cost T_a of any algorithm a in A_n (Lemma 3.1), due to a counting principle explained briefly at the end of this Section, and at length in Section 5. Compared with [39] or [13], note that, at a small additional cost, we replace convergence of expectations by \mathcal{L}_1 -convergence. Set

$$T_n(\omega) = \inf\{T_a(\omega) | a \in A_n\}.$$

We have:

Lemma 3.1

$$\frac{(V_n - T_n)_+}{\sqrt{n}} \xrightarrow{\mathcal{L}_1} 0.$$

and

Lemma 3.2

$$\frac{T_{Od(n)} - V_n}{\sqrt{n}} \xrightarrow{\mathcal{L}_1} 0$$

Having in mind Donsker's invariance principle, it is natural to think that the following holds

Theorem 3.3

$$\frac{V_n}{\sqrt{n}} \xrightarrow{law} I = \frac{1}{2} \int_0^1 \frac{dt}{M_1^{\mathcal{B}} - \mathcal{B}_t}.$$

However the result is not so straightforward, as I is not a continuous functional of the Brownian path. The proof of Theorem 3.3 is delayed to Section 6, where we exhibit a coupling of V_n/\sqrt{n} and I through a Skorohod embedding.

Theorem 1.1 is an immediate consequence of Lemma 3.2 and of Theorem 3.3. Let us prove Theorem 1.2:

Proof of Theorem 1.2: We have

$$\Pr(T_{Od(n)} \geq x\sqrt{n}) \geq \Phi_n(x) \geq \Pr(T_n \geq x\sqrt{n}), \quad (3.1)$$

the last inequality being due to the fact that $\forall a \in A_n, T_n \leq T_a$. On the other hand, as $T_{Od(n)} \geq T_n$, we obtain

$$\frac{(T_{Od(n)} - V_n)_+}{\sqrt{n}} \geq \frac{(V_n - T_n)_-}{\sqrt{n}}$$

so that Lemmata 3.2 and 3.1 yield

$$\frac{T_n - V_n}{\sqrt{n}} \xrightarrow{\mathcal{L}_1} 0.$$

Finally, owing to Theorem 3.3, both sides of (3.1) have the same limit, and

$$\lim_n \Phi_n(x) = \Pr(I \geq x). \quad \square$$

Remark : Let us explain briefly the (somewhat unexpected) appearance of V_n . If the speed of a traveler at point y of the line is bounded above by some function $z(y)$, then the duration t of the journey from point 0 to point x satisfies:

$$t \geq \int_0^x \frac{dy}{z(y)}. \quad (3.2)$$

But every algorithm a in A_n defines an increasing sequence \tilde{a} of positive integer-valued random variables $\{m_1, \dots, m_{T_a}\}$, the sequence \tilde{a} of positions probed by a during its runtime, and two consecutive positions m_i and m_{i+1} of \tilde{a} satisfy, as will be explained below:

$$m_{i+1} - m_i \leq 2M_n - S_{m_i} - S_{m_{i+1}}. \quad (3.3)$$

Moreover, due to slow variation of many of the paths of the random walk, with a high probability we can replace inequality (3.3) by

$$m_{i+1} - m_i \leq Z_{m_i}^{(n)}(1 + o(n)),$$

where

$$Z_k^{(n)} = 2(M_n - S_k + 1).$$

Thus $m_{i+1} - m_i$ can be seen as the speed of the algorithm at point m_i , and $Z_{m_i}^{(n)}$ as the speed limit, yielding a lower bound of V_n for the running time of the algorithm. So to prove Lemmata 3.2 and 3.1 we need a stochastic variant of inequality (3.2). This variant is given by Theorems 5.1 and 5.2 of Section 5. Note that we *do not* assume the m_i 's to be numbered chronologically: the lower bound V_n holds also for algorithms which *do not* probe the sample path from left to right.

Inequality (3.3) is better understood by looking at Figures 1.1(δ), or 3.1. We see on Figure 3.1 that the upper envelope of paths that agree with the result of probes at positions m_i and m_{i+1} has a maximum, on the interval $[m_i, m_{i+1}]$, that is given by

$$\frac{S_{m_i} + S_{m_{i+1}} + w_i}{2},$$

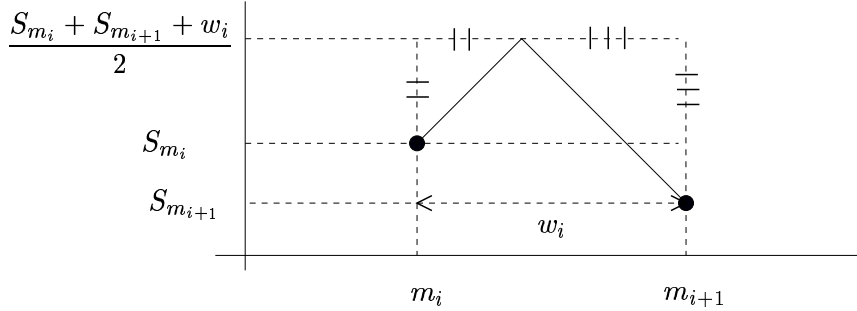


Figure 3.1

(see Figure 3.1). If (3.3) does not hold, we still do not know M_n for sure at the end of the runtime (see Figure 1.1(δ)).

4 Proof of Theorem 1.3

Theorem 1.3 is a consequence of the following remarkable property of I :

Theorem 4.1 :

$$\sqrt{SI} \stackrel{\text{law}}{=} \operatorname{arctanh} U + \operatorname{arctanh} V \quad (4.4)$$

where U and V are independent uniform variables on $[0, 1]$ and S is an exponential variable, with mean value 2, independent of I . The function $\operatorname{arctanh}$ is defined as usual by $\operatorname{arctanh}(u) = \frac{1}{2} \log \left(\frac{1+u}{1-u} \right)$.

In the sequel, $\operatorname{arctanh} U$ will be denoted by X, Y or Z .

Proof of Theorem 1.3 : On one hand:

$$\Pr(\sqrt{SI} \geq x) = \Pr(S \geq \frac{x^2}{I^2}) = E\left(\exp\left(-\frac{x^2}{2I^2}\right)\right)$$

and on the other hand:

$$\begin{aligned} \Pr(\operatorname{arctanh} U + \operatorname{arctanh} V \geq x) &= \Pr\left(\frac{U+V}{1+UV} \geq \tanh x\right) \\ &= \frac{1}{\sinh^2 x} \log(\cosh^2 x). \quad \square \end{aligned}$$

Proof of Theorem 4.1: Let $g = \sup\{s \leq 1 : \mathcal{B}_s = 0\}$ be the last zero in $[0, 1]$ of Brownian motion. The Brownian meander \mathcal{m} is the process

$$\mathcal{m}_t = \left| \frac{1}{\sqrt{1-g}} \mathcal{B}_{g+(1-g)t} \right| \text{ for } 0 \leq t \leq 1$$

Let θ be the unique time t in $]0, 1[$ such that

$$M_1^{\mathcal{B}} = B_\theta.$$

Using the decomposition due to Denisov (1983) (see also Biane & Yor (1988)) of the sample path $(\mathcal{B}_s)_{0 \leq s \leq 1}$:

$$\mathcal{B}_s = \sqrt{\theta} (\mathcal{m}_1^- - \mathcal{m}_{1-\frac{s}{\theta}}^-) I_{0 \leq s \leq \theta} + (\sqrt{\theta} \mathcal{m}_1^- - \sqrt{1-\theta} \mathcal{m}_{\frac{s-\theta}{1-\theta}}^+) I_{\theta \leq s \leq 1} \quad (4.5)$$

where $(\mathcal{m}_s^-, 0 \leq s \leq 1)$ and $(\mathcal{m}_s^+, 0 \leq s \leq 1)$ are two independent Brownian meanders also independent from θ . We get:

$$I = \int_0^1 \frac{du}{M_1^{\mathcal{B}} - \mathcal{B}_u} = \sqrt{\theta} \left(\int_0^1 \frac{du}{\mathcal{m}_u^-} \right) + \sqrt{1-\theta} \left(\int_0^1 \frac{du}{\mathcal{m}_u^+} \right)$$

Set

$$X = \sqrt{\theta S} \left(\int_0^1 \frac{du}{2 \mathcal{m}_u^-} \right)$$

and

$$Y = \sqrt{(1-\theta)S} \left(\int_0^1 \frac{du}{2m_u^+} \right).$$

Owing to (i) and (ii) below, X and Y are independent and identically distributed, and (4.4) reduces to proving that $\tanh(X)$ is uniform. This last fact relies upon (iii) and (iv):

i) θ , $1-\theta$ and g are arcsine distributed,

ii) $S\theta$ and $S(1-\theta)$ are independent, and satisfy:

$$\sqrt{S\theta} \stackrel{law}{=} \sqrt{S(1-\theta)} \stackrel{law}{=} |N| \quad (4.6)$$

where N denotes a Gaussian random variable with mean 0 and variance 1,

iii) $\frac{1}{2} \int_0^1 \frac{du}{m_u} \stackrel{law}{=} \sup_{0 \leq t \leq 1} |b_t|$, where $(b_t, 0 \leq t \leq 1)$ is a standard Brownian bridge,

iv) Z , defined by $Z \stackrel{def}{=} |N| \sup_{t \leq 1} |b_t|$, where N is independent of b , satisfies: $\Pr(Z \leq x) = \tanh(x)$.

Points (i) and (ii) are well-known, and (iii) can be found in Biane & Yor (1987) formula 5)c)i p. 69 (see also [7]); (iv) may be obtained using excursion theory (see Revuz & Yor (1994), Ch. XII, Exercise (4.24), question 3), but we shall give below a nice proof which is a special instance of methods developed in Carmona et al. [11]. Let $A_t = \sup_{s \leq t} \{\mathcal{B}_s^2\}$. Using the fact that $(b(t))_{0 \leq t \leq 1} = (\frac{1}{\sqrt{g}} \mathcal{B}_{tg})_{0 \leq t \leq 1}$ is a Brownian bridge independent of g , we obtain

$$A_g = g \cdot \sup_{0 \leq t \leq 1} \{b_t^2\}$$

and thus,

$$S A_g \stackrel{law}{=} Z^2$$

Set $\alpha_a = \inf\{s > 0 \mid A_s > a\}$ and $d_t = \inf\{u \geq t \mid \mathcal{B}_u = 0\}$. For $a \geq 0$, we have:

$$\{A_g < a\} = \{g < \alpha_a\} = \{1 < d_{\alpha_a}\}$$

Using the scaling property of Brownian motion we obtain:

$$d_{\alpha_a} \stackrel{law}{=} a d_{\alpha_1}$$

and thus

$$A_g \stackrel{law}{=} \frac{1}{d_{\alpha_1}}$$

Finally,

$$\Pr(Z \geq x) = \Pr(S A_g \geq x^2) = E\left[\exp\left(-\frac{x^2}{2}d_{\alpha_1}\right)\right]$$

But α_1 and $d_{\alpha_1} - \alpha_1$ are independent due to the Markov property, and

$$d_{\alpha_1} - \alpha_1 \stackrel{law}{=} \inf\{t > 0 | \mathcal{B}_t = 1\}$$

so, using the well known Laplace transforms of α_1 and $d_{\alpha_1} - \alpha_1$, we get

$$\Pr(Z \geq x) = \frac{1}{\cosh(x)} \exp(-x) = 1 - \tanh(x). \quad \square$$

Joint law of $M_1^{\mathcal{B}}$ and I

Through Skorohod embedding, it is straightforward to extend Theorem 3.3 to the following:

$$\frac{1}{\sqrt{n}}(V_n, M_n, M_n - S_n) \xrightarrow{law} (I, M_1^{\mathcal{B}}, M_1^{\mathcal{B}} - \mathcal{B}_1).$$

Furthermore, arguments similar to those of Theorem 4.1 (see the computations p.268 in Pitman & Yor (1996)) yield that

$$|N|\left(\int_0^1 \frac{du}{\mathcal{M}_u}, \mathcal{M}_1\right) \stackrel{law}{=} (2X, \tanh(X)\Gamma),$$

where \mathcal{M}_1 is the value of the meander $(\mathcal{M}_u)_{0 \leq u \leq 1}$ at time 1. As a consequence

$$\sqrt{S}(I, M_1^{\mathcal{B}}, M_1^{\mathcal{B}} - \mathcal{B}_1) \stackrel{law}{=} (X + Y, (\tanh X)\Gamma, (\tanh Y)\Gamma') \quad (4.7)$$

where X, Y, Γ and Γ' are independent, and Γ and Γ' are $\gamma_{2,1}$ -distributed random variables. From (4.7), one finds

$$E\left(\frac{1}{M^2} \exp\left(-\frac{h^2}{2M^2}\right) I_{\frac{I}{M} \leq \frac{a}{h}}\right) = \int_0^y \frac{du}{u^2} \exp(-h/u) \frac{y-u}{1-yu}$$

where $y = \tanh(a)$. This shows, in some sense, how M_n and $T_{Od(n)}$ are related.

Density of I

Writing Theorem 1.3 in the form

$$E\left[\exp\left(-\frac{a^2}{2I^2}\right)\right] = \frac{\log(\cosh^2 a)}{a^2} \frac{a^2}{\sinh^2 a}, \quad (4.8)$$

we observe the familiar Laplace transform $(a/\sinh(a))^2$ in $a^2/2$: indeed, if Σ_1 denotes the hitting time of 1 for a Bessel process of dimension 3, starting from 0, then (see Chung (1976) p.168-169)

$$E(e^{-\lambda^2/2 \Sigma_1}) = \frac{\lambda}{\sinh \lambda}.$$

So $(a/\sinh(a))^2$ appears as the Laplace transform of the sum Σ_2 of two independent copies of Σ_1 . As noted by Chung, Σ_1 and Σ_2 are surprisingly related to the supremum $M_1^{|b|}$ of the absolute value of the Brownian bridge, i.e. the asymptotic distribution of the Kolmogorov-Smirnov statistic, resp. to the maximum M_1^e of the normalized Brownian excursion, by:

$$\Sigma_1 \stackrel{law}{=} \frac{4}{\pi^2} (M_1^{|b|})^2, \quad \Sigma_2 \stackrel{law}{=} \frac{4}{\pi^2} (M_1^e)^2$$

on this topic, see also the nice developments by Louchard (1984). The distributions of these variables are discussed in details in [7] and we have kept the notation of that paper for ease of reference. The distribution of M_1^e is also the distribution or limit distribution of:

- the maximum of a Bernoulli excursion ([28], [47], [50]),
- the normalized height for several classes of trees of size n ([21],[20]),
- the range of the Brownian bridge ([34], [41], [44], [52]).

Denote by U a uniform r.v. and let

$$\Sigma_1^\sharp = \inf\{t : |\mathcal{B}_t| = 1\} \text{ and } G = \sup\{t < \Sigma_1^\sharp : \mathcal{B}_t = 0\}$$

then Williams' decomposition [53], [54] tells us that G and $(\Sigma_1^\sharp - G)$ are independent, that $\sup_{u \leq G} |\mathcal{B}_u| \stackrel{law}{=} V$ where V is uniform on $[0, 1]$, whereas Knight [32] showed that

$$E(e^{-\lambda^2/2 G}) = \frac{\tanh(\lambda)}{\lambda}, \quad E(e^{-\lambda^2/2 (\Sigma_1^\sharp - G)}) = \frac{\lambda}{\sinh(\lambda)}.$$

Theorem 4.2 *We have*

$$\frac{1}{I^2} \stackrel{law}{=} U.G + \Sigma_2 \stackrel{law}{=} H\Sigma_2^\sharp + \Sigma_2,$$

where, on the right hand side, U , G and Σ_2 are independent, H, Σ_2^\sharp and Σ_2 are independent, Σ_2^\sharp is the sum of two independent copies of Σ_1^\sharp and H is distributed as:

$$P(H \in dh) = \left(\frac{1}{\sqrt{h}} - 1 \right) dh \, I_{0 < h < 1}. \quad (4.9)$$

Proof : Owing to (4.8), the first identity in law reduces to

$$\begin{aligned} E\left[\exp(-\lambda UG)\right] &= \frac{1}{\lambda} E\left[\int_0^\lambda e^{-xG} dx\right] \\ &= \frac{1}{\lambda} \int_0^\lambda \frac{\tanh \sqrt{2x}}{\sqrt{2x}} dx \\ &= \frac{\log(\cosh \sqrt{2\lambda})}{\lambda}. \end{aligned}$$

To prove the second identity in law, we use $G \stackrel{law}{=} V^2 \Sigma_2^\sharp$ which follows from Williams' decomposition and $H = UV^2$ is easily shown to be distributed as in (4.9).

According to Biane & Yor (1987) p. 74, the density of G is:

$$\sum_{n=1}^{+\infty} \exp\left(-\frac{\tilde{n}^2 x}{2}\right), \quad x > 0, \quad \tilde{n} = \pi\left(n - \frac{1}{2}\right),$$

thus UG has density:

$$\theta_1(x) = \sum_{n=1}^{+\infty} \int_x^{+\infty} \exp\left(-\frac{\tilde{n}^2 u}{2}\right) \frac{du}{u}.$$

According to Chung (1976) or Kennedy (1976), we have:

$$\begin{aligned} \Pr(\Sigma_2 \leq x) &= \sum_{-\infty \leq n \leq +\infty} (1 - 4n^2 x^2) \exp(-2x^2 n^2) \\ &= \left(\frac{\pi^3}{2}\right)^{1/2} \theta'\left(\frac{\pi}{2x^2}\right), \end{aligned}$$

where θ denotes the usual Jacobi's (theta) function

$$\theta(x) = \sum_{-\infty < n < +\infty} e^{-\pi n^2 x} \quad (x > 0),$$

and where the second equality makes use of the Poisson identity:

$$\theta(x) = \frac{1}{\sqrt{x}} \theta\left(\frac{1}{x}\right)$$

(for some probabilistic insight on Poisson identity, and other related identities, see Biane & Yor (1987) pp. 72-74, or Smith & Diaconis (1988)).

Thus, the density of Σ_2 is given by:

$$\theta_2(x) = \frac{\partial}{\partial x} \sum_{-\infty < n < +\infty} (1 - n^2 \pi^2 x) e^{-n^2 \pi^2 x/2},$$

and the density Θ of $1/I^2$ is given by:

$$\Theta(x) = \int_0^x dy \theta_1(y) \theta_2(x-y), \quad (x \geq 0) \quad (4.10)$$

We finally obtain

Theorem 4.3 *If we denote by γ the density of I , we have:*

$$\gamma(u) = \frac{2}{u^3} \Theta\left(\frac{1}{u^2}\right) \quad (u > 0).$$

Using the power series expansion for $\frac{x \log(1+x)}{(1-x)^2}$, one can write the Laplace transform of $1/I^2$ taken at $a^2/2$:

$$\frac{\log(\cosh^2 a)}{\sinh^2 a} = \frac{2a}{\sinh^2 a} + 8 \sum_{n=1}^{\infty} b_n e^{-2na}$$

where:

$$\begin{aligned} b_n &= n \overline{H}_n - n \log 2 - 1_{\{n \text{ odd}\}}, \\ \overline{H}_n &= \sum_{j=1}^n \frac{(-1)^{j-1}}{j}, \end{aligned}$$

and we obtain the following formula, to be compared with (4.10):

$$\Theta(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \frac{\theta_2(x-t)}{\sqrt{t}} dt + 8 \sqrt{\frac{2}{\pi x^3}} \sum_{n=1}^{\infty} n b_n \exp\left(-\frac{2n^2}{x}\right).$$

Moments of I

Let Γ (resp. ζ) denote the Euler gamma (resp. Riemann zeta) function. According to the notations of Flajolet & Salvy [22], set

$$S_{1,p}^{-+} = \sum_{n=1}^{+\infty} \frac{\overline{H}_n}{n^p},$$

$\tilde{\zeta}(0) = 1/2$, $\tilde{\zeta}(1) = \log 2$ and for $k \geq 2$:

$$\tilde{\zeta}(k) = \sum_{n \geq 1} \frac{(-1)^{n+1}}{n^k} = (1 - 2^{1-k}) \zeta(k).$$

Proposition 4.4 $E(I) = \sqrt{\frac{8}{\pi}} \log 2$, and for $n \geq 2$

$$E(I^n) = \frac{4\Gamma(\frac{n+1}{2})}{\sqrt{2^n \pi}} \sum_{k=0}^n \tilde{\zeta}(k) \tilde{\zeta}(n-k) \quad (4.11)$$

which is also equal to:

$$E(I^n) = \frac{4\Gamma(\frac{n+1}{2})}{\sqrt{\pi} 2^n} \left(2S_{1,n-1}^{-+} - 2 \log 2 \zeta(n-1) + (n-1) \zeta(n) - \tilde{\zeta}(n) \right). \quad (4.12)$$

As a consequence, $E(I^2) = \log^2 2 + \frac{\pi^2}{12}$, and $\text{Var}(I) = (1 - \frac{8}{\pi}) \log^2 2 + \frac{\pi^2}{12} \simeq 0.0794565\dots$

Proof:

- Using (4.4) we obtain

$$E(I^n) = \frac{E((X+Y)^n)}{E(S^{n/2})}.$$

Straightforward computations yield

$$E(S^{n/2}) = \Gamma\left(\frac{n}{2} + 1\right) 2^{n/2}$$

and

$$E(X^k) = \frac{k!}{2^{k-1}} \tilde{\zeta}(k),$$

yielding (4.11).

- Thanks to Theorem 1.3, we have for every positive function f :

$$\begin{aligned} \int_0^{+\infty} f(a) \frac{2 \log(\cosh a)}{(\sinh a)^2} da &= E\left(\int_0^{+\infty} f(a) \exp\left(-\frac{a^2}{2I^2}\right) da\right) \\ &= E\left(I \int_0^{+\infty} f(xI) \exp\left(\frac{-x^2}{2}\right) dx\right). \end{aligned}$$

Hence, the choice $f(x) \equiv x^{n-1}$ yields:

$$E(I^n) = \frac{2^{1-\frac{n}{2}}}{\Gamma(n/2)} \int_0^{+\infty} a^{n-1} \frac{2 \log(\cosh a)}{\sinh^2 a} da.$$

Using the duplication formula for the Gamma function and the same power series expansions as in the preceding subsection, we obtain (4.12). \square

Note that we recover this way a surprising identity concerning Euler sums, due to Sitaramachandra Rao [46] and rederived by residue computations in Flajolet & Salvy [22]:

$$2S_{1,n-1}^- = 2 \log 2 \zeta(n-1) - (n-1)\zeta(n) + 2\tilde{\zeta}(n) + \sum_{k=1}^{n-1} \tilde{\zeta}(k)\tilde{\zeta}(n-k).$$

Laplace transform of I

Corollary 4.5 *The Laplace transform of I^2 is given by:*

$$\begin{aligned} \Psi_{I^2}\left(\frac{\lambda^2}{2}\right) &= E\left(\exp\left(-\frac{\lambda^2 I^2}{2}\right)\right) \\ &= E(J_o(2\lambda(X+Y))) = \frac{1}{\pi} \int_{-1}^1 \frac{dr}{\sqrt{1-r^2}} (E(\exp(2i\lambda r X)))^2 \end{aligned} \quad (4.13)$$

where X and Y have the same meaning as in Theorem 4.1, and moreover:

$$J_o(\xi) \stackrel{\text{def}}{=} \frac{1}{\pi} \int_{-1}^1 \frac{dr}{\sqrt{1-r^2}} \exp(i\xi r)$$

Proof of Corollary 4.5: We multiply both sides of the identity in law (4.4) by $\varepsilon\sqrt{g}$, where ε is a symmetric Bernoulli variable (i.e: $\Pr(\varepsilon = \pm 1) = \frac{1}{2}$), and g is arcsine distributed, independent of ε , and the pair (ε, g) is independent of either side of (4.4). Consequently, we obtain, on the new left-hand side of (4.4): NI , where N is a standard Gaussian variable, independent of I . Now, we take the characteristic functions, with respect to the argument λ , of both sides. On the left-hand side, we obtain:

$$E\left(\exp\left(-\frac{\lambda^2 I^2}{2}\right)\right)$$

whereas, on the right-hand side, we obtain

$$E(J_o(\lambda(X + Y))),$$

since the c.f. of $\varepsilon\sqrt{g}$ is $J_o(\xi) \equiv \frac{1}{\pi} \int_{-1}^1 \frac{dr}{\sqrt{1-r^2}} \exp(i\xi r)$. \square .

It may be of some interest to make formula (4.13) in Corollary (4.5) more explicit. For this purpose, we present the following results: we first introduce the incomplete beta function:

$$B_x(a, b) = \int_0^x u^{a-1}(1-u)^{b-1} du$$

We are now able to express the characteristic function of X in terms of $B_{1/2}$. Recall that the distribution of X is given by $P(X \leq x) = \tanh(x) I_{\{x \geq 0\}}$.

Proposition 4.6 *The c.f. of X is identical to:*

$$E(\exp(i\alpha X)) = 2B_{1/2}\left(1 - \frac{i\alpha}{2}, 1 + \frac{i\alpha}{2}\right) \quad (4.14)$$

so that

$$E(\cos(\alpha X)) = \frac{\frac{\pi\alpha}{2}}{\sinh\left(\frac{\pi\alpha}{2}\right)} ; \quad E(\sin(\alpha X)) = 2 \int_0^{1/2} \sin\left(\frac{\alpha}{2} \log(-1 + 1/u)\right) du$$

Proof:

i) We shall show the general identity:

$$\int_0^{+\infty} \exp(\alpha x) \frac{1}{(\cosh x)^\beta} dx = 2^{\beta-1} B_{1/2}\left(\frac{\beta-\alpha}{2}, \frac{\beta+\alpha}{2}\right) \quad (4.15)$$

for $\beta > 0$ and $Re \alpha < \beta$, $\alpha \in C$. Assuming for a moment that formula (4.15) holds, we deduce formula (4.14) by taking $\beta = 2$, and replacing α by $i\alpha$.

ii) We now prove formula (4.15): making the change of variable $t = \exp(-2x)$ on the left hand of (4.15), which we now denote $l(\alpha, \beta)$, we obtain:

$$l(\alpha, \beta) = 2^{\beta-1} \int_0^1 \frac{1}{(1+t)^\beta} t^{\frac{\beta-\alpha}{2}-1} dt,$$

and making a further change of variable $u = \frac{t}{1+t}$, we have:

$$l(\alpha, \beta) = 2^{\beta-1} B_{1/2} \left(\frac{\beta - \alpha}{2}, \frac{\beta + \alpha}{2} \right)$$

iii) Denoting by $C(\alpha)$ the left-hand side of (4.14), we deduce from (4.15) that:

$$\operatorname{Re} C(\alpha) = 2 \operatorname{Re} B_{1/2}(1 - i\alpha/2, 1 + i\alpha/2) = B(1 - i\alpha/2, 1 + i\alpha/2),$$

so that with the help of the classical formula of complements for the gamma function, we obtain:

$$\operatorname{Re} C(\alpha) = \frac{\frac{\pi\alpha}{2}}{\sinh(\frac{\pi\alpha}{2})}.$$

The formula for $\operatorname{Im} C(\alpha)$ is deduced directly from (4.14).

5 A counting principle

A variant of this principle is the key result in the proof of average case optimality in Odlyzko [39] as well as in Chassaing [13]. At the end of this Section, we shall use it to prove Lemmata 3.2 and 3.1. We believe Odlyzko's principle can be interesting in its own right: for instance, it gives a combinatorial proof of a surprising identity involving functionals of the Brownian excursion due to Biane & Yor (1987) (see Chassaing et al. [15]), and it gives also results about the maximal width of trees Chassaing & Marckert [14]. We try to give Odlyzko's principle in a general form, in order to spare the reader many references to very similar arguments in this paper, and also [15].

Let $(Z_k^{(n)})_{0 \leq k \leq n}$ denote an array of positive integer-valued random variables. Let α be a positive constant, less than 1, and let $G(n)$ and $H(n)$ be two sequences of positive numbers. We set

$$U_n = \sum_{k=1}^n \frac{1}{Z_k^{(n)}},$$

and

$$\tilde{U}_n = \sum_{Z_k^{(n)} \leq H(n)} \frac{1}{Z_k^{(n)}},$$

Note that $U_n \leq n$ so that $u_n = E(U_n)$ is well defined. We assume that there exists a set Ω_n of sample paths such that

$$\Pr(\bar{\Omega}_n) = o\left(\frac{u_n}{n}\right), \quad (5.16)$$

and that for any $\omega \in \Omega_n$

$$\forall k, l, |Z_{k+l}^{(n)} - Z_k^{(n)}| \leq l^\alpha G(n).$$

Furthermore we assume that

$$G(n) \ll H(n)^{1-\alpha},$$

and that

$$E(\tilde{U}_n) \ll u_n. \quad (5.17)$$

We consider a finite or denumerable set \tilde{A}_n of strictly increasing sequences a of positive integer-valued *random variables* $\{m_1, \dots, m_{T(a)}\}$, whose length $T(a)$ is *random*, such that $m_0 = 0$, $m_{T(a)} = n$, and such that the increments

$$w_i = m_{i+1} - m_i$$

satisfy

$$w_i \leq Z_{m_i}^{(n)} \vee H(n), \quad (5.18)$$

at least when $\omega \in \Omega_n$ (at the end of this Section, we shall choose $Z_k^{(n)} = 2(M_n - S_k + 1)(1 + o(n))$) in order to prove Lemmata 3.2 and 3.1). Let \tilde{T}_n denote the infimum of lengths of sequences in \tilde{A}_n .

Maybe hypothesis (5.16) and (5.17) will seem less arbitrary if one considers a simple problem of computation, where a strictly increasing sequence $u_k^{(\varepsilon)}$ satisfy $u_0^{(\varepsilon)} = 0$ and $u_{k+1}^{(\varepsilon)} - u_k^{(\varepsilon)} = \varepsilon z(u_k^{(\varepsilon)})$. The first index k_ε such that $u_k^{(\varepsilon)} \geq 1$ (see Figure 5.1) is easily seen to satisfy

$$\begin{aligned} k_\varepsilon &\sim \int_0^1 \sum_{k \geq 0} \frac{1}{\varepsilon z(u_k^{(\varepsilon)})} 1_{[u_k^{(\varepsilon)}, u_{k+1}^{(\varepsilon)}]}(x) dx \\ &\sim \frac{1}{\varepsilon} \int_0^1 \frac{dy}{z(y)}, \end{aligned}$$

under hypothesis of continuity of z , if z is positive. One can see hypothesis (5.16) as a stochastic analog of the latter. As regards hypothesis (5.17), obviously it is the stochastic analog of a – strong – integrability hypothesis for $1/z$, if z has a pole.

Theorem 5.1

$$\frac{(U_n - \tilde{T}_n)_+}{u_n} \xrightarrow{L_1} 0$$

Theorem 5.2 Assume that, for a sequence $a_n = (m_i^{(n)})_{0 \leq i \leq T(a_n)} \in \tilde{A}_n$,

$$E(\#\{i | Z_{m_i}^{(n)} \leq H(n)\} | I_{\Omega_n}) \ll u_n,$$

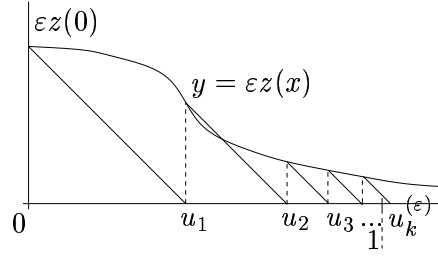


Figure 5.1

and that there exist a positive number δ , and a sequence ε_n of positive numbers decreasing to 0, such that for any integer n and any $\omega \in \Omega_n$, the sequence a_n satisfies

$$w_i^{(n)} \geq (1 - \varepsilon_n) Z_{m_i}^{(n)} \text{ if } (1 + \delta) Z_{m_i}^{(n)} \geq H(n).$$

Then

$$\frac{U_n - T(a_n)}{u_n} \xrightarrow{L_1} 0$$

Proof of Theorem 5.1

We shall prove that for n large enough, independently from the sequence $a \in \tilde{A}_n$, there exists a constant C such that:

$$(U_n - T(a))_+ \leq n I_{\tilde{\Omega}_n} + (\tilde{U}_n + C U_n G(n) H(n)^{\alpha-1}) I_{\Omega_n}, \quad (5.19)$$

so that:

$$(U_n - \tilde{T}_n)_+ \leq n I_{\tilde{\Omega}_n} + \tilde{U}_n + C U_n G(n) H(n)^{\alpha-1}.$$

Then hypothesis (5.16) and (5.17) yield Theorem 5.1 at once.

The obvious formula

$$T(a) = \sum_{i=0}^{T(a)-1} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{w_i}.$$

reflects the fact that $\frac{1}{w_i}$ is somehow the density of presence of points of the sequence a over the interval $[m_i, m_{i+1}]$. Inequality (5.18) entails that

$$T(a) \geq \sum_i \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_{m_i}^{(n)}}.$$

Theorem 5.1 follows from the fact that we can replace $\frac{1}{Z_{m_i}^{(n)}}$ by $\frac{1}{Z_k^{(n)}}$ in the last sum, due to the hypothesis (5.16) of slow variation of $Z_k^{(n)}$, obtaining thus a uniform lower bound U_n . This approximation is tight only if $Z_{m_i}^{(n)}$ and $Z_k^{(n)}$ are far enough from 0, but relation (5.17) takes care of the terms close to 0. So, in order to obtain (5.19) we bound $(U_n - T_a)_+$ by n for ω in $\bar{\Omega}_n$, and we proceed as follows for ω in Ω_n :

$$\begin{aligned} (U_n - T(a))_+ &= \left(\sum_i \sum_{k=m_i}^{m_{i+1}} \frac{w_i - Z_k^{(n)}}{Z_k^{(n)} w_i} \right)_+ \\ &\leq \sum_i \sum_{k=m_i}^{m_{i+1}} \frac{1}{Z_k^{(n)}} \left(\frac{w_i - Z_k^{(n)}}{w_i} \right)_+. \end{aligned}$$

- Assume that $w_i \geq \frac{H(n)}{3}$. For $k \in [m_i, m_{i+1}]$, we have

$$\left(\frac{w_i - Z_k^{(n)}}{w_i} \right)_+ \leq \left(\frac{Z_{m_i}^{(n)} - Z_k^{(n)}}{w_i} \right)_+ \leq \frac{G(n) w_i^\alpha}{w_i} \leq C G(n) H(n)^{\alpha-1}$$

- Assume that $Z_{m_i}^{(n)} - w_i \geq \frac{H(n)}{3}$. For $k \in [m_i, m_{i+1}]$, we have

$$w_i - Z_k^{(n)} \leq \frac{-H(n) + Z_{m_i}^{(n)} - Z_k^{(n)}}{3} \leq \frac{-H(n) + H(n)^\alpha G(n)}{3}$$

the last term being negative if n is large enough, independently of a and ω , so that $(w_i - Z_k^{(n)})_+ = 0$.

- Finally, let Δ be the set of indices i such that $w_i \leq \frac{H(n)}{3}$ and $Z_{m_i}^{(n)} - w_i \leq \frac{H(n)}{3}$. For $k \in [m_i, m_{i+1}]$, we have

$$Z_k^{(n)} = Z_k^{(n)} - Z_{m_i}^{(n)} + (Z_{m_i}^{(n)} - w_i) + w_i \leq \frac{H(n)^\alpha G(n)}{3} + \frac{2H(n)}{3} \leq H(n),$$

for n large enough so that

$$\begin{aligned} \left(\sum_{i \in \Delta} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} - \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{w_i} \right)_+ &\leq \sum_{i \in \Delta} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} \\ &\leq \tilde{U}_n. \quad \square \end{aligned}$$

Proof of Theorem 5.2

Similarly, let Δ' be the set of indices i such that $(1 + \delta) Z_{m_i}^{(n)} \leq H(n)$. For $\omega \in \Omega_n$, we have:

$$\left| \sum_{i \in \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} - \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{w_i} \right| \leq \#\Delta' + \sum_{i \in \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}}$$

But for $i \in \Delta'$ and $k \in [m_i, m_{i+1}]$, we have

$$\begin{aligned} Z_k^{(n)} &\leq Z_{m_i} + G(n)w_i^\alpha \\ &\leq \frac{H(n)}{1+\delta} + G(n)H(n)^\alpha \\ &\leq H(n), \end{aligned}$$

provided that n is large enough, so that

$$\sum_{i \in \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} \leq \tilde{U}_n,$$

and finally:

$$\left| \sum_{i \in \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} - \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{w_i} \right| \leq \#\{i | Z_{m_i}^{(n)} \leq H(n)\} + \tilde{U}_n.$$

On the other hand, for $i \notin \Delta'$ we have $(1+\delta)Z_{m_i}^{(n)} \geq H(n)$, so that for $k \in [m_i, m_{i+1}]$,

$$\begin{aligned} \left| \frac{w_i - Z_k^{(n)}}{w_i} \right| &\leq \frac{|w_i - Z_{m_i}^{(n)}|}{w_i} + \frac{|Z_{m_i}^{(n)} - Z_k^{(n)}|}{w_i} \\ &\leq \frac{\varepsilon_n}{1-\varepsilon_n} + G(n)w_i^{\alpha-1} \\ &\leq \frac{\varepsilon_n}{1-\varepsilon_n} + \left(\frac{1-\varepsilon_n}{1+\delta}\right)^{\alpha-1} G(n)H_n^{\alpha-1}, \end{aligned}$$

and

$$\begin{aligned} \left| \sum_{i \notin \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} - \frac{1}{w_i} \right| &\leq \sum_{i \notin \Delta'} \sum_{k=m_i+1}^{m_{i+1}} \frac{1}{Z_k^{(n)}} \left| \frac{w_i - Z_k^{(n)}}{w_i} \right| \\ &\leq U_n o(n). \end{aligned}$$

Finally:

$$\frac{|U_n - T(a_n)|}{u_n} \leq \frac{U_n o(n) + \tilde{U}_n + \#\{i | Z_{m_i}^{(n)} \leq H(n)\} I_{\Omega_n} + n I_{\bar{\Omega}_n}}{u_n}. \quad \square$$

Proof of Lemma 3.1

We set $\alpha = \frac{1}{2}$, $G(n) = c\sqrt{\log n}$, $H(n) = \log^3 n$, and for a in A_n , we consider the sequence $\tilde{a} = \{m_1, \dots, m_{T(\tilde{a})}\}$ obtained once we sort positions probed by a in increasing order, \tilde{a} being eventually supplemented by position n . So the finite set \tilde{A}_n of such sequences satisfy

$$\tilde{T}_n - 1 \leq T_n \leq \tilde{T}_n. \quad (5.20)$$

For $\omega \in \Omega_n$, (3.3) yields:

$$w_i \leq 2(M_n - S_{m_i}) + c\sqrt{w_i \log n}$$

and when $w_i \geq H(n)$,

$$w_i \leq 2\frac{M_n - S_{m_i} + 1}{1 - c/\log n}.$$

Thus, choosing:

$$Z_k^{(n)} = 2\frac{M_n - S_k + 1}{1 - c/\log n},$$

on one hand we see that

$$\frac{U_n - V_n}{u_n} \xrightarrow{L_1} 0$$

and on the other hand, owing to (5.20) and to Theorem 5.1,

$$\frac{(U_n + 1 - T_n)_+}{u_n} \xrightarrow{L_1} 0.$$

To finish the proof, we recall that, according to [39], [26] and [13]:

$$E(V_n) \simeq \sqrt{\frac{8}{\pi}} \log 2 \sqrt{n},$$

so that

$$u_n = \Theta(\sqrt{n}).$$

We also have to check hypotheses (5.16) and (5.17). First, for n large enough (else, let Ω_n be void), let Ω_d be the set $\Omega_d(n)$ of sample paths ω such that for any positive k and m

$$m + k \leq n \Rightarrow |S_{m+k} - S_m| \leq d\sqrt{k \log n}$$

for some d such that $d < c/2$. Due to Chernoff bounds (see for instance Bollobás (1985) p.12):

$$\forall x \geq 0, \forall k, \Pr(|S_k| > x) \leq 2 \exp\left(-\frac{x^2}{2k}\right) \quad (5.21)$$

we deduce as in [39]:

$$\Pr(\Omega_d(n)) \geq 1 - 2n^{2-\frac{d^2}{2}}. \quad (5.22)$$

so finally $c = 5$ does the job. To check hypothesis (5.17) we introduce the local time $D_q(n)$ of $M_n - S_k$ at height q , that is, $D_q(\omega) = \#\{k | 0 \leq k \leq n \text{ and } M_n - S_k = q\}$. According to [39], we have

Proposition 5.3 *There exists a constant C such that for every $n \in \mathbb{N}$*

$$E(D_q(n)) \leq C(q+1)$$

so that

$$\begin{aligned} \tilde{U}_n &\leq \frac{1-c/\log n}{2} \sum_{M_n - S_k \leq \log^3 n} \frac{1}{M_n - S_k + 1} \\ &\leq \sum_{q=0}^{\log^3 n} \frac{D_q(n)}{q+1} \end{aligned}$$

and

$$E(\tilde{U}_n) = O(\log^3 n). \quad \square$$

Proof of Lemma 3.2

Let us sketch a brief description of algorithm $Od(n)$ (for details we refer to [39], p. 285-286). Note that if at any step $Od(n)$ detects that ω belongs to $\bar{\Omega}_n$, then $Od(n)$ probes the n positions. In a first stage, $Od(n)$ searches a tight estimate M^* of M_n , that is an estimate that satisfies

$$M_n - M^* \leq cn^{\frac{1}{8}}(\log^{1/2} n) \quad (5.23)$$

for every $\omega \in \Omega_n$ (and $M_n \geq M^*$ for any ω), at a low cost. To do this, $Od(n)$ probes the sample path ω at positions $(jl)_{j=1, \dots, n/l}$ where $l = \lceil n^{1/2} \log n \rceil$, then it scans the vicinity of the "large" S_{jl} .

In the second stage, $Od(n)$ covers the sample path from left to right (independently of the probes of the first stage), as follows: assume that, at step t , $Od(n)$ just probed m . At step $t+1$:

- . If S_m is close to M^* ($M^* - S_m \leq n^{1/6}$), $Od(n)$ probes S_{m+1}
- . If S_m is far away from M^* ($M^* - S_m \geq n^{1/6}$), $Od(n)$ probes S_{m+k} where

$$k = 2(M^* - S_m) - 10c(M^* - S_m)^{1/2}(\log n)^{1/2}. \quad (5.24)$$

According to [39], this value of k ensures that $S_{m+k} + S_m + k \leq 2M^*$ if ω is in $\Omega_d(n)$, and that the random walk does not exceed M^* on $[m, m+k]$.

Let $a(n) = (m_1, \dots, m_{R(n)})$ be the sequence of positions probed during the second stage. According to [39], the first stage needs $O(\frac{\sqrt{n}}{\log n})$ probes, in the average, so that the proof of Lemma 3.2 reduces to:

$$\frac{R(n) - V_n}{\sqrt{n}} \xrightarrow{L_1} 0,$$

that is, it reduces to prove that $a(n)$ satisfies the hypothesis of Theorem 5.2. For α , $G(n)$ and Ω_n , we keep to the same definitions as in the preceding proof, but, due to the definition of $Od(n)$, we take $H(n) = 4n^{1/6}$. Hypothesis (5.18) is satisfied for

$$Z_k^{(n)} = 2(M_n - S_k + 1),$$

so that this time U_n is equal to V_n ; checking (5.17), we get $E(\tilde{U}_n) = O(n^{1/6})$. We have

$$\begin{aligned} \#\{i | Z_{m_i}^{(n)} \leq H(n)\} &\leq \#\{k | Z_k^{(n)} \leq H(n)\} \\ &\leq \sum_{q=0}^{H(n)} D_q, \end{aligned}$$

thus, due to Proposition 5.3

$$E(\#\{i | Z_{m_i}^{(n)} \leq H(n)\} I_{\Omega_c(n)}) = O(H(n)^2).$$

Using (5.23) and (5.24), one can see that for any $\omega \in \Omega_n$, and provided that $Z_{m_i}^{(n)} \geq 3n^{1/6}$ (i.e. we choose $\delta = 1/3$ in Theorem 5.2),

$$\begin{aligned} \frac{w_i}{Z_{m_i}^{(n)}} &\geq \frac{Z_{m_i}^{(n)} - 2 - 2cn^{1/8} \log^{1/2} n - 10c\sqrt{Z_{m_i}^{(n)} \log n}}{Z_{m_i}^{(n)}} \\ &\geq 1 - 10cn^{-1/12} \sqrt{\log n}, \end{aligned}$$

for n large enough (else let $\varepsilon_n = 1$), independent of $\omega \in \Omega_n$. \square

6 Proof of Theorem 3.3

Set

$$I_n = \frac{V_n}{\sqrt{n}} = \frac{1}{2\sqrt{n}} \sum_{k=1}^n \frac{1}{M_n - S_k + 1};$$

we shall now prove that I_n converges in distribution to I .

The operator T defined by

$$T(f) = \int_0^1 \frac{dx}{\max f - f(x)}$$

does not satisfy the continuity hypothesis for Donsker's invariance principle, so we exhibit a coupling of I_n and I through a Skorohod embedding, and prove that $I - I_n$ converges to 0 in probability. Let $(\mathcal{B}_t)_{t \geq 0}$ denote a standard Brownian motion and let $(S_k^{(n)})_{0 \leq k \leq n}$ be the embedded random walk associated to \mathcal{B} through the standard construction due to Knight (1962) (see Figures 6.1 and 6.2): consider the successive times $(\tau_k^{(n)})_{0 \leq k \leq n}$, defined

by $\tau_0^{(n)} = 0$ and for $k \geq 1$, $\tau_k^{(n)} = \inf \{t > \tau_{k-1}^{(n)}, |\mathcal{B}_t - \mathcal{B}_{\tau_{k-1}^{(n)}}| = n^{-1/2}\}$ where the Brownian motion has moved by $n^{-1/2}$ and set:

$$S_k^{(n)} = \sqrt{n}\mathcal{B}(\tau_k^{(n)}).$$

The sample path $(S_k^{(n)})_{k=0, \dots, n}$ has the same law as $\omega = (S_k(\omega))_{k=0, \dots, n}$ introduced in the beginning of the paper. Note that for every $k > 0$, $n(\tau_{k+1}^{(n)} - \tau_k^{(n)})$ has the same law as Σ_1^\sharp , the hitting time of 1 for a reflected Brownian motion, already defined in Section 4.

Consider the piecewise constant process $\beta_t^{(n)}$ equal to $\mathcal{B}(\tau_{k-1}^{(n)}) = n^{-1/2}S_{k-1}^{(n)}$ for t in the interval $]\tau_{k-1}^{(n)}, \tau_k^{(n)}]$; hence, by construction:

$$|\mathcal{B}(t) - \beta_t^{(n)}| \leq \frac{1}{\sqrt{n}}$$

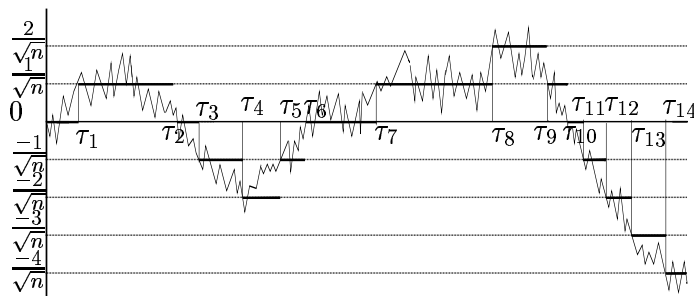


Figure 6.1 :The Skorohod embedding

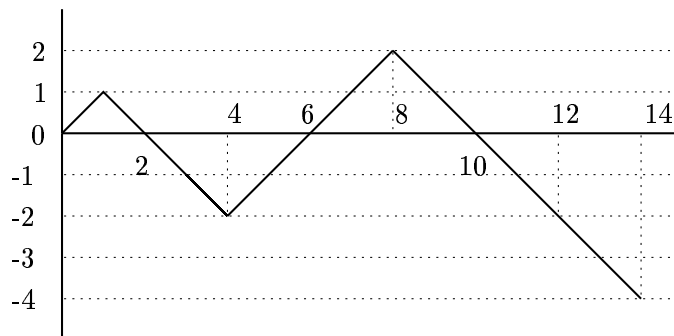


Figure 6.2 : The embedded random walk

Let $M_n^{(\beta)}$ denote the maximum of $(\beta_t^{(n)})_{0 \leq t \leq n}$, and set (for $0 < \varepsilon < 1/2$)

$$\begin{aligned} \tilde{I}_n &= \frac{1}{2\sqrt{n}} \sum_{k=1}^n \frac{1}{M_n^{(n)} - S_k + n^{-\varepsilon+1/2}}, \\ W_1 &= \int_0^{\tau_n^{(n)}} \frac{dt}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} - 2\tilde{I}_n \\ &= \sqrt{n} \sum_{k=1}^n \frac{\tau_k^{(n)} - \tau_{k-1}^{(n)}}{M_n^{(n)} - S_k^{(n)} + n^{-\varepsilon+1/2}} - 2\tilde{I}_n \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{n(\tau_k^{(n)} - \tau_{k-1}^{(n)}) - 1}{M_n^{(n)} - S_k^{(n)} + n^{-\varepsilon+1/2}}, \\ W_2 &= 2I - \int_0^{\tau_n^{(n)}} \frac{dt}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} \\ &= \int_0^1 \frac{dt}{M_1^{\mathcal{B}} - \mathcal{B}_t} - \int_0^{\tau_n^{(n)}} \frac{dt}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}}. \end{aligned}$$

We are to prove that $\tilde{I}_n - I_n$, W_1 and W_2 converge to 0 in probability. First

$$\begin{aligned} E(I_n - \tilde{I}_n) &\leq n^{-\varepsilon} E\left(\sum_{k=1}^n \frac{1}{(M_n^{(n)} - S_k^{(n)} + 1)^2}\right) \\ &= n^{-\varepsilon} E\left(\sum_{q=0}^n \frac{D_q(n)}{(q+1)^2}\right) \\ &= O(n^{-\varepsilon} \log n), \end{aligned}$$

the last line being due to Proposition 5.3. Since $(\tau_k^{(n)})_{0 \leq k \leq n}$ and $(\mathcal{B}(\tau_k^{(n)}))_{0 \leq k \leq n}$ are independent, and the r.v.s $n(\tau_k^{(n)} - \tau_{k-1}^{(n)}) - 1$ are independent, centered, with variance 1, we obtain

$$\begin{aligned} E(W_1^2) &= \frac{1}{n} E\left(\sum_{k=1}^n \frac{1}{(M_n^{(n)} - S_k^{(n)} + n^{-\varepsilon+1/2})^2}\right) \\ &\leq \frac{1}{n} E\left(\sum_{q=0}^n \frac{D_q(n)}{(q+1)^2}\right) \\ &= O\left(\frac{\log n}{n}\right), \end{aligned}$$

so that $W_1 \xrightarrow{\mathcal{L}_2} 0$. Next, we consider W_2 . Set $\tilde{\tau}_n = \tau_n^{(n)}$. Noting that $\tilde{\tau}_n$ has the same law as the average of n independent copies of Σ_1^\sharp , and that Σ_1^\sharp belongs to any \mathcal{L}_p (see for instance

Lépingle (1981)), it is routine to deduce that

$$\|\tilde{\tau}_n - 1\|_{\mathcal{L}_p} = O\left(\frac{1}{\sqrt{n}}\right). \quad (6.25)$$

We shall also need the following results:

Proposition 6.1 *For any positive η ,*

$$n^{-\eta+1/4} |M_n^{(\beta)} - M_1^{\mathcal{B}}| \xrightarrow{\text{proba}} 0$$

Proposition 6.2 *For $0 < \alpha \leq \frac{1}{2}$,*

$$E\left(\int_0^{\tilde{\tau}_n} \frac{I_{\{M_n^{(\beta)} - \beta_t \leq n^{-\alpha}\}}}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} dt\right) = O(n^{-\alpha})$$

Proposition 6.3 *For $\alpha > 0$,*

$$E\left(\int_0^1 \frac{I_{\{M_1^{\mathcal{B}} - \mathcal{B}_t \leq n^{-\alpha}\}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt\right) = O(n^{-\alpha})$$

Proofs of Propositions 6.1, 6.2 and 6.3 are delayed to the end of the Section. Now:

$$|W_2| \leq A + B + C$$

where

$$\begin{aligned} A &= \int_0^{1 \wedge \tilde{\tau}_n} \frac{|\beta_t^{(n)} - \mathcal{B}_t| + |M_n^{(\beta)} - M_1^{\mathcal{B}}| + n^{-\varepsilon}}{(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})(M_1^{\mathcal{B}} - \mathcal{B}_t)} dt \\ &\leq (2n^{-\varepsilon} + |M_n^{(\beta)} - M_1^{\mathcal{B}}|) \int_0^{1 \wedge \tilde{\tau}_n} \frac{1}{(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})(M_1^{\mathcal{B}} - \mathcal{B}_t)} dt \\ B &= I_{\tilde{\tau}_n \leq 1} \int_{\tilde{\tau}_n}^1 \frac{dt}{M_1^{\mathcal{B}} - \mathcal{B}_t} \\ C &= I_{\tilde{\tau}_n \geq 1} \int_1^{\tilde{\tau}_n} \frac{dt}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} \end{aligned}$$

By routine application of Hölder's inequality, we have successively:

$$B \leq (1 - \tilde{\tau}_n)^{1/p} \left(\int_{\tilde{\tau}_n}^1 \frac{dt}{(M_1^{\mathcal{B}} - \mathcal{B}_t)^q} \right)^{1/q} I_{\tilde{\tau}_n \leq 1}$$

and:

$$E[B] \leq E[|1 - \tau_n|]^{1/p} E\left[\int_0^1 \frac{dt}{(M_1^{\mathcal{B}} - \mathcal{B}_t)^q}\right]^{1/q},$$

the last term in both equations being finite provided that $q < 2$, so, due to (6.25), for any positive η , we have

$$n^{-\eta+1/4} B \xrightarrow{L_1} 0. \quad (6.26)$$

We also have

$$C \leq n^\varepsilon |\tilde{\tau}_n - 1|$$

which converges to 0 in any L_p thanks to (6.25). Finally, let

$$A \leq n^{-\varepsilon} \left(2 + n^\varepsilon |M_n^{(\beta)} - M_1^{\mathcal{B}}|\right) (A_1 + A_2) \quad (6.27)$$

where

$$\begin{aligned} A_1 &= \int_0^{\tilde{\tau}_n} \frac{dt}{(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})(M_1^{\mathcal{B}} - \mathcal{B}_t)} I_{\tilde{\tau}_n \geq 1} \\ A_2 &= \int_0^{\tilde{\tau}_n} \frac{dt}{(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})(M_1^{\mathcal{B}} - \mathcal{B}_t)} I_{\tilde{\tau}_n \leq 1} \end{aligned}$$

Note that the second term on the right hand side of (6.27) converges in probability to 2, due to Proposition 6.1, provided that $\varepsilon = 1/5$, for instance.

We have, for $\alpha < 1/2$,

$$\begin{aligned} A_1 &\leq n^\alpha \int_0^1 \frac{I_{\tilde{\tau}_n \geq 1} \text{ and } M_1^{\mathcal{B}} - \mathcal{B}_t \geq n^{-\alpha}}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} dt + n^\varepsilon \int_0^1 \frac{I_{M_1^{\mathcal{B}} - \mathcal{B}_t \leq n^{-\alpha}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt \\ &\leq \frac{n^{2\alpha}}{2} + n^\varepsilon \int_0^1 \frac{I_{M_1^{\mathcal{B}} - \mathcal{B}_t \leq n^{-\alpha}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt. \end{aligned}$$

Note that the second inequality holds since

$$M_n^{(\beta)} - \beta_t^{(n)} \geq M_1^{\mathcal{B}} - \mathcal{B}_t - \frac{2}{\sqrt{n}},$$

when $\tilde{\tau}_n \geq 1$. Thanks to Proposition 6.3,

$$E(A_1) \leq \frac{n^{2\alpha}}{2} + O(n^{\varepsilon-\alpha}).$$

Finally, for $\alpha = 1/20$ (for instance), we obtain:

$$n^{-\varepsilon} \left(2 + n^\varepsilon |M_n^{(\beta)} - M_1^{\mathcal{B}}|\right) A_1 \xrightarrow{\text{proba}} 0.$$

Similarly

$$\begin{aligned}
A_2 &\leq n^\alpha \int_0^{\tilde{\tau}_n} \frac{I_{M_n^{(\beta)} - \beta_t^{(n)} \geq n^{-\alpha}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt + \int_0^{\tilde{\tau}_n} \frac{I_{M_n^{(\beta)} - \beta_t^{(n)} \leq n^{-\alpha}}}{(M_1^{\mathcal{B}} - \mathcal{B}_t)(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})} dt \\
&\leq \frac{n^{2\alpha}}{2} + \int_0^{\tilde{\tau}_n} \frac{I_{M_n^{(\beta)} - \beta_t^{(n)} \leq n^{-\alpha}}}{(M_1^{\mathcal{B}} - \mathcal{B}_t)(M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon})} dt \\
&\leq \frac{n^{2\alpha}}{2} + n^\varepsilon \int_0^{\tilde{\tau}_n} \frac{I_{M_1^{\mathcal{B}} - \mathcal{B}_t \leq n^{-\varepsilon}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt + n^\varepsilon \int_0^{\tilde{\tau}_n} \frac{I_{M_n^{(\beta)} - \beta_t^{(n)} \leq n^{-\alpha}, M_1^{\mathcal{B}} - \mathcal{B}_t \geq n^{-\varepsilon}}}{M_n^{(\beta)} - \beta_t^{(n)} + n^{-\varepsilon}} dt,
\end{aligned}$$

where the second inequality holds since

$$M_1^{\mathcal{B}} - \mathcal{B}_t \geq M_n^{(\beta)} - \beta_t^{(n)} - n^{-1/2}$$

when $\tilde{\tau}_n \leq 1$. Owing to Propositions 6.2 and 6.3, we obtain

$$E(A_2) \leq \frac{n^{2\alpha}}{2} + O(1) + O(n^{\varepsilon-\alpha}).$$

Finally:

$$n^{-\varepsilon} \left(2 + n^\varepsilon |M_n^{(\beta)} - M_1^{\mathcal{B}}| \right) A_2 \xrightarrow{\text{proba}} 0. \quad \square$$

Proof of Proposition 6.1: Let η and a be positive. Assuming first that $\tilde{\tau}_n > 1$, we obtain:

$$|M_1^{\mathcal{B}} - M_n^{(\beta)}| \leq \frac{1}{\sqrt{n}} + M_{\tilde{\tau}_n}^{\mathcal{B}} - M_1^{\mathcal{B}}$$

Hence, for any positive $\phi(n)$,

$$\Pr(n^{-\eta+1/4} |M_n^{(\beta)} - M_1^{\mathcal{B}}| I_{\tilde{\tau}_n > 1} > a) \leq A_n + B_n,$$

where

$$\begin{aligned}
A_n &= \Pr(\tilde{\tau}_n - 1 > \phi(n)), \\
B_n &= \Pr(n^{-\eta+1/4} \left(\frac{1}{\sqrt{n}} + M_{1+\phi(n)}^{\mathcal{B}} - M_1^{\mathcal{B}} \right) > a) \\
&\leq \Pr(n^{-\eta+1/4} \left(\frac{1}{\sqrt{n}} + M_{\phi(n)}^{\mathcal{B}} \right) > a) \\
&= \Pr(n^{-\eta+1/4} \left(\frac{1}{\sqrt{n}} + \sqrt{\phi(n)} |N| \right) > a),
\end{aligned}$$

for N a standard Gaussian r.v., so the choice $\phi(n) = n^{\eta-1/2}$ gives the desired result. The case $\tilde{\tau}_n > 1$ can be proven with the same method. \square

Proof of Proposition 6.2 . Using Proposition 5.3, we have:

$$\begin{aligned} E\left(\int_0^{\tilde{\tau}_n} \frac{I_{\{M_n^{(\beta)} - \beta_t^{(n)} \leq n^{-\alpha}\}}}{M_n^{(\beta)} - \beta_t^{(n)} + \frac{1}{\sqrt{n}}} dt\right) &= \frac{1}{\sqrt{n}} E\left(\sum_{k=0}^n \frac{I_{\{M_n - S_k \leq n^{-\alpha+1/2}\}}}{M_n - S_k + 1}\right) \\ &= \frac{1}{\sqrt{n}} E\left(\sum_{j=0}^{n^{-\alpha+1/2}} \frac{D_q(n)}{q+1}\right) \\ &\leq \frac{C}{\sqrt{n}} n^{-\alpha+1/2} \quad \square \end{aligned}$$

Proof of Proposition 6.3: Using the Denisov decomposition (4.5), as in Section 4, we obtain:

$$\int_0^1 \frac{I_{\{M_1^{\mathcal{B}} - \mathcal{B}_t \leq n^{-\alpha}\}}}{M_1^{\mathcal{B}} - \mathcal{B}_t} dt = \sqrt{\theta} \int_0^1 \frac{I_{\{\mathcal{m}_s^- \sqrt{\theta} \leq n^{-\alpha}\}}}{\mathcal{m}_s^-} ds + \sqrt{1-\theta} \int_0^1 \frac{I_{\{\mathcal{m}_s^+ \sqrt{1-\theta} \leq n^{-\alpha}\}}}{\mathcal{m}_s^+} ds,$$

the two last terms, say A and B , being identically distributed. We have

$$A = \sqrt{\theta} \int_{0 \leq x \leq n^{-\alpha} \theta^{-1/2}} \frac{L_x}{x} dx$$

where L_x denotes the local time at level x of the Brownian meander. On the other hand, according to Durrett & Iglehart (1977):

$$\begin{aligned} E(L_x) &= 2 \int_0^x \exp\left(-\frac{(x+y)^2}{2}\right) dy \\ &= \sqrt{2\pi} \Pr(x \leq |N| \leq 2x), \end{aligned}$$

where N is standard Gaussian. Finally:

$$\begin{aligned} E(A) &= C_1 \int_0^1 \int_{0 \leq x \sqrt{y} \leq n^{-\alpha}} \Pr(x \leq |N| \leq 2x) \frac{dx}{x} \frac{dy}{\sqrt{1-y}} \\ &\leq C_2 \int_0^1 \int_{0 \leq x \sqrt{y} \leq n^{-\alpha}} e^{-\frac{x^2}{2}} dx \frac{dy}{\sqrt{1-y}} \\ &= C_2 \left(2 \int_0^{n^{-\alpha}} e^{-\frac{x^2}{2}} dx + \int_{n^{-\alpha}}^{+\infty} \int_0^{n^{-2\alpha} x^{-2}} \frac{dy}{\sqrt{1-y}} e^{-\frac{x^2}{2}} dx \right) \\ &\leq 2C_2 \left(n^{-\alpha} + \int_{n^{-\alpha}}^{+\infty} (1 - \sqrt{1 - n^{-2\alpha} x^{-2}}) e^{-\frac{x^2}{2}} dx \right) \end{aligned}$$

and:

$$\begin{aligned} \int_{n^{-\alpha}}^{+\infty} (1 - \sqrt{1 - n^{-2\alpha} x^{-2}}) e^{-\frac{x^2}{2}} dx &= n^{-\alpha} \int_1^{+\infty} (1 - \sqrt{1 - u^{-2}}) e^{-n^{-2\alpha} u^2} du \\ &\leq n^{-\alpha} \int_1^{+\infty} (1 - \sqrt{1 - u^{-2}}) du. \quad \square \end{aligned}$$

Acknowledgment

We thank Philippe Flajolet for calling our attention to the problem of the limit distribution of $T_{Od(n)}$.

References

- [1] Aho, A.V., Hopcroft, J.E., Ullman, J.D. (1974). *The design and analysis of computer algorithm*, Addison-Wesley.
- [2] *Fourth European Symposium on Algorithms*, (1998) Papers from the symposium held in Barcelona, September 1996, *Algorithmica* **21**(1), 1-154.
- [3] Alonso, L., Reingold, E.M., Schott, R. (1997). *The average-case complexity of determining the majority*. *SIAM J. Comput.* 26, No.1, 1-14.
- [4] Asmussen, S. (1987). *Applied probability and queues*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester.
- [5] Avriel, M., Wilde, D.J. (1966). *Optimal search for a maximum with sequences of simultaneous function evaluations*, *Manage. Sci.*, Ser. A 12, 722-731.
- [6] Baccelli, F., Brémaud, P. (1994). *Elements of queuing theory. Palm-martingale calculus and stochastic recurrences*. *Applications of Mathematics*, **26** Springer-Verlag, Berlin.
- [7] Biane, P., Pitman, J., Yor, M. (1999). *Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions*, preprint...
- [8] Biane, P., Yor, M. (1987). *Valeurs principales associées aux temps locaux browniens*, *Bull. Sci. Maths* **111**, 23-101.
- [9] Biane, P., Yor, M. (1988). *Quelques précisions sur le méandre brownien*, *Bull. Sci. Maths* **112**, 101-109.
- [10] Bollobás, B. (1985). *Random Graphs*, Academic Press p.12.
- [11] Carmona, P., Petit, F., Pitman, J., Yor, M. (1998). *On the laws of homogeneous functionals of the Brownian bridge*, Prépub. du Labo. Proba. Univ. Paris VI, to appear in *Studia Math. Hung.*
- [12] Chassaing, P. (1997). *Determining the majority: The biased case*, *Ann. Appl. Probab.* **7**, No.2, 523-544.
- [13] Chassaing, P. (1999) *How many questions are needed to find out the maximum of a Random Walk*, to appear in *Stoch. Proc. and Applns.*
- [14] Chassaing, P., Marckert, J.F., (1999) *Parking functions, empirical processes and the width of rooted labeled trees*, preprint Elie Cartan, Université Nancy I.

- [15] Chassaing, P., Marckert, J.F., Yor, M. *On the harmonic mean of depths of nodes of a random tree*, Preprint.
- [16] Chung, K.L. (1976). *Excursions in Brownian motion*, Ark. för Math., **14**, 155-177.
- [17] Cunto, W., Munro, J.I. (1989). *Average case selection*, J. Assoc. Comput. Mach. **36**, No.2, 270-279.
- [18] Denisov, I.V. (1983). *A random walk and a Wiener process near a maximum*, Th. Prob. App. **28** , 821-824.
- [19] Durrett, R., Iglehart, D.L. (1977), *Functionals of Brownian meander and Brownian excursion*, Ann. Probab. **5** , 130-135.
- [20] Flajolet, P., Gao, Z.C., Odlyzko, A.M., Richmond, B. (1993). *The distribution of heights of binary trees and other simple trees*, Combin. Probab. Comput. 2, no. 2, 145-156.
- [21] Flajolet, P., Odlyzko, A. (1982). *The average height of binary trees and other simple trees*, J. Comput. System Sci. **25**, no. 2, 171-213.
- [22] Flajolet, P., Salvy, B. (1998). *Euler sums and contour integral representations*, Experiment. Math. 7:1 15-35.
- [23] Flajolet, P., Poblete, P., Viola, A. (1998). *On the Analysis of Linear Probing Hashing*, Algoritmica, in press.
- [24] Grübel, R., Rösler, U. (1996). *Asymptotic distribution theory for Hoare's selection algorithm*, Adv. in Appl. Probab. **28**, no. 1, 252-269.
- [25] Hajek, B. (1987). *Locating the maximum of a simple random sequence by sequential search*, IEEE Trans. Inf. Theory IT-33, 877-881.
- [26] Hwang, H.K. (1997). *A constant arising from the analysis of algorithms for determining the maximum of a random walk*, Random Struct. Alg. **10**, 333-335.
- [27] Kao, M.Y., Reif, J.H., Tate, S.R., (1993). *Searching in an unknown environment: An optimal randomized algorithm for the cow-path problem*, Proc. 4th ACM-SIAM Symp. Discrete Math., **4** , 441-447.
- [28] Kennedy, D.P. (1976). *The distribution of the maximum Brownian excursion*, J. Appl. Probab. **13**, 371-376.
- [29] Kleinrock, L. (1975). *Queuing systems. Vol. I: Theory*, John Wiley & Sons.
- [30] Kleinrock, L. (1976). *Queuing systems. Vol. II: Computer applications*, John Wiley & Sons.
- [31] Knight, F.B. (1962). *On the random walk and Brownian motion*, Trans. Amer. math. Soc. **103**, 218-228.

- [32] Knight, F.B. (1969). *Brownian local times and taboo processes*, Trans. Am. Math. Soc. 143, 173-185.
- [33] Knuth, D.E. (1997). *The art of computer programming. Vol. 3: sorting and searching*, 2nd ed., Addison-Wesley.
- [34] Kuiper, N. H. (1960). *Tests concerning random points on a circle*, Indag. Math. **22**, 38-47.
- [35] Lépingle, D. (1981). *Une remarque sur les lois de certains temps d'atteinte*, Sémin. Pro. XV, Uni. Str. 1979/80, Lect. Notes Math. **850**, Springer, 669-670.
- [36] Louchard, G. (1984). *Kac's formula, Levy's local time and Brownian excursion*, J. Appl. Prob. **21**, 479-499.
- [37] Louchard, G., Kenyon, C., Schott, R. (1997). *Data structures' maxima*, SIAM J. Comput. **26**, no 4, 1006-1042.
- [38] Marckert, J.F. (1997). *An average case quasi-optimal algorithm for the search of zeros of a random walk*, Preprint Elie Cartan n.44.
- [39] Odlyzko, A.M. (1995). *Search for the maximum of a random walk*, Random Structures and Algorithms, **6**, 275-295.
- [40] Pitman, J.W., Yor, M. (1996). *Quelques identités en loi pour les processus de Bessel*, Hommage à P. A. Meyer et J. Neveu. Paris: Soc. Math. Fr., Astérisque. **236**, 249-276.
- [41] Pitman, J.W., Yor, M. (1998). *Path decomposition of a Brownian bridge related to the ratio of its maximum and amplitude*, Prépub. n 461 du Labo. Proba. Univ. Paris VI (Sept.1998), to appear in Studia Math. Hung. (1999).
- [42] Plaskota, L. (1996.) *Noisy information and computational complexity*, Cambridge University Press, Cambridge
- [43] Rösler, U. (1991). *A limit theorem for "Quicksort"*, RAIRO Inform. Thor. Appl. **25**, no. 1, 85-100.
- [44] Revuz, D., Yor, M. (1994). *Continuous Martingales and Brownian Motion*, Springer- 2nd edition.
- [45] Sedgewick, R., Flajolet, P. (1996). *An introduction to the analysis of algorithms*, Addison-Wesley.
- [46] R. Sitaramachandra Rao. (1987), *A formula of S. Ramanujan*, J. Number Theory 25:1, 1-19.
- [47] Smith, L., Diaconis, P (1988). *Honest Bernoulli Excursions*, J. Appl.Prob. **25**, 464-477.
- [48] Stoyan, D. (1983). *Comparison methods for queues and other stochastic models*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester.

- [49] Spitzer, F. (1957). *The Wiener-Hopf equation whose kernel is a probability density*, Duke math. J. **24**, 327-343.
- [50] Takács, L. (1991). *A Bernoulli excursion and its various applications*, Adv. in Appl. Probab. **23**, **3**, 557-585.
- [51] Traub, J.F., Wasilkowski, G.W., Wozniakowski, H. (1988). *Information-based complexity*, Comp. Sci. Scient. Comp. Boston, MA: Academic Press, Inc.
- [52] Vervaat, W. (1979). *A relation between Brownian bridge and Brownian excursion*, Ann. Probab. **7**, no. 1, 143-149.
- [53] Williams, D. (1974). *Path decomposition and continuity of local time for one-dimensional diffusions. I*. Proc. London math. Soc., III. Ser. **28**, 738-768.
- [54] Williams, D. (1979). *Diffusions, Markov processes, and martingales. Vol. 1. Foundations*. Wiley, Chichester, New York.
- [55] Yakowitz, S., Lugosi, E. (1990). *Random search in the presence of noise, with application to machine learning*, SIAM J. Sci. Stat. Comput. **11**, No.4, 702-712.
- [56] Yor, M. (1997). *Some aspects of Brownian motion. Part II. Some recent martingale problems*, Lectures in Mathematics, ETH Zürich, Birkhäuser Verlag, Basel.

CHAPITRE III

Stochastic properties of two algorithms searching the zeros of simple random walks, and applications

P. Chassaing¹, J.F. Marckert¹, M. Yor².

The main object of this paper is to prove the quasi optimality for the stochastic order of an algorithm searching the zeros of simple symmetric random walk. The study of the behavior of an other algorithm allows us to give a simple computation of the limit law of the couple height-width in binary trees of size n and the limit law of the harmonic mean of the depth of nodes in size n general trees.

1 Introduction

Consider $\omega = (S_k(\omega))_{k=0, \dots, n}$ a symmetric random walk. That is, $S_0(\omega) = 0$ and for any $0 \leq k \leq n-1$, $S_{k+1}(\omega) = S_k(\omega) \pm 1$, the 2^n such trajectories being equiprobable. The knowledge of one of the S_l 's for a l of our choice costs 1. The question is: How much must we pay to find each k such that $S_k = 0$?

The cost of an algorithm is naturally identified with the number of probes it uses. For example, the simplest and worst algorithm consists in probing each position l , for $1 \leq l \leq n$; its cost is n .

The average case analysis of algorithms for the search of all the zeros has been done: the double jump algorithm, $dj(n)$ (described in the next subsection) is already known to be quasi optimal in the average case (see Marckert [21]).

The first aim of this work is to study the limit law of the cost of algorithm $dj(n)$ and to prove its quasi-optimality for the stochastic order. The analysis of the simple jump algorithm, $sj(n)$, leads us to study again the properties of the relations between trees and Bernoulli excursions and thus, to derive asymptotic results on trees involving Brownian excursion.

1.1 The simple and double jump algorithms

The “simple-jump” algorithm $sj(n)$:

At step 1, we first probe position 2. Then we cover the walk from left to right. Assume that at step T , we probed position m . At step $T+1$:

- If $S_m = 0$ we probe $m+2$.
- If $S_m \neq 0$ we probe $m + |S_m|$,

and we proceed like this until $m + |S_m| > n$. On the following figure, we have blackened the probed points.

¹Laboratoire de Mathématiques, Institut Elie Cartan
BP 239 54 506 Vandoeuvre les Nancy Cedex

²Université Paris VI, Laboratoire de Probabilités
Tour 56 - 4, place Jussieu - 75252 Paris Cedex 05

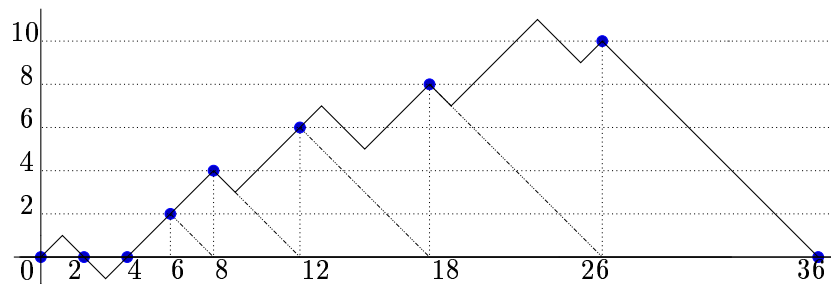


Figure 1.1 : simple jump algorithm

Theorem 1.1 Let $T_{sj(n)}$ be the cost of $sj(n)$. We have

$$\mathbb{E}(T_{sj(n)}) \sim \sqrt{\frac{2n}{\pi}} \log n.$$

The “double-jump” algorithm $dj(n)$:

At step 1, we probe position 2. Then we cover the sample path from left to right. Assume that at step T we probed position m . Two cases arise:

- a) $|S_m| \leq \log^{1+\varepsilon} n$. (ε being any fixed real number in the interval $]0,1[$)
- b) $|S_m| > \log^{1+\varepsilon} n$.

In Case a), at step $T+1$ we apply the simple-jump algorithm, that is, we probe position $m+2$ or $m+|S_m|$ depending on $S_m = 0$ or not.

In Case b), at step $T+1$ we probe $m+K$ in which

$$K = \lfloor 2|S_m| - c\sqrt{2 \log n |S_m|} \rfloor.$$

Here is the reason: assume that $S_m > 0$. Due to the bounded variation of ω , the minimum of the sample path on the interval $[m, m+K]$ is not smaller than

$$B_m = \frac{S_m + S_{m+K} - K}{2}.$$

$B_m > 0$ ensures that ω does not cancel on the interval $[m, m+K]$ and that we have found every zero on the left of $m+K$. The algorithm is ready for the next step.

If $B_m \leq 0$, ω may cancel in $[m, m+K]$ and then, we probe every position starting from m . Due to the moderate variation properties of the simple symmetric random walks (see [21] and Subsection 2.2), this situation occurs with a very small probability (for c large enough).

Stochastic properties of $dj(n)$

Let A_n be the set of all algorithms that find every zero of ω . For an algorithm a in A_n , let $T_a(\omega)$ be the number of probes it uses. Marckert [21] gives the lower bound of the average cost of T_a and proves that $dj(n)$ is quasi optimal in the average case:

Theorem 1.2

$$\bullet \min_{a \in A_n} \mathbb{E}(T_a) = \frac{1}{\sqrt{2\pi}} \sqrt{n} \log n + o(\sqrt{n} \log n)$$

and

$$\bullet \mathbb{E}(T_{dj(n)}) = \frac{1}{\sqrt{2\pi}} \sqrt{n} \log n + o(\sqrt{n} \log n).$$

“Quasi” means that $\mathbb{E}(T_{dj(n)})$ is equivalent to the lower bound. For an algorithm a , let us consider the tails of the distribution function of T_a :

$$\mathbb{P}(T_a \geq x).$$

This is the probability that algorithm a uses more than x probes. Theorem 1.2 suggests that the suitable normalization for T_a is $\sqrt{n} \log n$. So, we define

$$\Phi_n(x) = \min \left\{ \mathbb{P} \left(\frac{T_a}{\sqrt{n} \log n} \geq x \right) \mid a \in A_n \right\}.$$

$\Phi_n(x)$ is also the tail of a probability measure, but there does not necessarily exist an algorithm α_n such that for every $x \geq 0$,

$$\Phi_n(x) = \mathbb{P} \left(\frac{T_{\alpha_n}}{\sqrt{n} \log n} \geq x \right). \quad (1.1)$$

If such an algorithm α_n exists for every x , one will say that α_n is optimal for the stochastic order in A_n . This kind of optimality gives more guarantees than the only optimality in the average case (see discussion in Chapter 2 of Chassaing et al. [9]). Actually, $dj(n)$ satisfies (1.1) asymptotically. The two following theorems give the stochastic properties of $dj(n)$.

Theorem 1.3

$$\lim_n \mathbb{P} \left(\frac{T_{dj(n)}}{\sqrt{n} \log n} \geq x \right) = \mathbb{P} \left(\frac{|N|}{2} \geq x \right)$$

where N denotes a centered Gaussian variable with variance 1, and

Theorem 1.4

$$\lim_n \Phi_n(x) = \mathbb{P} \left(\frac{|N|}{2} \geq x \right).$$

We say that $dj(n)$ is *quasi optimal for the stochastic order*.

We have seen (Theorem 1.1 and Theorem 1.2) that $sj(n)$ is twice as expensive as $dj(n)$ in average. This fact is due to the assumptions made on the distribution of the random walks. Although, one can imagine that $sj(n)$ is very suitable for searching the root practically when it is possible to obtain some bounds on the derivatives (which often arises).

For related reading, one can find in Odlyzko [22] the case of the search of the maximum of symmetric random walks; Odlyzko provides a quasi-optimal algorithm in the average case; Chassaing [6] studies the search of the maximum of a non-symmetric random walk. Chassaing, Marckert & Yor [9] prove that Odlyzko’s algorithm is quasi-optimal for the stochastic order and characterize the limit law.

1.2 Trees and Brownian excursions

Denote by $T_B^{(n)}$ the set of *binary trees* with n leaves ($2n - 1$ nodes). For a tree $t \in T_B^{(n)}$, the profile of t is the sequence $(Z_i(t))_{i \geq 0}$ where $Z_i(t)$ is the number of nodes of t at distance i from the root ($Z_0(t) = 1, Z_1(t) = 2$). The width $W_B^{(n)}(t)$ of t is defined to be the maximum of the profile:

$$W_B^{(n)}(t) = \max_i \{Z_i(t)\}.$$

Denote by $H_B^{(n)}(t)$ the height of t . The limit law of the height (see [1, 2, 5, 15, 23, 29]) and of the width (see [8, 31]) are well known:

$$\frac{H_B^{(n)}}{\sqrt{2n}} \xrightarrow[n \rightarrow +\infty]{law} V, \quad (1.2)$$

$$\frac{W_B^{(n)}}{\sqrt{2n}} \xrightarrow[n \rightarrow +\infty]{law} \frac{V}{2} \quad (1.3)$$

where V has cumulative function of theta type:

$$F_V(x) = \mathbb{P}(V \leq x) = \sum_{-\infty < k < +\infty} (1 - k^2 x^2) \exp\left(-\frac{k^2 x^2}{2}\right). \quad (1.4)$$

Set

$$(H, W) = \left(\int_0^1 \frac{1}{e(s)} ds, \max_{0 \leq s \leq 1} e(s) \right)$$

where $(e(s))_{0 \leq s \leq 1}$ is a standard normalized Brownian excursion (see Subsection 3.1.1).

Theorem 1.5

$$\left(\frac{H_B^{(n)}}{\sqrt{2n}}, \frac{W_B^{(n)}}{\sqrt{2n}} \right) \xrightarrow[n \rightarrow +\infty]{law} (H, W).$$

The dependence between H and W reflects the dependence between $H_B^{(n)}$ and $W_B^{(n)}$.

Donati-Martin[12] has studied the couple $(\int_0^1 \frac{1}{e(s)} ds, \max_{0 \leq s \leq 1} e(s))$. Here are some of her results which characterize the couple (H, W) :

- Let $\lambda \geq 0, \alpha \geq 0$ then

$$\mathbb{E}\left(W \exp\left(-\frac{\lambda^2}{2W^2} - \frac{\alpha^2 H}{2W}\right)\right) = \sqrt{\frac{\pi}{2}} A^2(\lambda, \alpha)$$

where

$$A(\lambda, \alpha) = \frac{\exp(\lambda)}{\Phi(1 + \alpha^2/(2\lambda), 2, 2\lambda)}$$

and Φ denotes the confluent hypergeometric function :

$$\Phi(\alpha, \gamma, z) = \sum_{k=0}^{+\infty} \frac{(\alpha)_k z^k}{(\gamma)_k k!} \quad \text{for } |z| < +\infty, \gamma \neq 0, -1, -2, \dots$$

where $(\lambda)_k = \Gamma(\lambda + k)/\Gamma(\lambda)$.

• For $Re(s) > 1$, $Re(t) < 0$ and $Re(s + t) > 1$, we have:

$$\mathbb{E}(W^s H^t) = \sqrt{\frac{\pi}{2}} \frac{2^{\frac{5+t-s}{2}}}{\Gamma(-t)\Gamma(\frac{s+t-1}{2})} \int_0^{+\infty} \int_0^{+\infty} \lambda^{s+t-2} \alpha^{-(1+2t)} A^2(\lambda, \alpha) d\alpha d\lambda$$

Incidentally, a consequence of Theorem 1.5, (1.2) and (1.3) is the Biane & Yor [4, p. 69] identity for the normalized Brownian excursion:

$$\int_0^1 \frac{1}{e(s)} ds \stackrel{law}{=} 2 \max_{0 \leq s \leq 1} e(s). \quad (1.5)$$

The paper is organized as follows: Section 2 contains the stochastic analysis of $dj(n)$ and the proof of its quasi-optimality for the stochastic order. In Section 3, we give the proof of Theorem 1.5. Also, one can find in Section 3, the computation of the limit law of the harmonic mean of depth of nodes in size n general trees. Section 3 can be read independently of Section 2.

2 Quasi-Optimality of $dj(n)$ for the stochastic order

2.1 Sketch of the proofs of Theorems 1.3 and 1.4

Consider the random variable

$$V_n = \frac{1}{2} \sum_{k=1}^n \frac{1}{|S_k| + 1}.$$

The proof of the quasi optimality on the average case in [21] or for the stochastic order of $dj(n)$ results from the fact that we can compare the cost of any algorithm with V_n , as stated by the two following lemmas. Set

$$\hat{T}_n(\omega) = \inf \{T_a(\omega) | a \in A_n\}.$$

We have

Lemma 2.1

$$\frac{(V_n - \hat{T}_n)_+}{\sqrt{n} \log n} \xrightarrow[n \rightarrow +\infty]{L_1} 0.$$

This implies immediately that for every algorithm a in A_n , we have

Corollary 2.2

$$\frac{(V_n - T_a)_+}{\sqrt{n \log n}} \xrightarrow[n \rightarrow +\infty]{L_1} 0.$$

V_n is a “quasi lower bound” for the cost of the algorithms and $d_j(n)$ “achieves” it:

Lemma 2.3

$$\frac{|T_{d_j(n)} - V_n|}{\sqrt{n \log n}} \xrightarrow[n \rightarrow +\infty]{L_1} 0.$$

This normalization is not trivial since

$$\mathbb{E}(V_n) \simeq \frac{1}{\sqrt{2\pi}} \sqrt{n \log n} \quad (2.6)$$

see [21]. A consequence of Lemma 2.3 is that $\frac{T_{d_j(n)}}{\sqrt{n \log n}}$ and $\frac{V_n}{\sqrt{n \log n}}$ have the same limit law, which is given by

Theorem 2.4

$$\frac{V_n}{\sqrt{n \log n}} \xrightarrow[n \rightarrow +\infty]{law} \frac{|N|}{2}$$

where N denotes a centered Gaussian variable with variance 1.

Theorems 1.3 and 1.4 follow immediately Lemma 2.3, Corollary 2.2 and Theorem 2.4. The object of the rest of Section 2 is devoted to the proofs of Lemmas 2.1, 2.3 and of Theorem 2.4.

2.2 A counting principle

A variant of this principle is the key result in the proof of average case optimality in Odlyzko [22] as well as in Chassaing [6]. At the end of this section, we shall use it to prove Lemmata 2.1 and 2.3. We believe Odlyzko's principle is interesting in its own right: for instance, it leads to a simple proof of Theorem 1.5. We try to give Odlyzko's principle in general form, in order to spare the reader many calls to very similar arguments in this paper, and also in [9].

Let $(Z_k^{(n)})_{0 \leq k \leq n}$ denote an array of positive integer-valued random variables. Let α be a positive constant, less than 1, and let $G(n)$ and $H(n)$ be two sequences of positive numbers. We set

$$U_n = \sum_{k=1}^n \frac{1}{Z_k^{(n)}},$$

and

$$\tilde{U}_n = \sum_{Z_k^{(n)} \leq H(n)} \frac{1}{Z_k^{(n)}},$$

Note that $U_n \leq n$ so that $u_n = E(U_n)$ is well defined. We assume that there exists a set Ω_n of sample paths such that

$$\Pr(\bar{\Omega}_n) = o\left(\frac{u_n}{n}\right), \quad (2.7)$$

and that for any $\omega \in \Omega_n$

$$\forall k, l, |Z_{k+l}^{(n)} - Z_k^{(n)}| \leq l^\alpha G(n).$$

Furthermore we assume that

$$G(n) \ll H(n)^{1-\alpha},$$

and that

$$E(\tilde{U}_n) \ll u_n. \quad (2.8)$$

We consider a finite or denumerable set \tilde{A}_n of strictly increasing sequences a of positive integer-valued *random variables* $\{m_1, \dots, m_{T(a)}\}$, whose length $T(a)$ is *random*, such that $m_0 = 0$, $m_{T(a)} = n$, and such that the increments

$$w_i = m_{i+1} - m_i$$

satisfy

$$w_i \leq Z_{m_i}^{(n)} \vee H(n), \quad (2.9)$$

at least when $\omega \in \Omega_n$ (at the end of this Section, we shall choose $Z_k^{(n)} = 2(|S_k| + 1)(1 + o(n))$) in order to prove Lemmata 2.1 and 2.3). Let \tilde{T}_n denote the infimum of lengths of sequences in \tilde{A}_n .

Maybe hypothesis (2.7) and (2.8) will seem less arbitrary if one considers a simple problem of computation, where a strictly increasing sequence $u_k^{(\varepsilon)}$ satisfy $u_0^{(\varepsilon)} = 0$ and $u_{k+1}^{(\varepsilon)} - u_k^{(\varepsilon)} = \varepsilon z(u_k^{(\varepsilon)})$. The first index k_ε such that $u_k^{(\varepsilon)} \geq 1$ (see Figure 2.1) is easily seen to satisfy

$$\begin{aligned} k_\varepsilon &\sim \int_0^1 \sum_{k \geq 0} \frac{1}{\varepsilon z(u_k^{(\varepsilon)})} 1_{[u_k^{(\varepsilon)}, u_{k+1}^{(\varepsilon)}]}(x) dx \\ &\sim \frac{1}{\varepsilon} \int_0^1 \frac{dy}{z(y)}, \end{aligned}$$

under hypothesis of uniform continuity of z , if z is positive. One can see assumption (2.7) as a stochastic analog of the latter. As regards assumption (2.8), obviously it is the stochastic analog of a – strong – hypothesis of integrability for $1/z$, if z has a pole.

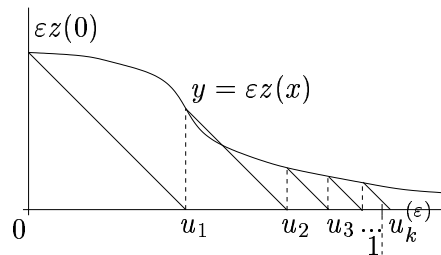


Figure 2.1

Theorem 2.5

$$\frac{(U_n - \tilde{T}_n)_+}{u_n} \xrightarrow{L_1} 0$$

Theorem 2.6 Assume that, for a sequence $a_n = (m_i^{(n)})_{0 \leq i \leq T(a_n)} \in \tilde{A}_n$,

$$E(\#\{i | Z_{m_i}^{(n)} \leq H(n)\} | I_{\Omega_n}) \ll u_n,$$

and that there exist a positive number δ , and a sequence ε_n of positive numbers decreasing to 0, such that for any integer n and any $\omega \in \Omega_n$, the sequence a_n satisfies

$$w_i^{(n)} \geq (1 - \varepsilon_n) Z_{m_i}^{(n)} \text{ if } (1 + \delta) Z_{m_i}^{(n)} \geq H(n).$$

Then

$$\frac{U_n - T(a_n)}{u_n} \xrightarrow{L_1} 0$$

The proof of this counting principle can be found in Chassaing et al. [9].

Proof of Lemma 2.1:

We set $\alpha = \frac{1}{2}$, $G(n) = c\sqrt{\log n}$, $H(n) = \log^{5/3} n$. Supposed (2.7) and (2.8) checked. For a in A_n , we consider the sequence $\tilde{a} = \{m_1, \dots, m_{T(\tilde{a})}\}$ obtained once we sort positions probed by a in increasing order, \tilde{a} being eventually supplemented by position n . So the finite set \tilde{A}_n of such sequences satisfy

$$\tilde{T}_n - 1 \leq \hat{T}_n \leq \tilde{T}_n. \quad (2.10)$$

For $\omega \in \Omega_n$, (2.12) yields:

$$w_i \leq 2(|S_{m_i}| + 1) + c\sqrt{w_i \log n}$$

and when $w_i \geq H(n)$,

$$w_i \leq 2 \frac{|S_{m_i}| + 1}{1 - c \log^{-1/3} n}.$$

Thus, choosing:

$$Z_k^{(n)} = 2 \frac{|S_k| + 1}{1 - c \log^{-1/3} n},$$

on one hand we see that

$$\frac{U_n - V_n}{u_n} \xrightarrow{L_1} 0$$

and on the other hand, owing to (2.10) and to Theorem 2.5,

$$\frac{(U_n + 1 - \hat{T}_n)_+}{u_n} \xrightarrow{L_1} 0.$$

u_n is given by equation (2.6) which gives the mean value of V_n :

$$u_n = \Theta(\sqrt{n} \log n).$$

Now, we have to check hypothesis (2.7) and (2.8).

A set of moderate variation: First, for n large enough (else, let Ω_n be void), let Ω_n be the set $\Omega_d(n)$ of sample paths ω such that for any positive k and m

$$m + k \leq n \Rightarrow |S_{m+k} - S_m| \leq d\sqrt{k \log n}$$

for some d such that $d < c/2$. Due to Chernoff bounds (see for instance Bollobás (1985) p.12):

$$\forall x \geq 0, \forall k, \Pr(|S_k| > x) \leq 2 \exp\left(-\frac{x^2}{2k}\right)$$

we deduce as in [22]:

$$\Pr(\Omega_d(n)) \geq 1 - 2n^{2-\frac{d^2}{2}}. \quad (2.11)$$

so finally $c = 5$ does the job. This is for (2.7).

Let consider u_i and u_{i+1} two consecutive probes and note $w_i = u_{i+1} - u_i$. Without loss of generality, suppose that ω is positive on $[u_i, u_{i+1}]$; on $[m_i, m_{i+1}]$, only four possible situations arise. See figures 2.2. In these four cases, we have for any i

$$S_{u_i} + S_{u_{i+1}} - w_i \geq -2 \quad (2.12)$$

Else one can not be sure that all the zeros are found. Then we have

$$w_i \leq Z_{m_i}^{(n)} \vee H(n)$$

To check assumption (2.8), we use Proposition 2.7 with $b = 5/3$. \square .

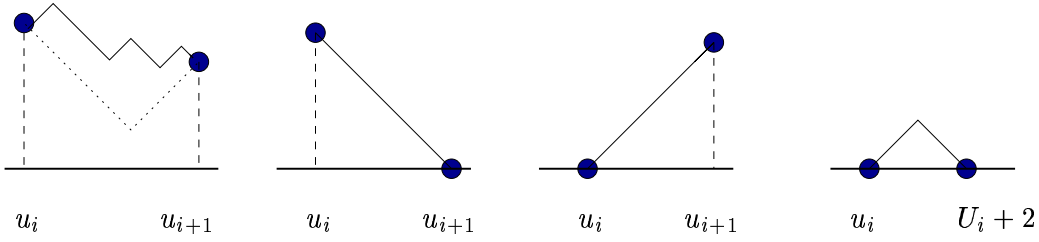


Figure 2.2

Proposition 2.7 For any positive number b , we have:

$$\frac{1}{2} \mathbb{E} \left(\sum_{h=1}^n \frac{\mathbb{I}_{\{h \mid |S_h| \leq 2 \log^b n\}}}{|S_h| + 1} \right) = O(\sqrt{n} \log \log n).$$

Proof: For any i , we have $\binom{h}{\lfloor \frac{h}{2} \rfloor} \geq \binom{h}{i}$. Thus

$$\begin{aligned} \mathbb{E} \left(\frac{\mathbb{I}_{\{h \mid |S_h| \leq \log^b n\}}}{|S_h| + 1} \right) &= \sum_{i=0, i \equiv h[2]}^{\lfloor \log^b n \rfloor} \frac{2 \binom{h}{\frac{h}{2} + \frac{i}{2}} 2^{-h}}{i + 1} \\ &\leq 2^{-h} \binom{h}{\lfloor \frac{h}{2} \rfloor} \sum_{i=0}^{\lfloor \log^b n \rfloor} \frac{1}{i + 1} \\ &\leq \frac{C(b)}{\sqrt{h}} \log \log n. \quad \square \end{aligned}$$

where $C(b)$ is a function depending only of b . This checks (2.8). \square

Equation (2.6) and Theorem 2.5 lead to:

$$\frac{(V_n - \hat{T}_n)_+}{\sqrt{n} \log n} \xrightarrow[n \rightarrow +\infty]{L_1} 0.$$

This ends the proof of Lemma 2.1 \square

Proof of Lemma 2.3 : Let $a(n) = (m_1, \dots, m_{R(n)})$ be the sequence of positions probed by $dj(n)$. Take $\varepsilon = 2/3$ in the double-jump algorithm. For α , $G(n)$ and Ω_n , we keep to the same definitions as in the preceding proof. Hypothesis (2.9) is satisfied for

$$Z_k^{(n)} = 2(S_k + 1),$$

so that this time U_n is equal to V_n ; by Proposition 2.7, we can notice that if $Z_{m_i} \geq (\log n)^{1+\varepsilon}$ then $w_i = \lfloor 2|S_{m_i}| - c\sqrt{2 \log n |S_{m_i}|} \rfloor$, and it is clear that there exists a sequence ε_n such that

$$w_i \geq (1 - \varepsilon_n) Z_{m_i}^{(n)}.$$

Now, denote as in [21], introduce the set $L_2 = \{i \mid Z_{m_i}^{(n)} \leq (\log n)^{1+\varepsilon}\} \cap \Omega_c(n)$. We have

Lemma 2.8

$$\mathbb{E}(\#L_2) \leq 6\sqrt{\frac{2n}{\pi}} \log^{\frac{1+\varepsilon}{2}} n.$$

Proof: See [21] Section 3.4. \square

2.3 Proof of Theorem 2.4

Denote by $\xi(n, l)$ the local time of ω at level l , that is:

$$\xi(n, l)(\omega) = \#\{k | 1 \leq k \leq n | S_k(\omega) = l\}.$$

Using

$$\frac{\xi(n, 0)}{\sqrt{n}} \xrightarrow{\text{law}} |N|$$

(see Révész [25] p.100), Theorem 2.4 follows immediately from

Proposition 2.9

$$\frac{1}{\sqrt{n} \log n} \sum_{k=1}^n \frac{1}{|S_k| + 1} - \frac{\xi(n, 0)}{\sqrt{n}} \xrightarrow{\text{proba}} 0$$

Proof :

$$\sum_{k=1}^n \frac{1}{|S_k| + 1} = \sum_{l=0}^n \frac{\xi(n, l) + \xi(n, -l)}{l + 1} = a_n + b_n + c_n + d_n$$

where

$$\begin{aligned} a_n &= \xi(n, 0) \left(1 + 2 \sum_{l=1}^{\lfloor c\sqrt{n \log n} \rfloor} \frac{1}{l + 1} \right), \\ b_n &= \sum_{l=\lfloor c\sqrt{n \log n} \rfloor + 1}^n \frac{\xi(n, l) + \xi(n, -l)}{l + 1}, \\ c_n &= \sum_{l=1}^{\lfloor \sqrt{n} \log^{-2} n \rfloor} \frac{\xi(n, l) + \xi(n, -l) - 2\xi(n, 0)}{l + 1}, \\ d_n &= \sum_{l=\lfloor \sqrt{n} \log^{-2} n \rfloor + 1}^{\lfloor c\sqrt{n \log n} \rfloor} \frac{\xi(n, l) + \xi(n, -l) - 2\xi(n, 0)}{l + 1}. \end{aligned}$$

We have

$$b_n \leq n \mathbb{I}_{\Omega_c}$$

and thus

$$\frac{b_n}{\sqrt{n} \log n} \xrightarrow{\text{proba}} 0.$$

Set

$$\delta_n = \sup_{|l| \leq \lfloor \sqrt{n} \log^{-2} n \rfloor} \left\{ \frac{|\xi(n, l) - \xi(n, 0)|}{\xi(n, 0)} \right\}.$$

According to Csáki & Földes [10], see also Révész [25] p. 122, we have

$$\delta_n \xrightarrow{\text{a.s.}} 0$$

Now,

$$\begin{aligned} \frac{|c_n|}{\sqrt{n} \log n} &\leq \frac{\xi(n, 0)}{\sqrt{n} \log n} \sum_{l=1}^{\lfloor \sqrt{n} \log^{-2} n \rfloor} \frac{|\xi(n, l) - \xi(n, 0)| + |\xi(n, -l) - \xi(n, 0)|}{\xi(n, 0)(l+1)} \\ &\leq \frac{2\xi(n, 0)}{\sqrt{n} \log n} \delta_n \sum_{|l| \leq \lfloor \sqrt{n} \log^{-2} n \rfloor} \frac{1}{l+1} \\ &\xrightarrow{\text{proba}} 0. \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left(\frac{d_n}{\sqrt{n} \log n}\right) &\leq \frac{4\mathbb{E}(\xi(n, 0))}{\sqrt{n} \log n} \sum_{l=\lfloor \sqrt{n} \log^{-2} n \rfloor + 1}^{\lfloor c\sqrt{n} \log n \rfloor} \frac{1}{l+1} \\ &\rightarrow 0. \end{aligned}$$

Finally

$$\frac{a_n}{\sqrt{n} \log n} - \frac{\xi(n, 0)}{\sqrt{n}} \xrightarrow{\text{proba}} 0. \quad \square$$

3 Relation between the search of zeros, trees and excursions

The one-to-one correspondences between discrete excursions and trees provide some efficient and fine methods to study the trees. A lot of functionals of trees (height, width, profile, total displacement...) can be expressed as functionals (maximum, local time, area...) of related discrete excursions (see Aldous [1] for a survey on this topic). Hence, the limit law of these functionals can often be expressed as the law of functionals of normalized Brownian excursion (maximum, area, local time...) by invariance principles (see [1]). In this paper, we use the properties of the Bernoulli excursions associated to the binary trees by the breadth first search. The algorithm $sj(n)$ gives a strong relation between the height of a binary tree t and its cost on the Bernoulli excursion associated to t by the b.f.s.. As a consequence, we can give a simple description of the limit law of the couple height-width in random binary trees, in term of Brownian excursion.

3.1 The couple height-width

In the case of labeled trees with n nodes, Aldous [3, Th. 3] proves that:

$$\frac{1}{\sqrt{n}}(H_L^{(n)}, W_L^{(n)}) \xrightarrow{law} \left(2 \max_{0 \leq t \leq 1} e(t), \frac{1}{2} \max_{t \geq 0} l(t) \right),$$

where l is the local time of the normalized Brownian excursion e , defined by:

$$\int_0^a l(x) dx = \int_0^1 \mathbb{I}_{[0,a]}(e(s)) ds.$$

The proof given by Aldous is also valid in the case of binary trees with n leaves and gives:

$$\frac{1}{\sqrt{2n}}(H_B^{(n)}, W_B^{(n)}) \xrightarrow{law} \left(2 \max_{0 \leq t \leq 1} e(t), \frac{1}{2} \max_{0 \leq t \leq 1} l(t) \right). \quad (3.13)$$

In the case of general trees with n nodes, it gives:

$$\frac{1}{\sqrt{n}}(H_G^{(n)}, W_G^{(n)}) \xrightarrow{law} \left(\sqrt{2} \max_{0 \leq t \leq 1} e(t), \frac{1}{\sqrt{2}} \max_{t \geq 0} l(t) \right).$$

Similar results can be obtained for other families of simple trees.

Using the result of Aldous (3.13), we can give a first proof of Theorem 1.5.

First proof of Theorem 1.5 : in order to relate Aldous result (3.13) involving the local time of Brownian excursion to Theorem 1.5 (where the results are expressed in term of Brownian excursion) we use the following Jeulin's identity.

Jeulin's identity

Let $(e(s))_{0 \leq s \leq 1}$ be a normalized Brownian excursion with local time $(l(s))_{s \geq 0}$. Define

$$L(y) = \int_0^y l(x) dx$$

and

$$k(t) = L^{-1}(t) = \sup \left\{ y \mid \int_0^y l(x) dx < t \right\}.$$

Jeulin [17] (see also [4, p. 70]) proved that the process $(\tilde{e}(s))_{0 \leq s \leq 1}$ defined by:

$$\tilde{e}(s) = \frac{1}{2} l(k(s)) \quad (3.14)$$

is itself a normalized Brownian excursion. It can be written

$$\frac{1}{2\tilde{e}} = (L^{-1})' = k',$$

with the consequence:

$$l\left(\int_0^s \frac{1}{2\tilde{e}(u)} du\right) = 2\tilde{e}(s). \quad (3.15)$$

A direct corollary of that is the identity:

$$\left(2 \max_{0 \leq t \leq 1} e(t), \frac{1}{2} \max_{t \geq 0} l(t)\right) = \left(\int_0^1 \frac{1}{\tilde{e}(s)} ds, \max_{0 \leq s \leq 1} \tilde{e}(s)\right). \quad (3.16)$$

The first component of (3.16) follows from (3.15) because

$$\inf\{x > 0 \mid l(x) = 0\} = \max_{0 \leq s \leq 1} e(s) = \int_0^1 \frac{1}{2\tilde{e}(s)} ds.$$

The second by taking the maximum at each side of (3.15). Thus, (3.13) is equivalent to Theorem 1.5. \square

Hence, this first proof of Theorem 1.5 (valid in the general case with minor modifications) relies on Jeulin's identity [17] and on Aldous [3] which relies itself on two particularly technical papers: Aldous [2] and Drmota & Gittenberger [13].

We give below a second proof of Theorem 1.5, direct and simple, that does not use any of the results mentioned above ([2, 3, 13, 17]).

3.1.1 Excursions

Let us consider a standard linear Brownian motion $\mathcal{B} = (\mathcal{B}_t)_{t \geq 0}$. Consider g the last zero before 1 of \mathcal{B} and d the first after:

$$\begin{aligned} g &= \sup\{0 \leq t \leq 1 \mid \mathcal{B}_t = 0\}, \\ d &= \inf\{t \geq 1 \mid \mathcal{B}_t = 0\}. \end{aligned}$$

Then, we have $g < 1 < d$ *a.s.* The process e defined for $0 \leq s \leq 1$ by

$$e(s) = \left| \frac{\mathcal{B}_{g+(d-g)s}}{\sqrt{d-g}} \right|$$

is called the standard *normalized Brownian excursion* (see Figure 3.1).

Consider $Es(2n)$ the subset of $2n$ steps simple symmetric random walks $\tilde{\omega} = (\tilde{S}_k(\tilde{\omega}))_{k=0, \dots, 2n}$ conditioned to satisfy:

$$\tilde{S}_0(\tilde{\omega}) = 0, \quad \tilde{S}_{2n}(\tilde{\omega}) = 0 \text{ and } \tilde{S}_k(\tilde{\omega}) > 0 \text{ for } k \in \{1, \dots, 2n-1\}.$$

$Es(2n)$ is called the set of Bernoulli excursions of size $2n$; its cardinality is given by the $n-1^{\text{th}}$ Catalan number:

$$C_{n-1} = \binom{2n-2}{n-1} \frac{1}{n}.$$

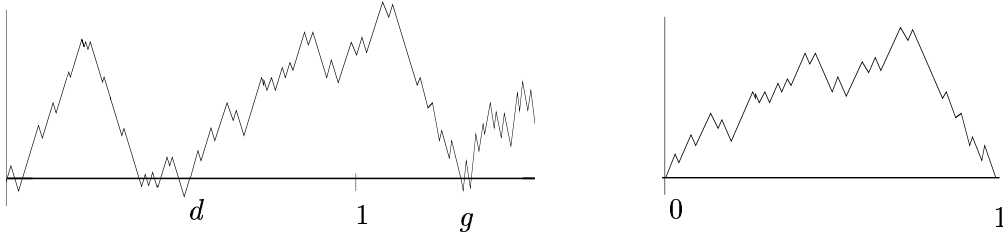


Figure 3.1

Thus, the C_{n-1} excursions of $Es(2n)$ are equiprobable and we will write $\tilde{\mathbb{P}}$ for the conditional probability \mathbb{P} to $Es(2n)$, that is:

$$\tilde{\mathbb{P}}(\tilde{\omega} \in A) = \mathbb{P}(\tilde{\omega} \in A \mid \tilde{\omega} \in Es(n)).$$

For any positive integers l, k , such that $k \leq l \leq 2n - k$, we have

$$\tilde{\mathbb{P}}(\tilde{S}_k = l) = \frac{n l^2}{k(2n - k)} \frac{\binom{k}{\frac{k-l}{2}} \binom{2n-k}{\frac{2n-k-l}{2}}}{\binom{2n-2}{n-1}} \mathbb{I}_{k \equiv l[2]}. \quad (3.17)$$

In this formula, $\frac{l}{k} \binom{k}{\frac{k-l}{2}}$ is the number of positive paths from $(0, 0)$ to (k, l) , $\frac{l}{2n-k} \binom{2n-k}{\frac{2n-k-l}{2}}$ the number of positive paths from (k, l) to $(2n, 0)$ (generalization of the ballot problem).

There exists, as in the case of simple symmetric random walks (see Subsection 2.2), a large subset of excursions in $Es(2n)$ satisfying a property of moderate variation. Denote by $\tilde{\Omega}_c(2n)$ the set of $2n$ -length excursions such that for any l, k in the set $\{0, 1, \dots, 2n\}$,

$$|\tilde{S}_k - \tilde{S}_l| \leq c \sqrt{|k - l| \log n}.$$

Lemma 3.1 *For every $\beta > 0$ there exist $c > 0$ such that, for n sufficiently large:*

$$\tilde{\mathbb{P}}(\tilde{\Omega}_c(2n)) \geq 1 - n^{-\beta}.$$

Proof : Following (2.11), we have:

$$\tilde{\mathbb{P}}(\tilde{\omega} \in \tilde{\mathcal{C}}\tilde{\Omega}_c(2n)) = \mathbb{P}(\tilde{\omega} \in \mathcal{C}\Omega_c(2n) / \tilde{\omega} \in Es(2n)) \leq \frac{\mathbb{P}(\mathcal{C}\Omega_c(n))}{\mathbb{P}(Es(n))} \leq \frac{2(2n)^{2-\frac{c^2}{2}}}{c_1 n^{-\frac{3}{2}}} \quad \square$$

The normalized Brownian excursion can be naturally viewed as the weak limit of *Bernoulli excursions*; see for example [1, 11, 27, 28, 30] for applications.

3.1.2 Expression of $(H_B^{(n)}, W_B^{(n)})$ in term of Bernoulli excursions

The breadth first search allows us to express the couple height-width in term of Bernoulli excursion.

Correspondence binary tree - excursion :

This correspondence associates with each $2n$ -length excursion $\tilde{\omega}$ a binary tree with $n - 1$ internal-nodes.

$$t_B^{(n)} : Es(2n) \longrightarrow T_B^{(n)}$$

$$\tilde{\omega} \longrightarrow t = t_B^{(n)}(\tilde{\omega})$$

Figure 3.2 shows a binary tree in correspondence with an excursion (breadth first search).

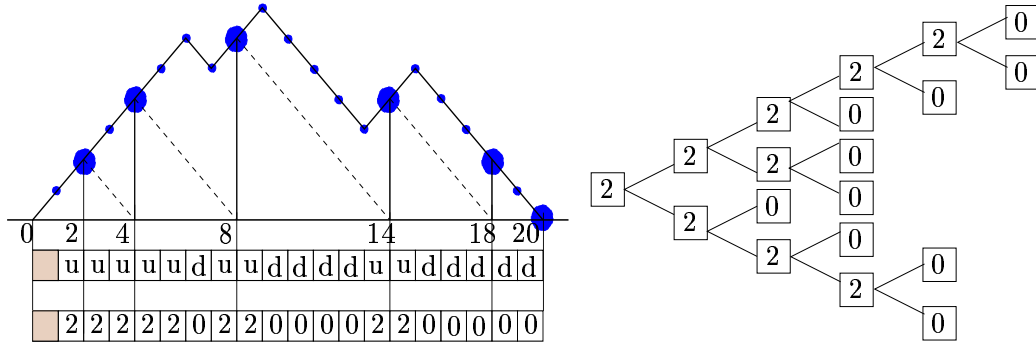


Figure 3.2 : Excursion - Binary tree

The height

With this correspondence, one can observe that $\tilde{T}_{sj(2n)}(\tilde{\omega})$, the number of probes used by the simple jump algorithm on the excursion $\tilde{\omega}$, and the height $H_B^{(n)}(t)$ are related by:

$$H_B^{(n)}(t) = \tilde{T}_{sj(2n)}(\tilde{\omega}) - 1. \tag{3.18}$$

Indeed, denote by $m_1 = 2, \dots, m_{\tilde{T}_{sj}}$ the (ordered) positions probed by sj on the first excursion of ω (which is assumed to be positive). By definition of sj , the $i + 1^{th}$ jump is S_{m_i} (for $i \geq 0$) so that

$$m_{i+1} = m_i + S_{m_i}.$$

The $i + 2^{th}$ jump, $S_{m_{i+1}}$ satisfies:

$$S_{m_{i+1}} = S_{m_i} + \sum_{h=m_i+1}^{m_i+S_{m_i}} X_h = \sum_{h=m_i+1}^{m_i+S_{m_i}} (1 + X_h) = \sum_{h=m_i+1}^{m_i+S_{m_i}} Y_h$$

where the r.v. Y_h are i.i.d. and

$$\mathbb{P}(Y_h = 0) = \mathbb{P}(Y_h = 2) = \frac{1}{2},$$

since ω is a simple symmetric random walk (that is, we have $S_0(\omega) = 0$; $S_k(\omega) = X_1(\omega) + \dots + X_k(\omega)$ where the $(X_i)_{i \geq 0}$ are i.i.d. r.v. with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$). If we condition the first excursion to be $2n$ long, the process $(Z_i)_{i \geq 1} = (S_{m_i})_{i \geq 1}$ is a Galton-Watson process with offspring distribution 0 and 2 with probability 1/2 (the initial population is $Z_1 = 2$) conditioned to have a total progeny of $2n - 1$ individuals. The number of jumps used by $sj(2n)$ on an excursion $\tilde{\omega}$ is simply the number of generations in the processus $(Z_i)_{i \geq 1}$; the height of the binary tree representation of this Galton-Watson process is exactly the number of generations minus 1.

Denote by \tilde{V}_{2n} the r.v.:

$$\tilde{V}_{2n}(\tilde{\omega}) = \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k(\tilde{\omega})},$$

where $\tilde{\omega} = (\tilde{S}_0(\tilde{\omega}), \dots, \tilde{S}_{2n}(\tilde{\omega}))$ is a Bernoulli excursion. The height $H_B^{(n)}(t)$ and \tilde{V}_{2n} are related by:

Lemma 3.2

$$\frac{1}{\sqrt{n}} |H_B^{(n)}(t) - \tilde{V}_{2n}(\tilde{\omega})| \xrightarrow{L_1} 0.$$

Proof : Using (3.18), we just prove

$$\frac{1}{\sqrt{n}} |\tilde{T}_{sj(2n)}(\tilde{\omega}) - \tilde{V}_{2n}(\tilde{\omega})| \xrightarrow{L_1} 0.$$

Denote by $(m_i)_{i=1, \dots, \tilde{T}_{sj(2n)}}$ the set of ordered probed points by sj and by $k_i = m_i - m_{i-1}$ the difference between two probes ($m_0 = 0$). The cost of $sj(2n)$ is exactly

$$\tilde{T}_{sj(2n)} = \sum_{i=0}^{\tilde{T}_{sj(2n)}} \sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i}.$$

Let α be a real number in $]0, 1/2[$.

$$\left| \sum_{i=0}^{\tilde{T}_{sj(2n)}} \sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} - \sum_{h=1}^{2n-1} \frac{1}{\tilde{S}_h} \right| \leq 2n^{1/2-\alpha} + A_n + B_n$$

where

$$\begin{aligned} A_n &\leq \sum_{i=0}^{\tilde{T}_{sj(2n)}} \left(\sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} - \frac{1}{\tilde{S}_h} \right) \mathbb{I}_{k_i \geq (\log^{1+\varepsilon} n)} \mathbb{I}_{[n^{1/2-\alpha}, 2n-n^{1/2-\alpha}]}(h) \\ B_n &\leq \sum_{i=0}^{\tilde{T}_{sj(2n)}} \left(\sum_{h=m_i+1}^{m_{i+1}} \frac{1}{k_i} - \frac{1}{\tilde{S}_h} \right) \mathbb{I}_{k_i \leq (\log n)^{1+\varepsilon}} \mathbb{I}_{[n^{1/2-\alpha}, 2n-n^{1/2-\alpha}]}(h) \end{aligned}$$

First, using (3.17), Stirling formula and that $\binom{k}{k/2} \geq \binom{k}{(k-l)/2}$ for any l and k ,

$$\begin{aligned} \mathbb{E}(B_n) &\leq \mathbb{E}(\#\{h|h \in [n^{1/2-\alpha}, 2n - n^{1/2-\alpha}], \tilde{S}_h \leq (\log n)^{1+\varepsilon}\}) \\ &\leq \sum_{h=n^{1/2-\alpha}}^{2n-n^{1/2-\alpha}} \sum_{l=1}^{\lfloor \log^{1+\varepsilon} n \rfloor} \tilde{\mathbb{P}}(\tilde{S}_h = l) \\ &= O(\log^{3+3\varepsilon} n) \end{aligned}$$

A simple computation gives that, we have:

$$\begin{aligned} \mathbb{I}_{\tilde{\Omega}_c(n)}(\tilde{\omega}) A_n &\leq (\log^{-\varepsilon/2} n) \tilde{V}_{2n}, \\ \mathbb{I}_{\tilde{\Omega}_c(n)}(\tilde{\omega}) A_n &\leq n. \end{aligned}$$

We choose c large enough in order to have $\beta > 3/2$ in Lemma 3.1, and Lemma 3.2 is proved. \square

The width

Thanks to the b.f.s., the profile of t can be read on $\tilde{\omega}$:

$$\begin{aligned} Z_0(t) &= 1 \\ Z_1(t) &= \tilde{S}_2(\tilde{\omega}) \\ Z_2(t) &= \tilde{S}_{2+Z_1(t)}(\tilde{\omega}) \\ Z_{k+1}(t) &= \tilde{S}_{2+Z_1(t)+\dots+Z_k(t)}(\tilde{\omega}). \end{aligned}$$

Then denoting by $l(k) = 2 + Z_1(t) + \dots + Z_k(t)$, we have

$$W_B^{(n)}(t) = \max_k \tilde{S}_{l(k)}(\tilde{\omega}).$$

As a consequence, we have:

$$\tilde{M}_{2n} \stackrel{def}{=} \max_k \tilde{S}_k(\tilde{\omega}) \geq \max_k \tilde{S}_{l(k)}(\tilde{\omega}) = W_B^{(n)}(\tilde{\omega}).$$

The moderate variation of the paths (cf. Lemma 3.1) yields:

Lemma 3.3

$$\frac{1}{\sqrt{n}} \|\tilde{M}_{2n}(\tilde{\omega}) - W_B^{(n)}(t)\|_1 = O(n^{-1/4} \sqrt{\log n})$$

Proof: We have

$$0 \leq \mathbb{E}(\tilde{M}_{2n}(\tilde{\omega}) - W_B^{(n)}(t)) = \mathbb{E}(\tilde{M}_{2n} - \max_k \tilde{S}_{l(k)}).$$

We consider an index $K(\tilde{\omega})$ such that $\tilde{M}_{2n}(\tilde{\omega}) = \tilde{S}_K(\tilde{\omega})$. There exists an integer $i(\tilde{\omega})$ such that

$$l(i) \leq K \leq l(i+1) \wedge n.$$

Then,

$$\begin{aligned} \mathbb{E}(\tilde{M}_{2n} - \max_k \tilde{S}_{l(k)}) &\leq \mathbb{E}(\tilde{S}_K - \tilde{S}_{l(i)}) \\ &\leq \mathbb{E}(\mathbb{I}_{\Omega_c} c \sqrt{\log(2n)} |K - l(i)| + n \mathbb{I}_{\mathfrak{C}\Omega_c}) \\ &\leq \mathbb{E}(c \sqrt{\log(2n)} \tilde{M}_{2n} \mathbb{I}_{\Omega_c}) + n \mathbb{P}(\mathfrak{C}\Omega_c) \\ &= O(n^{1/4} \sqrt{\log n}), \end{aligned}$$

for a suitable c , by Jensen inequality. \square

The couple height-width

Using Lemmas 3.2 and 3.3, we have:

$$\frac{1}{\sqrt{n}} \left\| (H_B^{(n)}(t), W_B^{(n)}(t)) - (\tilde{V}_{2n}(\tilde{\omega}), \tilde{M}_{2n}(\tilde{\omega})) \right\|_1 = o(1).$$

Thus, $\frac{1}{\sqrt{n}}(H_B^{(n)}, W_B^{(n)})$ and $\frac{1}{\sqrt{n}}(\tilde{V}_{2n}, \tilde{M}_{2n})$ have the same limit law. The proof of Theorem 1.5 reduces to the proof of Proposition 3.4.

Proposition 3.4

$$\frac{1}{\sqrt{2n}} (\tilde{V}_{2n}, \tilde{M}_{2n}) \xrightarrow{law} \left(\int_0^1 \frac{1}{e(s)} ds, \max_{0 \leq s \leq 1} e(s) \right).$$

To prove Proposition 3.4 we need the following Lemma:

Lemma 3.5 *Let $(X_n)_n$ and $(X_n^{(a)})_{n,a}$ be two families of \mathbb{R}^2+ valued r.v. on the same probability space \mathcal{P}_n such that:*

$$(X_n^{(a)})_n \xrightarrow[n \rightarrow +\infty]{law} X^{(a)}.$$

Suppose that there exists a probability space \mathcal{P} and a \mathbb{R}^2+ valued r.v. X on \mathcal{P} such that, for any $a > 0$,

$$\begin{aligned} \|X_n - X_n^{(a)}\|_1 &\leq C_1 a \quad \text{for any } a > 0, \\ \|X - X^{(a)}\|_1 &\leq C_2 a \quad \text{for any } a > 0 \end{aligned}$$

for two constants C_1 and C_2 . Then

$$X_n \xrightarrow{law} X.$$

Proof : For any $a > 0$ and $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2_+$, we have:

$$\begin{aligned} |\mathbb{E}(e^{-\lambda \cdot X_n}) - \mathbb{E}(e^{-\lambda \cdot X})| &\leq |\mathbb{E}(e^{-\lambda \cdot X_n}) - \mathbb{E}(e^{-\lambda \cdot X_n^{(a)}})| + |\mathbb{E}(e^{-\lambda \cdot X_n^{(a)}}) - \mathbb{E}(e^{-\lambda \cdot X^{(a)}})| \\ &\quad + |\mathbb{E}(e^{-\lambda \cdot X^{(a)}}) - \mathbb{E}(e^{-\lambda \cdot X})| \end{aligned}$$

Where $u \cdot v$ is the scalar product of u with v . Thus,

$$\limsup_n |\mathbb{E}(e^{-\lambda \cdot X_n}) - \mathbb{E}(e^{-\lambda \cdot X})| \leq \|\lambda\|_\infty C_1 a + 0 + \|\lambda\|_\infty C_2 a.$$

Since this relation is true for any $a > 0$, Lemma 3.5 is proved. \square

Proof of Proposition 3.4 : Denote by e_n the continuous process

$$e_n(t) = \frac{\tilde{S}_{[2nt]} + \{2nt\}(\tilde{S}_{[2nt]+1} - \tilde{S}_{[2nt]})}{\sqrt{2n}}, \quad 0 \leq t \leq 1,$$

where $[x]$, (resp. $\{x\}$) is the integer (resp. fractional) part of x . When $n \rightarrow +\infty$, it converges weakly to the Brownian excursion (see Gikhman & Skorohod [16]). We have

$$\tilde{M}_{2n}/\sqrt{2n} = \max_{0 \leq t \leq 1} e_n(t)$$

and

$$\frac{\tilde{V}_{2n}}{\sqrt{2n}} = \frac{1}{2n} \sum_{k=1}^{2n-1} \frac{1}{e_n(k/2n)}.$$

Define $\tilde{V}_{2n}^{(a)}$ by

$$\tilde{V}_{2n}^{(a)} = \frac{1}{2n} \sum_{k=1}^{2n-1} \frac{\mathbb{I}_{\{e_n \geq a\}}}{e_n(k/2n)}.$$

Set

$$\begin{aligned} X_n &= \frac{1}{\sqrt{2n}}(\tilde{V}_{2n}, \tilde{M}_{2n}), \\ X_n^{(a)} &= \frac{1}{\sqrt{2n}}(\tilde{V}_{2n}^{(a)}, \tilde{M}_{2n}), \\ X &= \left(\int_0^1 \frac{1}{e(s)} ds, \max_{0 \leq s \leq 1} e(s) \right), \\ X^{(a)} &= \left(\int_0^1 \frac{\mathbb{I}_{\{e(s) \geq a\}}}{e(s)} ds, \max_{0 \leq s \leq 1} e(s) \right). \end{aligned}$$

Proposition 3.4 is equivalent to

$$X_n \xrightarrow{law} X.$$

We shall prove this convergence using Lemma 3.5.

The convergence of $X_n^{(a)}$ to $X^{(a)}$ when n goes to ∞ results from the continuity of the functional. To conclude, we just need to prove the two following lemmas:

Lemma 3.6 *There exists a positive constant C_1 such that, for any $a > 0$,*

$$\|X_n - X_n^{(a)}\|_1 = \frac{1}{\sqrt{2n}} \mathbb{E} \left(\sum_{k=1}^{2n-1} \frac{I_{\tilde{S}_k \leq a\sqrt{2n}}}{\tilde{S}_k} \right) \leq C_1 a.$$

Lemma 3.7 *There exists a positive constant C_2 such that, for any $a > 0$,*

$$\|X - X^{(a)}\|_1 = \mathbb{E} \left(\int_0^1 \frac{\mathbb{I}_{e(s) \leq a}}{e(s)} ds \right) \leq C_2 a.$$

Proof of Lemma 3.6 : Using Formula (3.17), we have

$$\begin{aligned} \frac{1}{\sqrt{2n}} \mathbb{E} \left(\sum_{k=1}^{2n-1} \frac{I_{\tilde{S}_k \leq a\sqrt{2n}}}{\tilde{S}_k} \right) &= \frac{1}{\sqrt{2n}} \sum_{k=1}^{2n-1} \sum_{l=1}^{a\sqrt{2n}} \frac{\tilde{\mathbb{P}}(\tilde{S}_k = l)}{l} \\ &\leq c_1 \cdot \frac{n}{\sqrt{n} \binom{2n-2}{n-1}} \sum_{k=1}^n \sum_{l=1}^{a\sqrt{2n}} \frac{l \binom{k}{\frac{k-l}{2}} \binom{2n-k}{\frac{2n-k}{2}}}{k(2n-k)} \mathbb{I}_{k \equiv l[2]} \\ &\leq c_2 \frac{1}{\sqrt{n}} \sum_{l=1}^{a\sqrt{2n}} \sum_{k=1}^n \frac{l \binom{k}{\frac{k-l}{2}}}{2^k k} \mathbb{I}_{k \equiv l[2]}. \end{aligned}$$

Note that

$$\begin{aligned} \frac{l \binom{k}{\frac{k-l}{2}}}{2^k k} \mathbb{I}_{k \equiv l[2]} &= \mathbb{P}(S_k = l) \frac{l}{k} \\ \mathbb{P}(S_k = l) \frac{l}{k} &= \frac{l}{k} \frac{1}{2\pi} \int_0^{2\pi} \cos^k(t) \cos(lt) dt \\ &= \frac{2}{\pi} \int_0^{\pi/2} \sin(lt) \cos^{k-1}(t) \sin t dt \text{ for } k \geq 1. \end{aligned}$$

Thus,

$$\begin{aligned} \frac{1}{\sqrt{2n}} \mathbb{E} \left(\sum_{k=1}^{2n-1} \frac{I_{\tilde{S}_k \leq a\sqrt{2n}}}{\tilde{S}_k} \right) &\leq \frac{c_3}{\sqrt{n}} \sum_{l=1}^{a\sqrt{2n}} \int_0^{\pi/2} \sin(lt) \sin t \frac{1 - \cos^n t}{1 - \cos t} dt \\ &= \frac{c_3}{\sqrt{n}} \sum_{l=1}^{a\sqrt{2n}} \int_0^{\pi/2} \sin(lt) \frac{(1 + \cos t)(1 - \cos^n t)}{\sin t} dt. \quad (3.19) \end{aligned}$$

Let us expand this sum and bound its terms. Set

$$\begin{aligned} I_l &= \int_0^{\pi/2} \frac{\sin(lt) \cos t}{\sin t} dt, \\ J_l &= \int_0^{\pi/2} \frac{\sin(lt)}{\sin t} dt. \end{aligned}$$

We notice that

$$J_l = I_{l-1} + \frac{\sin((l-1)\frac{\pi}{2})}{l-1} \text{ and } I_l = J_{l-1} + \frac{\sin(\frac{\pi l}{2})}{l}.$$

So J_l and I_l are uniformly bounded. We have

$$\begin{aligned} \left| \int_0^{\pi/2} \frac{\sin(lt)}{\sin t} \cos^n t \, dt \right| &\leq \int_0^{\pi/2} \left| \frac{\sin(lt)}{\sin t} \right| \cos^n t \, dt \\ &\leq l \int_0^{\pi/2} \cos^n t \, dt \\ &= \frac{l}{2} \sqrt{\pi} \frac{l \Gamma(\frac{n+1}{2})}{\Gamma(n/2 + 1)} \end{aligned}$$

This last term is uniformly bounded for $l \in \{1, \dots, a\sqrt{2n}\}$ because

$$\frac{\sqrt{n} \Gamma(\frac{n+1}{2})}{\Gamma(n/2 + 1)}$$

converges when $n \rightarrow +\infty$. So each term of the sum in (3.19) is bounded and the proof is complete. \square

Proof of Lemma 3.7 : According to [14, Prop. 3.4],

$$\mathbb{E} \left(\int_0^1 \mathbb{I}_{e(s) \leq a} \, ds \right) = 1 - 2e^{-2a^2},$$

so, we have:

$$\mathbb{E} \left(\int_0^1 \frac{\mathbb{I}_{e(s) \leq a}}{e(s)} \, ds \right) \leq 8a. \quad \square$$

3.2 Harmonic mean of depth of nodes in size n general trees

Denote by $T_G^{(n)}$ the set of rooted, plane, *general trees* with n unlabeled vertices. For a tree $t \in T_G^{(n)}$, denote by $N_G^{(n)}(t)$ the set of its nodes (different from the root). For s a node of $N_G^{(n)}(t)$, denote by $h(s)$ its distance to the root. The trees of $T_G^{(n)}$ are assumed equiprobable. We have:

Theorem 3.8 (*Harmonic mean of depth of nodes*)

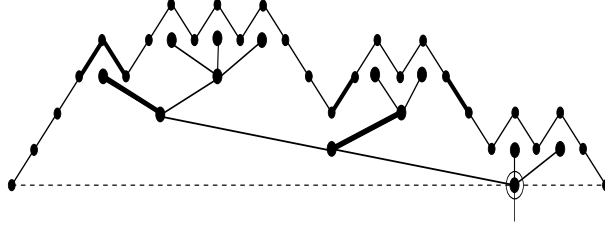
$$\sqrt{\frac{2}{n}} \left(\sum_{s \in N_G^{(n)}(t)} \frac{1}{h(s)} \right) \xrightarrow[n \rightarrow +\infty]{law} V.$$

Using Lemma 3.2 we just have to prove that $\frac{1}{\sqrt{n}} \sum_{s \in N_G^{(n)}(t)} \frac{1}{h(s)}$ and $\frac{1}{2\sqrt{n}} \tilde{V}_{2n}$ have the same limit law. It will follow from (3.21) and Lemma 3.9 below.

Correspondence general tree - excursion:

This classical correspondence associates with each excursion of $Es(2n)$, a general tree with n nodes. The nodes are visited in the same order as in the depth first search; one can consider the height of a fly walking clockwise around the tree from the root to the root; see Cormen et al. [7] chap. 23, Takács [27] or Le Gall [20] p. 454 (see Figure 3.3).

$$\begin{aligned} t_G^{(n)} : Es(2n) &\longrightarrow T_G^{(n)} \\ \tilde{\omega} &\longrightarrow t_G^{(n)}(\tilde{\omega}) \end{aligned}$$

**Figure 3.3** : Excursion - General tree

This correspondence associates with each edge of t , two steps of the excursion $\tilde{\omega}$ (the two extreme steps of the excursion are figured by the root). So, with each node at height $k > 0$, one can associate two steps $[\tilde{S}_{k_1}, \tilde{S}_{k_1+1}]$, $[\tilde{S}_{k_2}, \tilde{S}_{k_2+1}]$ of the excursion such that

$$\tilde{S}_{k_1} \wedge \tilde{S}_{k_1+1} = \tilde{S}_{k_2} \wedge \tilde{S}_{k_2+1} = k. \quad (3.20)$$

Then,

$$\sum_{s \in N_G^{(n)}(t_G^{(n)}(\tilde{\omega}))} \frac{1}{h(s)} = \frac{1}{2} \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k(\tilde{\omega}) \wedge \tilde{S}_{k+1}(\tilde{\omega})}. \quad (3.21)$$

Lemma 3.9

$$\frac{1}{\sqrt{n}} \left\| \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k \wedge \tilde{S}_{k+1}} - \tilde{V}_{2n} \right\|_1 = O\left(\frac{\log n}{\sqrt{n}}\right).$$

Proof :

$$0 \leq \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k \wedge \tilde{S}_{k+1}} - \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k} \leq \sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k^2}$$

$$\begin{aligned} \mathbb{E}\left(\sum_{k=1}^{2n-1} \frac{1}{\tilde{S}_k^2}\right) &= \sum_{k=1}^{2n-1} \sum_{l=1}^{k \wedge (2n-k)} \frac{\tilde{\mathbb{P}}(\tilde{S}_k = l)}{l^2} \\ &\leq 2 \sum_{k=1}^n \sum_{l=1}^k \frac{n}{k(2n-k)} \frac{\binom{k}{(k-l)/2} \binom{2n-k}{(2n-k-l)/2}}{\binom{2n-2}{n-1}} \mathbb{1}_{k \equiv l[2]} \\ &\leq 2 \sum_{k=1}^n \frac{n}{k(2n-k) \binom{2n-2}{n-1}} \sum_{l=1}^k \binom{k}{(k-l)/2} \binom{2n-k}{(2n-k-l)/2} \end{aligned}$$

To conclude, note that

$$\sum_{l=1}^k \binom{k}{(k-l)/2} \binom{2n-k}{(2n-k-l)/2} \leq \binom{2n}{n}. \quad \square$$

3.2.1 Table of convergences

In the case of simple symmetric random walks, we have seen:

$$\begin{aligned} \frac{T_{dj(n)}}{\sqrt{n} \log n} &\xrightarrow{\text{law}} |N| \\ \frac{V_n}{\sqrt{n} \log n} &\xrightarrow{\text{law}} |N| \end{aligned}$$

In the case of excursion $\tilde{\omega}$ of length $2n$:

Random variables								
	$\frac{H_B^{(n)}}{\sqrt{2n}}$	$\frac{W_B^{(n)}}{\sqrt{2n}}$	$\frac{H_G^{(n)}}{\sqrt{n}}$	$\frac{W_G^{(n)}}{\sqrt{n}}$	$\sqrt{\frac{2}{n}} \left(\sum_{s \in N_G^{(n)}(t)} \frac{1}{h(s)} \right)$	$\frac{\tilde{M}_{2n}}{\sqrt{2n}}$	$\frac{\tilde{V}_{2n}}{\sqrt{2n}}$	$\frac{\tilde{T}_{sj(2n)}}{\sqrt{2n}}$
Limit law	V	$\frac{V}{2}$	$\frac{V}{\sqrt{2}}$	$\frac{V}{\sqrt{2}}$	V	$\frac{V}{2}$	V	V

where V has for distribution function:

$$F_V(x) = \mathbb{P}(V \leq x) = \sum_{-\infty < k < +\infty} (1 - k^2 x^2) \exp\left(-\frac{k^2 x^2}{2}\right).$$

References

- [1] D. Aldous, (1991) *The continuum random tree. II: An overview*, Stochastic analysis, Proc. Symp., Durham. UK 1990, Lond. Math. Soc. Lect. Note Ser. 167, 23-70.
- [2] D. Aldous, (1993) *The continuum random tree. III* Ann. of Probab. 21, No.1, 248-289.
- [3] D. Aldous, (1998) *Brownian excursion conditionned on its local time*, Elect. Comm. in Probab., **3** , 79-90.
- [4] P. Biane, M. Yor, (1987) *Valeurs principales associées aux temps locaux browniens*, Bull. Sci. Maths 111, 23-101.
- [5] N. De Bruijn, D. Knuth, S. Rice, (1972) *The average height of planted plane trees*, Graph Theory and Computing (R-C. Read, Ed),15-22, Acedemic Press, New York.
- [6] P. Chassaing, (1999) *How many questions are needed to find out the maximum of a Random Walk*, in press, Stoch. Proc. and Applns.
- [7] T.H. Cormen, E. Leiserson, R.L. Rivest, (1990) *Introduction to Algorithm*, MIT press.
- [8] P. Chassaing, J.F. Marckert, (1999) *Parking functions, empirical processes and the width of rooted labeled trees*, preprint Elie Cartan, Université Nancy I.
- [9] P. Chassaing, J.F. Marckert, M. Yor, (1999) *A Stochastically Quasi-Optimal Search Algorithm for the Maximum of the Simple Random Walk*, To appear in Ann. of App. Prob.
- [10] E. Csàki, A. Földes, (1987) *A note on the stability of the local time of a Wiener process*, Stoch. Proc. Appl. 25, 203-213.
- [11] E. Csàki, S.G. Mohanty, (1986) *Some joint distributions for conditional random walks*, Can. J. Stat. 14, 19-28.
- [12] C. Donati-Martin, (1999) *Some remarks about the identity in law for the Bessel bridge*
 $\int_0^1 \frac{ds}{r(s)} \stackrel{(law)}{=} 2 \sup_{s \leq 1} r(s)$, prepub.
- [13] M. Drmota, B. Gittenberger, (1997) *On the profile of random trees*, Random Structures Algorithms 10, no. 4, 421-451.
- [14] R.T. Durrett, D.L. Iglehart, (1977) *Functionals of Brownian meander and Brownian excursion*, Ann. Probab. 5, 130-135.
- [15] P. Flajolet, A. Odlyzko, (1982) *The average height of binary trees and other simple trees*, J. Comp. and Sys. Sci., Vol. 25, No.2.
- [16] I.I. Gikhman, A.V.Skorohod, (1969) *Introduction to the theory of random processes*, W.B. Saunders, Philadelphia.

-
- [17] Th. Jeulin, (1980) *Semi-martingales et grossissement d'une filtration* Lecture Notes in Mathematics, 833, Springer-Verlag.
- [18] D.P. Kennedy, (1976) *The distribution of the maximum Brownian excursion*, J. appl. Probab. 13, 371-376.
- [19] D.E. Knuth, (1973) *The Art of Computer Programming*, second edition, Volume 1, Addison-Wesley pub.
- [20] J.F. Le Gall, (1986) *Une approche élémentaire des théorèmes de décomposition de Williams*, Lect. Notes Math. 1204, Sémin. Prob. XX, p. 447-464, Springer.
- [21] J.F. Marckert, (1997) *An average case quasi optimal algorithm for the search of all the zeros of a random walk*, Preprint Elie Cartan, 44.
- [22] A.M. Odlyzko, (1995) *Search for the maximum of a random walk*, Ran. Struct. Alg., Vol. 6, p. 275-295.
- [23] A. Renyi, G. Szekeres, (1967) *On the height of trees*, J. Aust. Math. Soc. 7, 497-507.
- [24] D. Revuz, M.Yor, (1990) *Continuous martingales and Brownian motion*, Springer-Verlag.
- [25] P. Révész, (1990) *Random walk in random and non-random environments*, World Scientific. xiv.
- [26] R. Sedgewick, Ph. Flajolet, (1996) *An introduction to the analysis of algorithms*, Amsterdam: Addison-Wesley, xv, 492 p.
- [27] L. Takács, (1991) *A Bernoulli excursion and its various applications*, Adv. Appl. Prob., 23, 557-585
- [28] L. Takács, (1991) *On the distribution of the number of vertices in layers of random trees*, J. Appl. Math. Stochastic Anal. 4, No.3, 175-186.
- [29] L. Takács, (1992) *On the total heights of random rooted trees*, J. Appl. Probab. 29, No.3, 543-556.
- [30] L. Takács, (1992) *Random walk processes and their various applications*, Pro. The. App., Essays to the Mem. of J. Mogyorodi, Math. Appl. 80, 1-32 .
- [31] L. Takács, (1993) *Limit distributions for queues and random rooted trees*, J. Appl. Math. Stoch. Ana., **6** , No.3, p.189 - 216.

CHAPITRE IV

Parking functions, empirical processes and the width of rooted labeled trees ¹

Philippe Chassaing & Jean-François Marckert

Institut Elie Cartan, INRIA, CNRS
 Université Henri Poincaré,
 BP 239, 54 506 Vandoeuvre Cedex, France.
 chassain@iecn.u-nancy.fr, marckert@iecn.u-nancy.fr

Abstract. This paper provides tight bounds for the moments of the width of rooted labeled trees with n nodes, answering an open question of Odlyzko and Wilf (1987). Essential to the proof are the one to one correspondences between trees and parking functions, and also a coupling between parking functions and empirical processes of mathematical statistics. Our result turns out to be a consequence of the strong convergence of the empirical process to the Brownian bridge (Komlos, Major and Tusnady, 1976).

1 Introduction

1.1 Rooted labeled trees

A size n labeled tree is a connected graph with set of vertices $\{1, 2, 3, \dots, n\}$, and with $n - 1$ edges. If we specify that one vertex is to be the root we have a *rooted* labeled tree. According to Cayley (1889) the number of such trees is n^{n-1} .

For ω chosen at random in the set Ω_n of size n rooted labeled trees, let $Z_k^{(n)}(\omega)$ denote the number of nodes at distance k from the root of ω , and let $H_n(\omega)$ denote the maximum distance of a node from the root, the *height* of ω ; $(Z_k^{(n)})_{k \geq 0}$ is the *profile* of the tree. The *width* $W_n(\omega)$ is defined by:

$$W_n = \max_{0 \leq k \leq H_n} Z_k^{(n)}.$$

Odlyzko and Wilf (1987) used a surprising Perron-Frobenius-like theory to derive asymptotics for the cumulative function of W_n . They also proved that

$$C_1 \sqrt{n} \leq E(W_n) \leq C_2 \sqrt{n \log n},$$

and left the first term in the asymptotic of $E(W_n)$ as an open question. Denote by $l(t)$ the local time of the normalized Brownian excursion e . at level t , i.e.

$$l(t) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^1 I_{[t, t+\varepsilon]} e(s) ds.$$

Aldous (1992) conjectured that $t \mapsto Z_{\lfloor \sqrt{nt} \rfloor}^{(n)} / \sqrt{n}$ would converge weakly, as a stochastic process, to $t \mapsto 1/2 \ell(t/2)$. Aldous's conjecture was settled in 1997 by Drmota and Gittenberger. As noted by these last authors, their result entails the weak convergence of W_n / \sqrt{n}

¹**Key words.** Rooted labeled trees, moment, width, Brownian excursion, empirical processes, hashing with linear probing, parking.

A.M.S. Classification. 05C05, 60J65, 60J80, 62G30.

to the maximum m of the Brownian excursion, as $1/2 \ell(t)$ is itself a normalized Brownian excursion under a change of time (see Jeulin [23], p. 264): for an account of these and other nice results about Brownian excursion, see [7]. Previously, the weak convergence of W_n/\sqrt{n} to m was proved directly by Takács (1993).

However weak convergence does not answer completely the question of Odlyzko & Wilf, as it does not yield convergence of the first moment, and even less the speed of this convergence. The aim of our paper is to fill this gap. Note that the results of Aldous, Drmota & Gittenberger, and Takács are actually about *general* simple trees. Rooted labeled trees are a special case of simple trees, but an important one (see [28] p. 389-390, [33], [34] or [44]).

1.2 Results

It is part of history of probability that:

$$\Pr(m \leq x) = \sum_{-\infty < k < +\infty} (1 - 4k^2 x^2) e^{-2k^2 x^2}$$

and, for $r > 1$:

$$\mu_r = E(m^r) = 2^{-r/2} r(r-1) \Gamma\left(\frac{r}{2}\right) \zeta(r), \quad \mu_1 = \sqrt{\frac{\pi}{2}},$$

(see [12], [13], [27], [32], [7], [46]). We shall say that m is *theta-distributed* by reference to Jacobi's Theta function [6] [8]. Incidentally, it is also well known that theta-distributed random variables occur as a limit for the height of trees: see Renyi and Szekeres (1967) for rooted labeled trees, Flajolet and Odlyzko (1982) for general simple trees. Aldous (1993) finally explains both Flajolet and Odlyzko's results and Takács' results (1992) about the total path length of simple trees, by proving convergence of the rescaled search depth process of a simple tree, to the normalized Brownian excursion (see Section 2 of [1]).

We prove that, at the price of enlarging the probability spaces, there exists a sequence of theta-distributed random variables m_n satisfying:

Theorem 1.1 For $p \geq 1$,

$$E(|n^{-1/2} W_n - m_n|^p) \leq C_p n^{-p/4} (\log n)^{p/2}.$$

in which C_p does not depend on n .

yielding, as a corollary:

Theorem 1.2 For $p \geq 1$,

$$|E(n^{-p/2} W_n^p) - E(m^p)| \leq C'_p n^{-1/4} \sqrt{\log n}$$

The proof uses essentially the breadth first search random walk (see Spencer (1997) or Aldous (1997)), also exploited in Takács (1993): Takács' approach leads to convergence of moments of the width for binary trees, or general unlabeled trees, due to a clever use of the ballot theorem.

For rooted labeled trees, Takács' approach needs an additional ingredient: we use a close connection between rooted labeled trees and *empirical processes* of mathematical statistics, which, we believe, has interest in itself. For instance, it gives an $O(n)$ algorithm for the generation of a random rooted labeled tree, an alternative to the $O(n)$ algorithm using Prufer-Knuth's correspondence (see Moon [33] Chap. 2, Knuth [28] p.389-391, or Alonso and Schott [5] p. 25-30). It also gives an alternative proof of the recent result of Flajolet, Viola & Poblete (1998), about the total displacement for linear probing hashing.

1.3 Motivations

One of the motivations of Odlyzko and Wilf, when they study the width of labeled trees, is to give a tight estimate for the average bandwidth of this class of tree.

Denote by $E(T)$ (resp. $V(T)$) the sets of edges (resp. vertices) of a tree T . Assume $\#V(T) = n$ and let Q_n be the set of one-to-one mappings between $V(T)$ and $\{1, \dots, n\}$. The *bandwidth* $\beta(T)$ of T is defined by:

$$\beta(T) = \min_{f \in Q_n} \left(\max_{(u,v) \in E(T)} |f(u) - f(v)| \right)$$

Let $W(T)$ (resp. $H(T)$) denote the width (resp. the height) of T . Height, width and bandwidth of T are related by:

Proposition 1.3

$$\frac{n-1}{2 H(T)} \leq \beta(T) \leq 2 W(T).$$

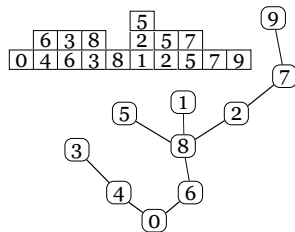
For a proof see [35]. Analysis of the width is also a first step towards the study of the average complexity of drawing trees nicely (see [48]).

2 Rooted labeled trees, parking functions and empirical processes

2.1 The breadth first search random walk

A first correspondence between parking functions and acyclic functions was discovered by Schutzenberger (1968) (see also [20] [21]). For seek of convenience we shall describe the merely equivalent connection between rooted labeled trees and parking functions, borrowing an important ingredient to [47], [3], among others: the *breadth first search random walk*. We start the breadth first search of the rooted labeled tree with vertex 0, that is, we maintain a *queue* Q that is initially (0). At each stage we take the vertex x at the head of the queue, remove x from the queue, and add all "new" neighbors of x at the end of the queue, in increasing order.

At step 0, the search produces the set A_0 of neighbors of vertex 0, so that after step 0 the queue contains exactly the elements of A_0 , but not 0 anymore. At step 1, the search produces the set A_1 of new neighbors of the smallest element x in A_0 , so that after step 1 the



queue contains $A_0 \cup A_1 - \{x\}$. Let A_k denote the set of new elements in the queue after step k , and let $x_k = \#A_k$. A labeled tree with vertices $\{0, 1, 2, 3, \dots, n\}$, rooted at 0, is described by a sequence of disjoint sets $(A_i)_{0 \leq i \leq n}$, whose union is $\{1, 2, \dots, n\}$, and whose cardinalities $x_i = \#A_i$ satisfy the following set of constraints:

$$\begin{aligned}
 x_0 &\geq 1, \\
 x_0 + x_1 - 1 &\geq 1, \\
 &\dots \\
 x_0 + x_1 + \dots + x_k - k &\geq 1, \\
 &\dots \\
 x_0 + x_1 + \dots + x_{n-1} - n + 1 &\geq 1, \\
 x_0 + x_1 + \dots + x_n - n &= 0.
 \end{aligned} \tag{2.1}$$

Constraints 2.1 are necessary and sufficient conditions for a tree to be connected, or for the queue to become empty only after step n :

$$y_k = x_0 + x_1 + \dots + x_k - k$$

denotes the number of vertices in the queue after step k . We have clearly

$$y_k - y_{k-1} = x_k - 1.$$

Note that $Z_1^{(n+1)} = y_0$, $Z_2^{(n+1)} = y_{Z_1^{(n+1)}}$, and more generally

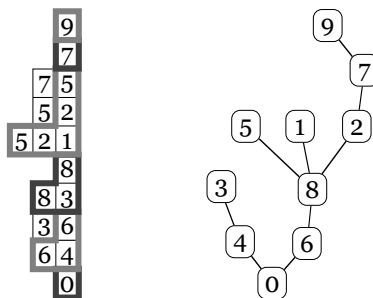
$$Z_{k+1}^{(n+1)} = y_{Z_1^{(n+1)} + Z_2^{(n+1)} + \dots + Z_k^{(n+1)}}.$$

At step $Z_1^{(n+1)} + Z_2^{(n+1)} + \dots + Z_k^{(n+1)}$, we explore the last vertex at a distance k from the root, and the queue contains exactly the vertices at distance $k + 1$ from the root. Actually this is Kendall's embedding of a Galton-Watson process in the process of queue lengths, between two successive idle periods, when studying a single-server queue (see [25], [50], Section 4, or [38]), with only one difference: we assume that exactly $n + 1$ clients arrive in this interval.

The main point is that the width W_n , defined by

$$W_n = \max_k Z_k^{(n+1)},$$

is the maximum of a sample of y_i , and our proof of the convergence of moments will be in two steps, one of them being that this sample is "representative", in the sense that



Theorem 2.1 For any $p \geq 1$

$$\|W_n - \max_k y_k\|_p = O(n^{1/4} \sqrt{\log n}),$$

due to slow variation of the sequence $y = (y_k)_{0 \leq k \leq n}$ (see Section 3).

For the second step, quoting Spencer (1997), we note that if a rooted labeled tree (or a parking function, as we shall see later) ω is drawn at random, the corresponding sequence $(x_i(\omega))_{0 \leq i \leq n}$ has the distribution of independent Poisson random variables with mean value 1 conditioned to satisfy constraints 2.1: the number of n -tuples $(A_i)_{0 \leq i \leq n}$ with cardinalities $(x_i)_{0 \leq i \leq n}$,

$$\frac{n!}{x_1! x_2! \dots x_n!},$$

is proportional to the product of Poisson probabilities $e^{-1}/x_i!$. In other words, the corresponding unlabeled tree is a Galton-Watson tree with Poisson(1) progeny, constrained to have $n + 1$ nodes. In terms of Galton-Watson processes, A_k is the progeny of the k^{th} node visited by the breadth first search (see also [1], Section 2.1, where the lexicon between simple trees and Galton-Watson trees is clearly stated and explained).

As consequences, the sequence $y = (y_k)_{0 \leq k \leq n}$ is the discrete excursion with length n associated with a random walk whose increments are $x_i - 1$, and, according to [24], [17]:

$$\frac{1}{\sqrt{n}}(y_{[nt]})_{0 \leq t \leq 1} \xrightarrow{\text{weakly}} (e(t))_{0 \leq t \leq 1},$$

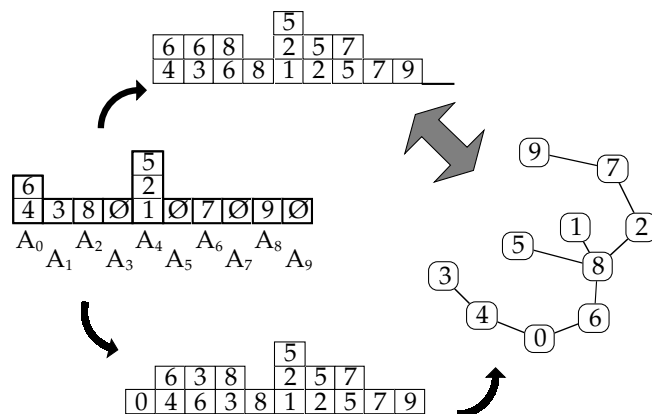
in which e denote a normalized Brownian excursion. Thus

$$\frac{\max_k y_k}{\sqrt{n}} \xrightarrow{\text{weakly}} m = \max_{0 \leq t \leq 1} e(t).$$

Instead of weak convergence we want the more demanding convergence of moments: to this aim, we describe in the next subsections the connection between rooted labeled trees and empirical processes, through parking functions.

2.2 Rooted labeled trees and parking functions

In hashing with linear probing, or parking (see [29] Section 6.4, [40] or [19]), we consider the case with n cars and $n + 1$ places $\{0, 1, 2, \dots, n\}$, car c_k parking on place p_k if p_k is still



empty, that is, if a car with a smaller index did not park on place p_k before. Else car c_k tries places $(p_k + 1) \bmod n + 1, (p_k + 2) \bmod n + 1, \dots$. We consider *confined* sequences (see Knuth (1998)), or *parking functions* (see Foata and Riordan (1974), or Françon (1975)), that is sequences $\omega = (p_k)_{1 \leq k \leq n}$ such that the last empty place is place n . Such a parking function ω is alternatively characterized by a sequence $(A_i)_{0 \leq i \leq n}$, as previously for rooted labeled trees, but here:

$$A_i = \{k \mid p_k = i\}$$

is the set of cars whose first try is place i (x_i still denotes $\#A_i$).

In this interpretation, let y_k denote the number of cars that tried, successfully or not, to park on place k . We have $y_0 = x_0$, while for $k \geq 1$:

$$y_k = y_{k-1} - 1 + x_k,$$

since place $k - 1$ is occupied by some car c_i and among the y_{k-1} cars that visited place $k - 1$, only c_i won't visit place k . Constraints (2.1) just say that the only empty place is place n . As a consequence, each of the $(n + 1)^{n-1}$ sequences $(A_i)_{0 \leq i \leq n}$ defines at the same time a unique rooted labeled tree with size $n + 1$ and a unique confined parking function for n cars on $n + 1$ places.

Incidentally, in this interpretation, the *total displacement* $D(n, n - 1)$ for hashing with linear probing, studied in [29], [31], and [19], can be written

$$D(n, n - 1) = -n + 1 + n \int_0^1 y_{\lfloor nt \rfloor} dt,$$

so, using the weak convergence of $\frac{y_{\lfloor nt \rfloor}}{\sqrt{n}}$ to the normalized Brownian excursion, we obtain:

$$\frac{D(n, n - 1)}{n\sqrt{n}} \xrightarrow{\text{weakly}} \int_0^1 e(t) dt.$$

recovering partly the result of [19], concerning the convergence of the total displacement towards the Airy law. However, in order to obtain the convergence of moments, as in [19], we need the connection between parking functions and empirical processes.

2.3 Short and loose historic of empirical processes

Consider a sequence of independent random variables $(U_i)_{1 \leq i \leq n}$, each of them uniform on $[0, 1]$. Let $F_n(t)$ denote the *empirical distribution* for $(U_i)_{1 \leq i \leq n}$, defined for $t \in [0, 1]$ by:

$$F_n(t) = \frac{\#\{i | 1 \leq i \leq n \text{ and } U_i \leq t\}}{n}$$

The foundations of mathematical statistics were layed by Glivenko & Cantelli, Kolmogorov & Smirnov, Doob, Donsker and many others, when they studied how the empirical distribution function converges towards the distribution function $F(t) = t$ of the uniform law (for historical background, see [41] p. 117-118, [15], or the encyclopedic Shorack & Wellner). Specially, following a suggestion of Doob, Donsker (and many others after him) studied the *empirical process*:

$$\alpha_r(t) = \sqrt{r}(F_r(t) - F(t)).$$

and established that:

Theorem 2.2

$$(\alpha_r(t))_{t \in [0,1]} \xrightarrow{\text{weakly}} (b(t))_{t \in [0,1]},$$

$b(t)$ being the normalized Brownian bridge.

Many relevant statistics (such as sample mean, variation or median ...) are functionals of F_n , and their speed of convergence towards the true mean, median, etc ... , is revealed by the empirical process, what makes it fundamental in mathematical statistics (see [45]).

Thus the error term is of order $O(\sqrt{r})$, more precisely, the error is $\sqrt{r}T_r$, in which $T_r = \sup\{|\alpha_r(t)| \mid t \in [0, 1]\}$, and T_r converges to the Kolmogorov (1931) statistic $T = \sup_{0 \leq t \leq 1} |b(t)|$:

$$\begin{aligned} \Pr(T \leq x) &= \sum_{-\infty}^{+\infty} (-1)^n e^{-n^2 x} \\ &= \prod_{n=1}^{+\infty} \frac{1 - e^{-nx}}{1 + e^{-nx}}. \end{aligned}$$

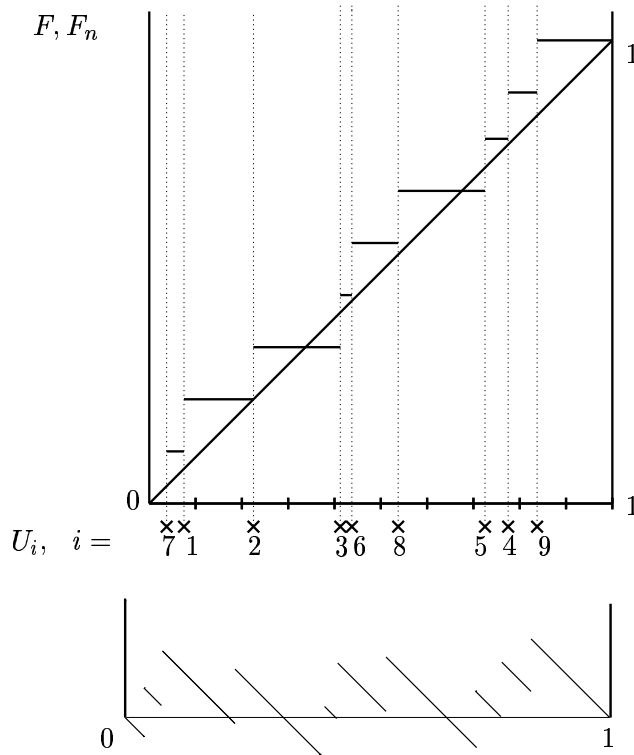
Incidentally, the connection between Kolmogorov distribution and theta-distribution is still incompletely understood (see [12], [13], [32], [7]).

The next major step in the study of the speed of convergence was:

Theorem 2.3 (Komlos, Major & Tusnady, 1976) *Given U_1, U_2, \dots uniform on $[0, 1]$ and independent, there exists a sequence $(b_n)_{n \geq 1}$ of normalized Brownian bridges such that for all n and x ,*

$$\Pr \left(\sup_{0 \leq t \leq 1} |\alpha_n(t) - b_n(t)| \geq \frac{A \log n + x}{\sqrt{n}} \right) \leq M e^{-\mu x},$$

where A, M and μ are positive absolute constants.



In other terms, we can write:

$$F_n(t) = F(t) + \frac{b_n(t)}{\sqrt{n}} + \frac{c_n(t)}{n},$$

in which $c_n(t)$ denotes $\sqrt{n}(\alpha_n(t) - b_n(t))$, and satisfies:

$$\Pr\left(\sup_{0 \leq t \leq 1} |c_n(t)| \geq A \log n + x\right) \leq M e^{-\mu x},$$

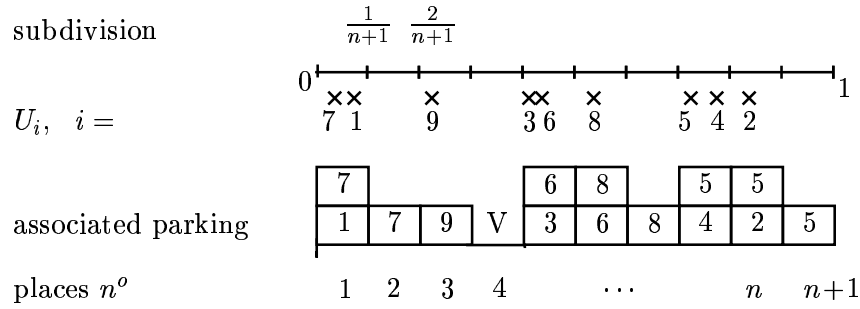
with the consequence that:

$$\left\| \|\alpha_n - b_n\|_\infty \right\|_p = O\left(\frac{\log n}{\sqrt{n}}\right).$$

KMT's Theorem is the last ingredient we need to estimate $\|W_n\|_p$.

2.4 Parking functions and empirical processes

Let us realize hashing with linear probing, or parking, for a table with size $n + 1$, with the help of a sequence of independent random variables $(U_i)_{1 \leq i \leq n}$, each of them uniform on $[0, 1]$: let the first try of car c_i be the place $p_i = \lceil (n + 1)U_i \rceil$ (see figure). In this setting, the parking functions, or hashing sequences, $\omega = (p_k)_{1 \leq k \leq n}$, are not assumed anymore to be “confined”, and the $(n + 1)^n$ parking functions are clearly equiprobable. Let $V(\omega)$ denote the last empty place: V is random uniform on $Z_n = \{0, 1, 2, \dots, n\}$.



A natural question arises: does V depend on the chronology or does it depend only on the sequence $(x_i)_{0 \leq i \leq n}$? and is there a convenient way of computing it, without running all the process of dispatching of cars on places? The answer is given in Proposition 2.4, and yields a surprisingly tight coupling between the sequence $(y_k)_{0 \leq k \leq n}$ and the empirical process α_n generated with the help of $(U_i)_{1 \leq i \leq n}$. Let $0 \leq T(n) \leq n$ be defined by:

$$\alpha_n\left(\frac{T(n)}{n+1}\right) = \min_{0 \leq k \leq n} \alpha_n\left(\frac{k}{n+1}\right).$$

Proposition 2.4 $T(n)$ is uniformly distributed on $\{0, 1, 2, \dots, n\}$, and

$$T(n) = V.$$

Proof: Let $b(n, i) = \sqrt{n} \alpha_n(i/n + 1)$. For $k = 1, 2, \dots, n + 1$, we have:

$$\begin{aligned} b(n, V+k) - b(n, V) &= x_{V+1} + x_{V+2} + \dots + x_{V+k} - k \frac{n}{n+1} \\ &= y_{V+k} + k - 1 - k \frac{n}{n+1} \\ &= y_{V+k} + \frac{k}{n+1} - 1, \end{aligned} \tag{2.2}$$

the second equality, as already seen in Subsection 2.2, due to the fact that $y_{V+1} = x_{V+1}$, but for $k \neq V$, $y_{k+1} = y_k - 1 + x_{k+1}$. For $k = 1, 2, \dots, n$, place $V+k$ is not empty, yielding that $y_{V+k} \geq 1$. As a consequence, the last term in (2.2) is positive, so that $T(n)$ is well defined and is equal to V . \square

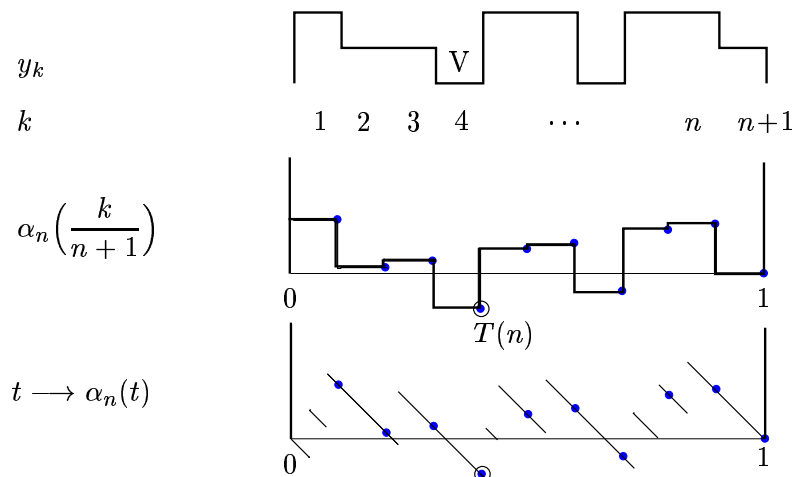
As a byproduct of (2.2), we obtain:

Corollary 2.5

$$(y_{V+k})_{0 \leq k \leq n} = (w_k)_{0 \leq k \leq n}$$

in which

$$w_k = \frac{n+1-k}{n+1} + \sqrt{n} \left(\alpha_n\left(\left\{\frac{k+T(n)}{n+1}\right\}\right) - \alpha_n\left(\frac{T(n)}{n+1}\right) \right).$$



If we define, for $\omega = (p_k)_{1 \leq k \leq n}$,

$$T\omega = (1 + p_k)_{1 \leq k \leq n},$$

the sequence $\tilde{y} = (y_{V(\omega)+k}(\omega))_{0 \leq k \leq n}$ is invariant under T , while $V(T\omega) = 1 + V(\omega)$. It follows that V is uniform and independent of \tilde{y} , so that, on one hand, the conditional distribution of \tilde{y} given that $V = n$ is the same as the unconditional distribution of \tilde{y} . On the other hand, the conditional distribution of \tilde{y} given that $V = n$ is the distribution of the sequence $(y_k)_{0 \leq k \leq n}$ under the hypothesis of equiprobability of *confined* parking schemes, as in Subsection 2.2. We obtain finally that, for rooted labeled trees, or for confined parking functions, that is, in the setting of Subsection 2.2:

Theorem 2.6

$$(y_k)_{0 \leq k \leq n} \stackrel{(law)}{=} (w_k)_{0 \leq k \leq n}.$$

This connection between parking functions and empirical processes is close in spirit to a coding of parking functions given page 14 of [20]. Besides that, the correspondence trees-parking schemes of Subsection 2.2 is close to the one explained ibidem page 17. This explicit coupling also reminds of similarities between the Cayley tree function, tightly related with parking functions, or the Borel distribution (see [4] Section 2.2, or [14] Section 2.7) on one hand, and expressions omnipresent in Chap. 9 of Shorack & Wellner about empirical processes, on the other hand (see Exercise 1, p.345 or formulas of Birnbaum & Pyke p.386). For more on these surprising connections, see a recent preprint by Pitman and Stanley [39].

2.5 Generation of a random rooted labeled tree

In other words, Theorem 2.6 amounts to say that one can generate a random confined parking function, or a random rooted labeled tree, with the help of (U_1, \dots, U_n) , computing first $T(n)$ and setting:

$$A_i = \{k \mid \lfloor (n+1)U_k \rfloor = T(n) + i\}.$$

This algorithm does not compare unfavorably to the algorithm based on the Prufer-Knuth's correspondence between labeled trees and n -tuples of $\{0, 1, \dots, n\}^n$ (see Moon [33] Chap. 2, Knuth [28] p.389-391, or Alonso and Schott [5] p. 25-30), as it takes $O(n)$ to compute the subsets A_i and $O(n)$ to draw the corresponding labeled tree, given the subsets A_i .

In the next Sections we assume that the random rooted labeled trees are generated using this algorithm. As a consequence, in Theorem 2.6, it is an equality between random variables that holds, and not merely an equality between probability distributions.

3 Bounds for moments

Let us build a sequence m_n of theta-distributed random variables such that

Theorem 3.1 *For any $p \geq 1$*

$$\left\| m_n - \frac{\max_k y_k}{\sqrt{n}} \right\|_p = O\left(\frac{\log n}{\sqrt{n}}\right).$$

Together with Theorem 2.1, it will yield the desired bounds for $\|W_n\|_p$.

3.1 Proof of Theorem 3.1

The next step uses strong convergence: Theorem 2.3 asserts the existence, on the same probability space as $(U_k)_{k \geq 1}$, of a sequence of Brownian bridges $(b_n)_{n \geq 1}$, that approximate closely the sequence $(\alpha_n)_{n \geq 1}$. We use:

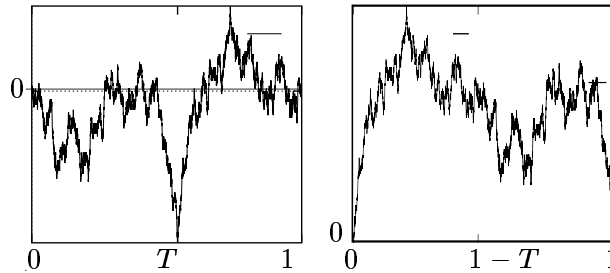


Figure 3.1 *Vervaat's decomposition*

Theorem 3.2 (Vervaat, 1979) *Let $b = (b(t))_{0 \leq t \leq 1}$ be a Brownian bridge, and let T be the almost surely unique point such that $b(T) = \inf_{0 \leq t \leq 1} b(t)$. Then T is uniform and $e = (e(t))_{0 \leq t \leq 1}$, defined by $e(t) = b(\{T + t\}) - b(T)$, is a normalized Brownian excursion, independent of T ,*

to build from $(b_n)_{n \geq 1}$ a sequence of normalized Brownian excursions $(e_n)_{n \geq 1}$, whose maximums are precisely the random variables

$$\begin{aligned} m_n &= \sup_{0 \leq t \leq 1} b_n(t) - \inf_{0 \leq t \leq 1} b_n(t) \\ &= \sup_{0 \leq t \leq 1} e_n(t) \end{aligned}$$

of Theorem 1.1. Set:

$$\begin{aligned} D_n &= \sqrt{n} \sup_{0 \leq t \leq 1} \left| \alpha_n \left(\frac{\lfloor (n+1)t \rfloor}{n+1} \right) - \alpha_n(t) \right|, \\ C_n &= \sup_{0 \leq t \leq 1} |c_n(t)|. \end{aligned}$$

Due to the construction of y_k in Subsection 2.5, we have

$$\begin{aligned} \left| \frac{\max_k y_k}{\sqrt{n}} - m_n \right| &\leq \left| \frac{\max_k y_k}{\sqrt{n}} - \left(\sup_{0 \leq t \leq 1} \alpha_n(t) - \inf_{0 \leq t \leq 1} \alpha_n(t) \right) \right| + \frac{2C_n}{\sqrt{n}} \\ &\leq \frac{1 + 2D_n + 2C_n}{\sqrt{n}}, \end{aligned}$$

in which the second inequality follows from Theorem 2.6. Due to Theorem 2.3, C_n belongs to any L_p , and it satisfies:

$$\|C_n\|_p \leq K_p \log n.$$

The same is true for D_n , as we have:

Proposition 3.3 *For suitable A , C and K ,*

$$\Pr(D_n \geq C \log n + x) \leq A n^{1-KC} e^{-Kx},$$

yielding Theorem 3.1. \square

Proof of Proposition 3.3. Let $Y_{k,n}$ denote the number of cars whose first try is place k , and set

$$Z_n = \max_{0 \leq k \leq n} Y_{k,n}.$$

We have:

$$\begin{aligned} \left| \alpha_n \left(\frac{\lfloor (n+1)t \rfloor}{n+1} \right) - \alpha_n(t) \right| &\leq \frac{\sqrt{n}}{1+n} + \frac{Z_n}{\sqrt{n}} \\ &\leq \frac{1 + Z_n}{\sqrt{n}}, \end{aligned}$$

and:

$$\begin{aligned} \Pr(Z_n \geq C \log n + x) &\leq n \Pr(Y_{1,n} \geq C \log n + x) \\ &\leq n E[\exp(KY_{1,n})] \exp(-KC \log n) e^{-Kx} \\ &\leq n \left(1 + \frac{e^K - 1}{n+1} \right)^n \exp(-KC \log n) e^{-Kx} \\ &\leq A n^{1-KC} e^{-Kx}, \end{aligned}$$

the third inequality due to the fact that $Y_{k,n}$ follows a binomial distribution with parameters $(n, 1/n+1)$. \square

Note that if $D(n, n-1)$ denotes the total displacement of the parking scheme derived from U_1, U_2, \dots, U_n , we have

$$D(n, n-1) = -n + 1 + n \int_0^1 y_{[nt]} dt.$$

Thus

$$\left| \frac{D(n, n-1)}{n\sqrt{n}} - \int_0^1 e_n(t) dt \right| \leq 2 \frac{1 + D_n + C_n}{\sqrt{n}}.$$

This leads to an error term $\frac{\log n}{\sqrt{n}}$ in the convergence of the k^{th} moment of the total displacement to the k^{th} moment of the Airy law, error term that does not compare too unfavorably to the error term $(1/\sqrt{n})$ obtained by Flajolet et al. [19].

3.2 Proof of Theorem 2.1

Set

$$M_n = \max_k y_k.$$

The aim of this section is to bound the difference between M_n and W_n . Essentially, we follow the line of proof of [50] (see formula 63, page 200). We improve Takács's bounds with the help of Petrov's Theorem 3.5: let $\Omega_c(n)$ be the set of sequences $y = (y_k)_{k=0, \dots, n}$ that satisfy

$$|y_{m+k} - y_m| \leq c\sqrt{k \log n}$$

whenever $k \geq 0$, $m \geq 0$ and $m+k \leq n$. We have:

Proposition 3.4 *Given a positive number α there exists a constant c , not depending on n , such that*

$$1 - \Pr(y \in \Omega_c(n)) = o(n^{-\alpha}).$$

The proof of Proposition 3.4 is given at the end of the Subsection. Recall that $Z_i^{(n+1)} = y_{m(i)}$, in which $m(1) = 0$ and $m(i+1) = m(i) + Z_i^{(n+1)}$. Consider an integer k such that $y_k = M_n$: for some i , $k \in [m_i, m_{i+1}[$, so that:

$$\begin{aligned} 0 \leq M_n - W_n &\leq M_n - Z_i^{(n+1)} \\ &\leq I_{\Omega_c(n)} c\sqrt{(k - m(i)) \log n} + n(1 - I_{\Omega_c(n)}) \\ &\leq I_{\Omega_c(n)} c\sqrt{Z_i^{(n+1)} \log n} + n(1 - I_{\Omega_c(n)}) \\ &\leq I_{\Omega_c(n)} c\sqrt{M_n \log n} + n(1 - I_{\Omega_c(n)}) \\ &\leq I_{\Omega_c(n)} c\sqrt{c\sqrt{n} \log^{3/2} n} + n(1 - I_{\Omega_c(n)}), \end{aligned}$$

and, by a suitable choice of c in Proposition 3.4:

$$E(|W_n - M_n|^p) = O(n^{p/4}(\log n)^{3p/4}).$$

This last estimate holds true under the general hypothesis of finite exponential moments for the progeny, but, for Theorem 2.1, we need to decrease the exponent of $\log n$ to $p/2$. The last but one inequality gives:

$$E(|W_n - M_n|^p) \leq \tilde{c}E(M_n^{p/2}) \log^{p/2} n + n^p(1 - \Pr \Omega_c(n)).$$

Theorem 2.6 entails

$$M_n \leq \sqrt{n}(1 + 2 \sup_t |\alpha_n(t)|).$$

Finally, owing to Dvoretzky-Kiefer-Wolfowitz inequality:

$$P(\sup_t |\alpha_n(t)| \geq x) \leq 4\sqrt{2} \exp(-2x^2)$$

(see [45] p.354-356), we obtain:

$$E(M_n^{p/2}) \leq \tilde{K}_p n^{p/4}. \quad \square$$

Proof of Proposition 3.4. Let $(z_k)_{0 \leq k \leq n}$ be a sequence of independent random variables, Poisson-distributed with mean 1, and let $t = (t_k)_{0 \leq k \leq n}$ be the random walk with increments $z_k - 1$. Let $\Delta(n)$ denote the set of sample paths y that satisfy constraints (2.1). As a consequence of Spencer's key remark,

$$\begin{aligned} \Pr(y \notin \Omega_c(n)) &= \Pr(t \notin \Omega_c(n) \mid t \in \Delta(n)) \\ &\leq \frac{\Pr(t \notin \Omega_c(n))}{\Pr(t \in \Delta(n))}. \end{aligned}$$

Due to Otter's formula ([36], see also [50], [38]), we have:

$$\Pr(t \in \Delta(n)) = \frac{1}{n} \Pr(t_n = 0),$$

so due to the standard local limit theorem (see for instance Ibragimov & Linnik Chap. 4, Th. 4.2.1) we obtain:

$$\Pr(t \in \Delta(n)) = \Theta(n^{-3/2}).$$

Thus we are to prove Proposition 3.4 only for the unconditioned random walk t , but this is a consequence of:

Theorem 3.5 (Petrov, 1954) *Let Y_k be a random walk with increments X_k satisfying simultaneously*

- $E(X_k) = 0$, and

- for some positive constant α , $E(e^{\alpha|X_k|}) < +\infty$,

then:

i) there exist two positive real constants g and T such that

$$E(\exp(tX_1)) \leq \exp(gt^2) \text{ for } |t| < T,$$

ii) for $(Y_k)_{k \geq 1}$ defined as above, we have

$$\begin{aligned} \Pr(|Y_k| \geq x) &\leq 2 \exp\left(-\frac{x^2}{4kg}\right) \quad \text{if } 0 \leq x \leq kgT, \\ &\leq 2 \exp\left(-\frac{Tx}{2}\right) \quad \text{if } x \geq kgT. \end{aligned}$$

(see [37] p.52-55). Theorem 3.5 yields:

$$\begin{aligned} \Pr(t \notin \Omega_c(n)) &\leq \Pr(\exists m, k \mid |t_{m+k} - t_m| \geq c\sqrt{k \log n}) \\ &\leq n \sum_{k=1}^n \Pr(|t_k| \geq c\sqrt{k \log n}) \\ &\leq n \sum_{k=1}^{\frac{c^2 \log n}{4T^2 g^2}} 2 \exp\left(-\frac{c^2 \log n}{4g}\right) + n \sum_{k=\frac{c^2 \log n}{4T^2 g^2}}^n 2 \exp(-T\sqrt{k \log n}/2) \\ &\leq \frac{c^2 \log n}{2T^2 g^2} n^{1-c^2/4g} + 2n^{1-c/4g}. \end{aligned}$$

For c large enough, the last term is $o(n^{-\alpha})$. □

4 Concluding remarks

Convergence of moments of the width extends easily to binary trees: the breadth-first search random walk for a binary tree is a ruin sequence, and in the correspondence between ruin sequences and *general trees* (see [44], p. 268-271), the maximum of the ruin sequence is within a constant of the height of the corresponding general tree. Thus we can use

Theorem 4.1 (*Flajolet-Odlyzko, 1982*) *The r^{th} moment of the height of a general tree with n nodes is asymptotic to $2^{-r/2} n^{r/2} \mu_r$,*

instead of Theorem 3.1, to obtain convergence of moments of the width of binary trees. However, compared with Theorem 1.2, we lose the speed of convergence.

According to [50], asymptotics for the moments of the width of binary trees, or of general trees, can also be obtained through closed form formulas for the distribution function of the maximum of the breadth-first search random walk (see (48) and (49) p. 197 of [50]), using results of [9], [26], and a weaker form of Theorem 2.1 (see p. 200-201 of [50]). In a work in

progress, Cyril Banderier and Philippe Flajolet study carefully asymptotics of the maximum of the breadth-first search random walk for general simple trees with finite degree. Together with Theorem 2.1, their work gives asymptotics for moments of the width of general simple trees with finite degree.

The results of Subsections 2.4 and 3.1 are extended in [11] in order to study the “emergence of a giant block” of consecutive cars for a parking function. An interesting phenomenon of coalescence of blocks appears, reminiscent of the coalescence of connected components observed in [3] for the random graph process during its phase transition. From Subsection 2.4, a moment of thought suggests that the width of the largest block is $n - o(n)$ if the number of empty places is $o(\sqrt{n})$, while it is simultaneously $n - \Theta(n)$ and $\Theta(n)$ if the number of empty places is $\Theta(\sqrt{n})$ (see [11] for more precise statements).

References

- [1] D. Aldous, (1991) *The continuum random tree. II: An overview*. Stochastic analysis, Proc. Symp., Durham/UK 1990, Lond. Math. Soc. Lect. Note Ser. 167, 23-70.
- [2] D. Aldous, (1993) *The continuum random tree. III* Ann. of Probab. 21, No.1, 248-289.
- [3] D. J. Aldous, (1997) *Brownian excursions, critical random graphs and the multiplicative coalescent*. Ann. Probab. 25, No.2, 812-854.
- [4] D. J. Aldous, (1999) *Deterministic and Stochastic Models for Coalescence (Aggregation, Coagulation): A Review of the Mean-Field Theory for Probabilists*. Bernoulli 5 3-48.
- [5] L. Alonso, R. Schott, (1995) *Random generation of trees*. Kluwer Academic Publishers, Boston, MA.
- [6] R. Bellman, (1961) *A brief introduction to theta functions*. Holt, Rinehart and Winston, New York.
- [7] P. Biane, M. Yor, (1987) *Valeurs principales associées aux temps locaux browniens*, Bull. Sci. Maths 111, 23-101.
- [8] J.M. Borwein, P.B. Borwein, (1998) *Pi and the AGM. A study in analytic number theory and computational complexity*. John Wiley & Sons, Inc., New York.
- [9] N.G. de Bruijn, D.E. Knuth, S.O. Rice, (1972) *The average height of planted plane trees*. In: *Graph Theory and Computing*. Edited by R.C. Read, Academic Press, New York, 15-22.
- [10] A. Cayley, (1889) *A theorem on trees*. Quarterly Journal of Pure and Applied Math. 23, 376-378.
- [11] Ph. Chassaing, G. Louchard, (1999) *Phase transition for parking blocks, Brownian excursion and coalescence*. Preliminary version.
- [12] K.L. Chung, (1976) *Excursions in Brownian motion*, Ark. för Math., vol. 14, p.155-177.
- [13] K.L. Chung, (1982) *A cluster of great formulas*. Acta Math. Acad. Sci. Hungar. **39**, no. 1-3, 65-67.
- [14] P.C. Consul, (1989) *Generalized Poisson distributions. Properties and applications*. Statistics: Textbooks and Monographs, 99. Marcel Dekker, Inc., New York.
- [15] M. Csörgö, P. Révész, (1981) *Strong approximations in probability and statistics*. Probability and Mathematical Statistics. Academic Press, Inc., New York-London.
- [16] M. Drmota, B. Gittenberger, (1997) *On the profile of random trees.*, Random Structures Algorithms 10, no. 4, 421-451.

- [17] R.T. Durrett, D.L. Iglehart, D.R. Miller, (1977) *Weak convergence to Brownian meander and Brownian excursion.*, Ann. Probability **5**, no. 1, 117–129.
- [18] P. Flajolet, A. Odlyzko, (Oct 1982) *The average height of binary trees and other simple trees*, J. Comp. and Sys. Sci., Vol. 25, No.2.
- [19] P. Flajolet, P. Poblete, A. Viola, (1998). *On the Analysis of Linear Probing Hashing*. Algoritmica Vol. **22** (4): 490-515.
- [20] D. Foata, J. Riordan, (1974). *Mappings of acyclic and parking functions*. Aequationes math. Vol. **10**: 10-22.
- [21] J. Françon, (1975), *Acyclic and parking functions*. J. Combinat. Theory, Ser. A Vol. **18**: 27-35.
- [22] I.A. Ibragimov, Yu.V. Linnik, (1971) *Independent and stationary sequences of random variables*. Groningen, The Netherlands: Wolters-Noordhoff Publishing Company .
- [23] Th. Jeulin, (1980) *Semi-martingales et grossissement d'une filtration*. Lecture Notes in Mathematics, 833, Springer-Verlag.
- [24] W. D. Kaigh, (1976), *An invariance principle for random walk conditioned by a late return to zero*. Ann. Probability **4**, no. 1, 115–121.
- [25] D.G. Kendall, (1951). *Some problems in the theory of queues*, J. of the Roy. Stat. Soc. **B 13**, 151-185.
- [26] R. Kemp, (1979). *On the average stack size of regularly distributed binary trees*. In *Automata, Lang. and Progr.*, Sixth Colloquium, Graz, Austria 1979. Ed. by H.A. Maurer, Lect. Notes in Comp. Sci. **71**, Springer, 340-355.
- [27] D.P. Kennedy, (1976). *The distribution of the maximum Brownian excursion*, J. Appl. Probab. **13**, 371-376.
- [28] D. E. Knuth, (1997) *The art of computer programming. Vol. 1: fundamental algorithms*, 2nd ed., Addison-Wesley.
- [29] D. E. Knuth, (1997) *The art of computer programming. Vol. 3: sorting and searching*, 2nd ed., Addison-Wesley.
- [30] D. E. Knuth, (1998) *Linear Probing and Graphs*. Algoritmica Vol. **22** (4): 561-568.
- [31] A.G. Konheim, B. Weiss, (1966) *An occupancy discipline and applications*. SIAM J. Appl. Math. Vol. **14** (6): 1266-1274 .
- [32] G. Louchard, (1984). *Kac's formula, Levy's local time and Brownian excursion*, J. Appl. Prob. **21**, p. 479-499.
- [33] J.W. Moon, (1970) *Counting labelled trees*. Canadian Mathematical Monographs, No. 1, Canadian Mathematical Congress, Montreal.

- [34] A. Nijenhuis, H.S. Wilf, (1975) *Combinatorial algorithms*. Academic Press, New York-London.
- [35] A.M. Odlyzko, H.S. Wilf, (1987) *Bandwidths and profiles of trees.*, J. Comb. Theory, Ser. B 42, 348-370.
- [36] R. Otter, (1949) *The multiplicative process*. Ann. Math. Statist., Baltimore Md. 20, 206-224 .
- [37] V.V. Petrov, (1975) *Sums of independant random variables*, Springer .
- [38] J. Pitman, (1998) *Enumerations of trees and forests related to branching processes and random walks*. In *Microsurveys in Discrete Probability*, ed. by D. Aldous and J. Propp. DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI.
- [39] J. Pitman, R. Stanley, *A polytope related to empirical distributions, plane trees, parking functions, and the associahedron*. Technical Report 560, Dept. Statistics, U.C. Berkeley, June 1999. available at <http://stat-www.berkeley.edu/users/pitman/bibliog.html>
- [40] B. Pittel, (1987) *Linear probing: the probable largest search time grows logarithmically with the number of records*. J. Algorithms, Vol. 8 (2): 236–249.
- [41] D. Pollard, (1984) *Convergence of stochastic processes*. Springer Series in Statistics. Springer-Verlag, New York-Berlin.
- [42] A. Renyi & G. Szekeres, (1967) *On the height of trees*. J. Aust. Math. Soc. 7, 497-507.
- [43] M. P. Schutzenberger, (1968) *On an enumeration problem*. J. Combinatorial Theory Vol. 4, 219–221.
- [44] R. Sedgewick, P. Flajolet, (1996) *An introduction to the analysis of algorithms*, Addison-Wesley.
- [45] G. R. Shorack, J. A. Wellner, (1986). *Empirical processes with applications to statistics*. Wiley.
- [46] L. Smith, P. Diaconis, (1988) *Honest Bernoulli Excursions*, J. Appl.Prob. 25, p.464-477.
- [47] J. Spencer, (1997) *Enumerating graphs and Brownian motion*. Commun. Pure Appl. Math. Vol. 50 (3), 291-294.
- [48] K. J. Supowit, E. M. Reingold, (1982/83) *The complexity of drawing trees nicely*. Acta Inform. 18, no. 4, 377–392.
- [49] L. Takács, (1992) *On the total heights of random rooted trees*, J. Appl. Probab. 29, No.3, 543-556.
- [50] L. Takács, (1993) *Limit distributions for queues and random rooted trees*, J. Appl. Math. Stoch. Ana., 6 , No.3, p.189 - 216.

Appendice

On va prouver ici la convergence des moments de la largeur des arbres binaires (conditionnés à avoir n noeuds et divisée par \sqrt{n}) vers les moments correspondant du maximum de l'excursion brownienne normalisée. Plus formellement, on va prouver le

Théorème 4.2

$$E\left(\left(n^{-1/2} W_n^B\right)^r\right) \xrightarrow[n \rightarrow +\infty]{} r(r-1)\Gamma(r/2)\zeta(r)2^{-r/2}$$

(Où la formule est valable pour $r = 1$ en prenant la limite $\sqrt{\pi/2}$).

Ce cas a déjà été traité dans Takács [50]. La preuve que l'on propose ici est légèrement différente et reprend de nombreux arguments utilisés dans le cas des arbres étiquetés.

La première étape consiste à remarquer que la largeur W^B d'un arbre peut être approximée par le maximum de l'excursion associée via le parcours en largeur. Le parcours en largeur fournit une bijection entre l'ensemble F_n des arbres binaires à n noeuds et l'ensemble $Es(n+1)$ des excursions de Bernoulli de longueur $n+1$ (voir Sedgewick & Flajolet [44] et Figure 4.1). On conserve les mêmes notations pour les variables $(x_i)_{i \geq 0}$, ainsi que pour les

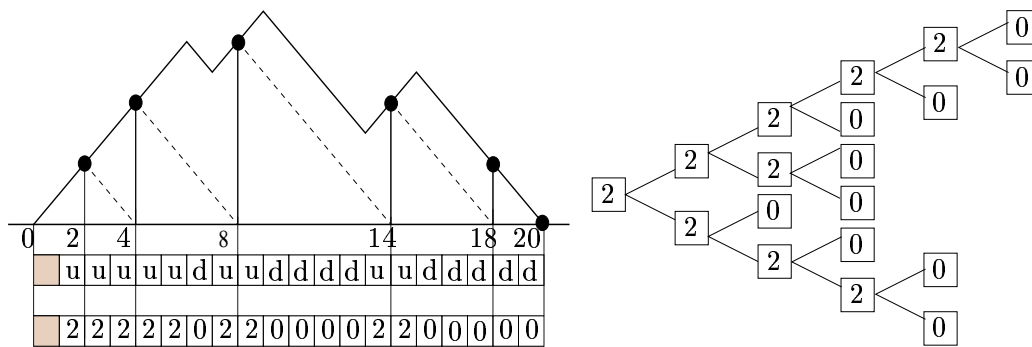


Figure 4.1 : Excursions et arbres binaires

$(Z_i)_{0 \leq i \leq H_n}$. Les nombres x_0, x_1, \dots, x_{n-1} satisfont aussi à l'ensemble de contraintes (2.1) (du Chapitre IV), avec $x_0 = 2$. L'excursion de Bernoulli $(S_k)_{k=0, \dots, n+1}$ associée à ces nombres $(S_k)_{k=0, \dots, n+1}$ (par le parcours en largeur d'abord) est définie comme suit :

$$\begin{aligned} S_0 &= 0, S_1 = 1 \text{ (Construction artificielle)} \\ S_2 &= x_0 \\ S_3 &= x_0 + (x_1 - 1) \\ &\vdots \\ S_n &= x_0 + (x_1 - 1) + \dots + (x_{n-2} - 1) \\ S_{n+1} &= 0 \end{aligned}$$

Comme pour les arbres étiquetés, on a : $Z_0 = S_2, Z_1 = S_{2+Z_0}, \dots, Z_{k+1} = S_{2+Z_1+\dots+Z_k}, \dots$

Dans le cas des arbres étiquetés, les x_i étaient des Poissons $\mathcal{P}(1) - 1$ (conditionnés à former une excursion); dans le cas des arbres binaires, les incréments sont des variables aléatoires uniformes sur l'ensemble $\{-1, 1\}$ (conditionnées à former une excursion). Donc, on peut prouver un analogue du théorème 2.1 pour les arbres binaires :

Proposition 4.3

$$\|W_n^B - \max_{k=0, \dots, n+1} S_k\|_p = O(n^{1/4} \sqrt{\log n}).$$

On peut toutefois simplifier la preuve de l'analogue du théorème 2.1 en utilisant les bornes de Chernov (Bollobás p.12.) à la place du théorème de Petrov ainsi que les nombres de Catalan pour le calcul de $Pr(S \in \Delta(n))$ au lieu du théorème local central limite. La convergence faible du processus

$$\left(\frac{S_{[(n+1)t]}}{\sqrt{n}}\right)_{0 \leq t \leq 1}$$

vers l'excursion Brownienne normalisée $(e(t))_{0 \leq t \leq 1}$ ainsi que la proposition 4.3 impliquent que

$$\frac{W_n}{\sqrt{n}} \xrightarrow{\text{loi}} m$$

où m est le maximum de l'excursion Brownienne normalisée.

Pour obtenir les moments de W_n , dans le cas des arbres étiquetés, on a utilisé un argument de couplage avec le processus empirique. Ici, on va utiliser le fait que le maximum des excursions de Bernoulli et la hauteur des arbres généraux (dont on connaît la loi et les moments limites) sont fortement liés.

On utilise donc une bijection bien connue, qui lie l'ensemble $Es(n+1)$ des excursions et l'ensemble $G_{(n+1)/2}$ des arbres ordonnés avec racine ayant $(n+1)/2$ noeuds (ces arbres "généraux" sont les arbres de Galton-Watson dont la loi de progéniture est géométrique ($\mathbb{P}(Z_1 = k) = p^k(1 - q)$) conditionnés à avoir $(n+1)/2$ noeuds. On pourra se référer aux annexes et au chapitre IV).

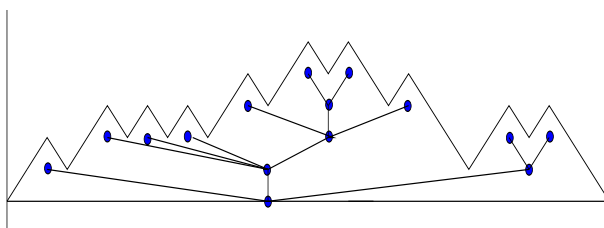


Figure 4.2 :Excursion et arbres en bijection.

Notons $H_G^{((n+1)/2)}$ la hauteur des arbres ordonnés ayant une racine. L'ensemble des arbres de $G_{(n+1)/2}$ et l'ensemble des excursions $Es(n+1)$ étant munis de la loi uniforme, on peut

voir aisément que :

$$H_G^{((n+1)/2)} \stackrel{loi}{=} M_n - 1$$

Les moments limites de H_G sont connus :

Théorème 4.4 (*Flajolet & Odlyzko*)(87)

$$\lim_n \mathbb{E} \left(\left(\frac{W_n}{\sqrt{n}} \right)^r \right) = \lim_n \mathbb{E} \left(\left(\frac{H_G^{(n/2)}}{\sqrt{n}} \right)^r \right) = r(r-1)\Gamma(r/2)\zeta(r)2^{-r/2}$$

En utilisant le théorème 4.1 et la Proposition 4.3, on termine aisément la preuve du théorème 4.2. \square

CHAPITRE V

La montée de barre optimale en présence d'une échéance aléatoire.

P. Chassaing, J.F. Marckert

Soit une famille d'algorithmes $(A(t))_{t>0}$ de durée t non interruptibles et une échéance aléatoire T de loi μ connue. Exécuter jusqu'à terme un algorithme A_t rapporte $f(t)$. On étudie comment construire un algorithme interruptible optimal en moyenne en exécutant successivement une sous famille de $(A_t)_{t>0}$, la règle étant que le résultat d'un algorithme mené à terme remplace le résultat précédent et qu'un algorithme interrompu ne rapporte rien.

1 Introduction

1.1 Cadre général

La résolution de problèmes avec limitation de ressources est un nouveau thème de recherche en Intelligence Artificielle. Son objectif principal est de fournir des méthodes permettant de déterminer la meilleure stratégie pour résoudre une tâche donnée en fonction de contraintes qui lui sont imposées. Le résultat escompté de ces travaux est la conception de systèmes robustes ayant des capacités d'auto-adaptation face à un environnement en constante évolution. Les applications sont nombreuses, notamment dans des domaines critiques nécessitant des traitements répartis ou temps réel tels que l'assistance à la surveillance (de patients en médecine, de divers sous-systèmes embarqués à bord d'avions, de satellites, de voitures, etc.), l'assistance à la conduite (de procédés industriels, de véhicules autonomes, d'avions, de satellites, de réseaux de communication, etc.), l'assistance à la prise de décision (notamment sur les marchés financiers) ou à la recherche d'informations réparties sur plusieurs sites (voir [4]).

Un des problèmes récurrents dans ce genre d'application est le suivant : on doit traiter une tâche de manière la plus précise possible (analyser une image, un son, prendre une décision...) sans savoir exactement le temps dont on dispose pour se faire. On est alors confronté au fait que, le plus souvent, un algorithme interrompu en cours d'exécution ne fournit aucun résultat, même pas un résultat approximatif.

Pour parer ce problème on construit un algorithme interruptible en utilisant une suite d'algorithmes non interruptibles de la façon suivante : on exécute à la suite, des algorithmes $(A(t_i))_{i \geq 1}$ de durée t_i croissante fournissant des approximations de plus en plus précises du résultat recherché. Ainsi, si l'on est interrompu à un temps T au moment où on a déjà exécuté jusqu'à terme $A(t_1), A(t_2), A(t_3)$ mais que l'exécution de $A(t_4)$ n'est pas terminée, on disposera au moins du résultat de $A(t_3)$.

On va supposer par la suite que l'on dispose d'une famille $(A(t))_{t>0}$ d'algorithmes, t désignant le temps nécessaire pour que l'algorithme $A(t)$ produise son résultat $a(t)$. On affecte alors au résultat $a(t)$, une valeur $f(t)$. La fonction f sera supposée croissante. On va donc tenter d'exécuter jusqu'à terme un algorithme $A(s_1)$; si on y parvient, on tentera d'exécuter un algorithme $A(s_2)$, $s_2 > s_1$ etc.... Un algorithme interrompu avant terme ne donne pas de résultat, de valeur $f(0)$ et chaque algorithme est *indépendant* des autres dans le sens où

l'algorithme $A(s_n)$ est de durée s_n quels que soient les calculs effectués auparavant.

On suppose que l'on est interrompu à un temps T aléatoire, la loi μ de T étant connue. Au moment T , on dispose donc du résultat obtenu par le dernier algorithme non interrompu $A(s_{i_T^*})$ où i_T^* est défini par :

$$i_T^* = \max\{i \mid s_1 + s_2 + \dots + s_i \leq T\}. \quad (1.1)$$

Le but que l'on se fixe alors, est de choisir judicieusement la suite $(s_i)_{i \geq 1}$ de manière à maximiser la valeur *moyenne* du résultat obtenu. La question se formule donc ainsi : quelle suite $s_1, s_2, \dots, s_n, \dots$ maximise

$$v = \mathbb{E}(f(s_{i_T^*})) ?$$

On peut écrire v de manière équivalente :

$$v = \max_{s_1 > 0, s_2 > s_1, \dots} \{ \mathbb{P}(T \leq s_1)f(0) + \mathbb{P}(T \geq s_1)f(s_1) + \mathbb{P}(T \geq s_1 + s_2)(f(s_2) - f(s_1)) \\ + \mathbb{P}(T \geq s_1 + s_2 + s_3)(f(s_3) - f(s_2)) + \dots \}$$

A propos de telles problématiques, on pourra voir les articles de Zilberstein (et al.) [5], [6], [7] qui traitent du même type d'optimisation mais lorsque l'on ne connaît pas la loi de l'échéance et que l'on s'intéresse donc à une optimalité dans le "pire des cas". [6] reprend d'ailleurs le même genre d'argument que Baeza et al. [1] qui s'intéressaient à la meilleure stratégie pour trouver un objet sur la droite réelle (entre autres) lorsque sa position n'est pas connue (ni la loi de sa position).

Sans parler pour l'instant de problèmes d'unicité, on peut voir qu'une suite $(s_i)_{i \geq 1}$ maximisant la valeur de v doit satisfaire à deux conditions antagonistes :

- Elle doit être suffisamment "grande", car si on perd du temps à tenter des algorithmes de petites valeurs, on risque de ne plus avoir de temps pour des algorithmes de valeurs supérieures.
- Elle doit comporter des "petites" valeurs pour *garantir* un résultat, $f(0)$ pouvant être particulièrement réhibitoire dans certains cas.

De plus, il est clair que la suite $(s_i)_{i \geq 1}$ dépend fondamentalement de la fonction f . On peut par contre remarquer qu'il existe une borne supérieure relativement petite pour V :

Proposition 1.1 V est majoré par $\mathbb{E}(f(T))$.

Preuve : La première démonstration est intuitive : si l'on connaît à l'avance la date T de l'échéance, on choisira comme première tentative $s_1 = T$, aucune autre tentative n'étant possible ensuite. On obtiendra donc comme récompense

$$\mathbb{E}(f(s_1)) = \mathbb{E}(f(T)).$$

Si on ne connaît pas T à l'avance, le résultat est forcément moins bon. Voici un argument plus formel. Une suite $s = (s_i)_{i \geq 1}$ étant choisie, la récompense (aléatoire) est donnée par :

$$\phi_s(T) = \sum_{n \geq 1} f(s_n) \mathbb{I}_{s_1 + \dots + s_n \leq T < s_1 + \dots + s_{n+1}}$$

Donc, on a évidemment, pour T dans $[s_1 + \dots + s_n, s_1 + \dots + s_{n+1}[$:

$$\phi_s(T) = f(s_n) \leq f(s_1 + \dots + s_n) \leq f(T),$$

d'où

$$v = \mathbb{E}(\phi_s(T)) \leq \mathbb{E}(f(T))$$

1.2 Quelques exemples

Nous allons voir dans cette partie des exemples suffisamment simples qui permettent de trouver la suite optimale s_1, s_2, \dots . Dans tous les cas que l'on va considérer, on va supposer que $\mathbb{E}(T) = 1$.

- Si T peut prendre deux valeurs $0 < t_1 < 1$ et $t_2 > 1$; on va voir que l'on a d'autres alternatives que tenter d'exécuter $A(t_1)$ ou $A(t_2)$.

Une analyse rapide montre que l'on doit choisir entre tenter tout de suite $A(t_2)$ ou tenter $A(s_1)$ avec $s_1 \leq t_1$ puis $A(t_2 - s_1)$ (si $t_2 - s_1 > s_1$). On aura alors :

$$v = \max_{0 \leq s_1 \leq t_1} \{ \mathbb{P}(T = t_1)f(s_1) + \mathbb{P}(T = t_2)f(t_2 - s_1)\mathbb{I}_{t_2 - s_1 > s_1} \}.$$

- Si T suit une loi uniforme dans l'intervalle $[1 - a, 1 + a]$, de la même façon, on remarque que

$$v = \max_{0 \leq s_1 \leq 1 - a} \left\{ \max_{1 - a - s_1 \leq x \leq 1 + a - s_1} \{ \mathbb{P}(T \leq s_1 + x)f(s_1) + \mathbb{P}(T > s_1 + x)f(x) \} \right\}.$$

On peut remarquer que pour certaines fonctions de paiement f , il n'y a pas unicité du x qui maximise v . Par exemple si T suit une loi uniforme dans l'intervalle $[0, 2]$ avec $f(t) = t$; on s'aperçoit que la quantité v optimale vaut $1/2$ et peut être obtenue de plusieurs manières :

- en prenant $s_1 = 1$ (et donc, aucune autre tentative n'est possible ensuite)
- en prenant $0 \leq s_1 < 1$ et $s_2 = 1$.

Dans ce cas particulier, il n'existe pas de manière d'obtenir la même récompense moyenne avec trois tentatives (ou plus).

- Si la suite des algorithmes exécutables est réduite à $\{A_{t_1}, \dots, A_{t_n}\}$ pour des temps t_i donnés, le problème se résout complètement, puisque la suite maximisant la moyenne est une partie de l'ensemble $\{A_{t_1}, \dots, A_{t_n}\}$; il n'y a qu'à les essayer toutes...

D'une manière générale on voit que le problème posé est complexe à résoudre et ceci même numériquement. En effet, même si T est à valeur dans un compact $[A, B]$ le nombre de tentatives à choisir n'est pas du tout clair. Si le nombre k de tentatives est connu et s'il est grand, on voit que chercher les valeurs des t_i revient à chercher le maximum d'une fonction de k variables sous certaines contraintes. On va dans la suite s'intéresser à un cas particulier, révélateur des difficultés que présente ce genre de problème lorsque le support de T est \mathbb{R}^+ .

2 Le cas T de loi exponentielle

On va prendre comme fonction valeur f , la fonction

$$f(t) = t$$

et pour loi de l'échéance T , la loi exponentielle de moyenne 1, c'est à dire, que la queue de distribution de T est donnée par :

$$\mathbb{P}(T \geq x) = e^{-x}.$$

La loi exponentielle est souvent utilisée pour formaliser des instants de panne ou des lois d'échéance pour la raison qu'elle est "sans mémoire", c'est à dire que la durée résiduelle sachant que l'échéance n'est pas survenue avant le temps t a même loi que la durée initiale :

$$\mathbb{P}(T \geq t + s \mid T \geq t) = \mathbb{P}(T \geq s).$$

Si on a réussi à exécuter jusqu'à terme un algorithme $A(s_1)$, on peut recommencer un autre algorithme $A(s_2)$ et cet algorithme ira ou non à terme indépendamment de la valeur de s_1 . Ceci "simplifie" en quelque sorte les choses puisque les probabilités conditionnelles deviennent de vraies probabilités mais d'un autre coté on n'a acquis aucune information du passé.

La suite $(s_k)_{k \geq 1}$ étant positive et croissante, on note $x = (x_i)_{i \geq 1}$ la suite positive constituée de ses accroissements de sorte que

$$s_k = x_1 + \dots + x_k.$$

Le profit moyen associée à la suite $x = (x_i)_{i \geq 1}$ est

$$v(x) = x_1 e^{-w_1} + x_2 e^{-w_2} + \dots + x_k e^{-w_k} + \dots \quad (2.2)$$

où

$$w_k = s_1 + s_2 + \dots + s_k = kx_1 + (k-1)x_2 + \dots + x_k$$

car

$$v(x) = x_1 \mathbb{P}(T \geq s_1) + x_2 \mathbb{P}(T \geq s_1 + s_2) + \dots + x_n \mathbb{P}(T \geq s_1 + s_2 + \dots + s_n) + \dots .$$

v vérifie alors

$$v = \sup\{v(x) \mid x \in E\}$$

où E est l'ensemble des suites à termes positifs. Le problème consiste donc à calculer v ainsi qu'une suite x^* réalisant ce maximum, c'est à dire telle que $v = v(x^*)$.

2.1 Une équation de Bellman

Supposons que $s_n = a$ soit fixé ; on a :

$$\begin{aligned} v_n(x) &= v(x) - (x_1 e^{-w_1} + x_2 e^{-w_2} + \dots + x_n e^{-w_n}) \\ &= e^{-w_n} (\tilde{x}_1 e^{-\tilde{w}_1} e^{-a} + \tilde{x}_2 e^{-\tilde{w}_2} e^{-2a} \dots + \tilde{x}_k e^{-\tilde{w}_k} e^{-ka} + \dots) \\ &= e^{-w_n - a} w(a, \tilde{x}) \end{aligned}$$

où les différents tildes désignent les quantités associées à la suite décalée de x :

$$\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_k, \dots) = (x_{n+1}, x_{n+2}, \dots, x_{n+k}, \dots)$$

et où $w(a, x)$ est défini par :

$$w(a, x) = x_1 e^{-w_1} + x_2 e^{-w_2} e^{-a} + \dots + x_k e^{-w_k} e^{-(k-1)a} + \dots$$

On peut noter que

$$w(0, x) = v(x)$$

$e^{-a} w(a, \tilde{x})$ s'interprète comme ce qu'on obtient de plus que a en moyenne si on a déjà a et si la séquence des tentatives effectuées est :

$$(a + \tilde{s}_1, a + \tilde{s}_2, \dots, a + \tilde{s}_k, \dots)$$

Posons

$$V(a) = a + e^{-a} \max\{w(a, x) | x \in E\}.$$

$V(a)$ s'interprète comme le mieux que l'on puisse obtenir de plus que a en moyenne si on a déjà a . Le but est donc de calculer $V(0)$. L'équation de *programmation dynamique* s'écrit :

$$\begin{aligned} V(a) &= \max\{a(1 - e^{-y}) + e^{-y} V(y) | y > a\} \\ &= a + e^{-a} \max\{(V(a+x) - a)e^{-x} | x > 0\} \end{aligned} \quad (2.3)$$

On pourra voir au sujet des équations de programmation dynamique les nombreux articles de Bellman et ses livres, en particulier [2]. Notons

$$\Psi(x) = V(x) - x.$$

On a donc

$$\Psi(a) = e^{-a} \max\{w(a, x) | x \in E\}$$

et on peut dès maintenant établir que :

Proposition 2.1

- a) $x \rightarrow e^x \Psi(x)$ est décroissante
- b) Ψ est convexe

c) Ψ est majoré par 1

Preuve : On a

$$e^a \Psi(a) = \max\{w(a, x) | x \in E\}.$$

Pour $x \in E$, $a \rightarrow w(a, x)$ est une fonction convexe, décroissante en a . Donc

$$a \rightarrow \max\{w(a, x) | x \in E\}$$

est également convexe et décroissante. Puisque e^{-x} et $w(a, x)$ sont deux fonctions convexes, positives et décroissantes, leur produit est convexe. Pour prouver c) on utilise le fait que Ψ est décroissante et la Proposition 1.1 avec $\mathbb{E}(f(T)) = 1$. \square

Mais aussi, de l'égalité (2.3), on déduit que $\Psi : \mathbb{R}^+ \rightarrow \mathbb{R}$ est solution d'une équation dite de Bellman (2.4) très importante pour la suite :

$$\bar{\Psi}(a) = e^{-a} \max\{(x + \bar{\Psi}(a + x))e^{-x} | x > 0\} \quad (2.4)$$

Une étude complète de Ψ nous fournira aussi bien $\Psi(0) = V(0) = v$ que la suite optimisante x^* . On va donc procéder de la manière suivante : dans la partie 2.2, on va montrer que Ψ est l'unique solution convexe décroissante à l'équation de Bellman (2.4) ; dans les parties 2.3, 2.4, 2.5 on va montrer que toute solution d'un système différentiel, et d'une équation fonctionnelle non standards est également solution de l'équation de Bellman. Dans les parties 2.6, 2.7, 2.8 on va "prouver" qu'il existe une solution décroissante et convexe à l'équation différentielle non standard et étudier ses propriétés ce qui nous conduira à calculer la suite $(s_i)_{i \geq 1}$ qui réalise v . Malheureusement, une étape dans une démonstration manque toujours...

2.2 "Unicité" de la solution de l'équation de Bellman

Le but de cette partie est de prouver le

Théorème 2.2 *Il existe une seule fonction décroissante convexe (sur \mathbb{R}^+) solution de l'équation de Bellman (2.4).*

Soit H_b l'espace des applications ϕ convexes décroissantes de $[b, +\infty[$ dans \mathbb{R}^+ et T_b l'opérateur défini sur H_b par :

$$(T_b \phi)(y) = e^{-y} \max\{(x + \phi(x + y))e^{-x} | x \geq 0\}.$$

Lemme 1 *T_b est contractant sur H_b ; de plus, si ϕ et ψ sont deux fonctions de H_b , on a :*

$$\|T_b \phi - T_b \psi\|_\infty \leq e^{-b} \|\phi - \psi\|_\infty$$

Preuve : . Montrons d'abord que H_b est stable par T_b . Tout d'abord, il est facile de voir que $T_b \phi$ est aussi décroissant. Montrons que $T_b \phi$ est convexe. Posons

$$\Theta(y) = \max\{(x + \phi(x + y))e^{-x} | x \geq 0\}.$$

Alors Θ est convexe ; en effet :

$$\begin{aligned}\Theta(a) + \Theta(b) &\geq 2 \max \left\{ \left(x + \frac{\phi(x+a) + \phi(x+b)}{2} \right) e^{-x} \mid x \geq 0 \right\} \\ &\geq 2 \max \left\{ \left(x + \phi \left(x + \frac{a+b}{2} \right) \right) e^{-x} \mid x \geq 0 \right\} \\ &\geq 2\Theta \left(\frac{a+b}{2} \right)\end{aligned}$$

. Soit ϕ et ψ deux fonctions de H_b ; on note g et h les fonctions définies par :

$$\begin{aligned}g(x) &= (x + \phi(x+a))e^{-x} \\ h(x) &= (x + \psi(x+a))e^{-x}\end{aligned}$$

On a

$$\begin{aligned}|T_b\phi - T_b\psi|(a) &= e^{-a} | \max\{g(x) \mid x \geq 0\} - \max\{h(x) \mid x \geq 0\} | \\ &\leq e^{-a} \max\{|g(x) - h(x)| \mid x \geq 0\} \\ &\leq e^{-a} \max\{|\phi(x+a) - \psi(x+a)| e^{-x} \mid x \geq 0\} \\ &\leq e^{-a} \|\phi - \psi\|_\infty. \quad \square\end{aligned}$$

Corollaire 2.3 *Il existe une unique solution Φ positive convexe et décroissante à l'équation de Bellman sur $]0, +\infty[$.*

On ne peut pas prolonger de façon unique cette solution Φ sur $[0, +\infty[$ car en choisissant $\Phi(0)$ quelconque tel que :

$$\Phi(0) \geq \max \{ (x + \Phi(x))e^{-x} \mid x > 0 \}$$

alors Φ satisfera à

$$\Phi(0) = \max \{ (x + \Phi(x))e^{-x} \mid x \geq 0 \}$$

Proposition 2.4 $\bar{\Psi}$ est l'extension de Φ à $[0, +\infty[$ qui satisfait à

$$\bar{\Psi}(0) = \max \{ (x + \Phi(x))e^{-x} \mid x > 0 \}.$$

C'est donc la plus petite extension de Φ à $[0, +\infty[$.

Preuve :

$$\begin{aligned}\bar{\Psi}(0) &= \sup \{ x_1 e^{-w_1} + x_2 e^{-w_2} + x_3 e^{-w_3} + \dots \mid x = (x_i)_{i \geq 1} \in E \} \\ &= \sup \{ \sup \{ x_1 e^{-w_1} + x_2 e^{-w_2 - a} + x_3 e^{-w_3 - 2a} + \dots \mid a > 0 \} \mid x = (x_i)_{i \geq 1} \in E \} \\ &= \sup \{ \sup \{ x_1 e^{-w_1} + x_2 e^{-w_2 - a} + x_3 e^{-w_3 - 2a} + \dots \mid x = (x_i)_{i \geq 1} \in E \} \mid a > 0 \} \\ &= \sup \{ \bar{\Psi}(a) \mid a > 0 \} \\ &= \sup \{ \Phi(a) \mid a > 0 \} \\ &= \sup \{ \sup \{ (x + \Phi(a+x))e^{-x} \mid x > 0 \} \mid a > 0 \} \\ &= \sup \{ \sup \{ (x + \Phi(a+x))e^{-x} \mid a > 0 \} \mid x > 0 \} \\ &= \sup \{ (x + \Phi(x))e^{-x} \mid x > 0 \}. \quad \square\end{aligned}$$

De la Proposition 2.1, on déduit immédiatement que

Corollaire 2.5 Ψ est l'unique solution décroissante convexe, continue en 0 de l'équation de Bellman.

Ψ vérifie de plus :

Proposition 2.6

$$e^{-x-1} \leq \Psi(x) \leq \frac{e^{-x}}{2}, \quad \text{pour tout } x \in \mathbb{R}^+.$$

Preuve : On considère H'_b l'espace des fonctions ϕ convexes décroissantes de $[b, +\infty[$ dans \mathbb{R}^+ telles que

$$\frac{e^{-x}}{e} \leq \phi(x) \leq \frac{e^{-x}}{2} \quad \text{pour tout } x \in [b, +\infty[.$$

On va montrer que H'_b est stable pour T_b , la contraction se montrant comme sur H_b . Soit ϕ appartenant à H'_b et $y \geq b$, on a :

$$(T_b\phi)(y) \leq e^{-y} \max \left\{ \left(x + \frac{e^{-x}}{2}\right)e^{-x} \mid x > 0 \right\} \leq \frac{e^{-y}}{2}$$

et

$$(T_b\phi)(y) \geq e^{-y} \max \left\{ (x + e^{-x-y-1})e^{-x} \mid x > 0 \right\} \geq e^{-y-1}. \quad \square$$

On peut par exemple déduire de cette proposition que la valeur recherchée v vérifie :

$$v = \Psi(0) \in [e^{-1}, 1/2].$$

2.3 Lien avec un système différentiel "exotique"

Notons

$$\tau(a) = \inf \left\{ \operatorname{argmax} \left\{ (x + \Psi(a+x))e^{-x} \mid x > 0 \right\} \right\}.$$

où $\operatorname{argmax}\{f(x) \mid x > 0\}$ est l'ensemble des valeurs positives de x où f atteint son maximum. On peut remarquer que Ψ étant convexe décroissante, Ψ est également continue et ainsi la fonction $x \rightarrow (x + \Psi(a+x))e^{-x}$ est continue et bornée d'après la Proposition 2.1. Elle possède donc un maximum et (au moins) un endroit où il est atteint : τ est donc bien défini. De plus, il est facile de prouver que τ prend ses valeurs dans $[0, 1]$, car $\Psi(a+x)e^{-x}$ est décroissante sur \mathbb{R}^+ ($\forall a > 0$) et xe^{-x} est décroissante sur $[1, +\infty[$.

On va supposer pour l'instant que Ψ est dérivable, ce qui est le cas sauf au plus sur un ensemble dénombrable de points puisque Ψ est convexe. En écrivant que $\tau(a)$ annule la dérivée de $(x + \Psi(a+x))e^{-x}$, on obtient :

$$\tau(a) = 1 + \Psi'(a + \tau(a)) - \Psi(a + \tau(a)). \quad (2.5)$$

Puisque

$$\Psi(a) = e^{-(a+\tau(a))} (\tau(a) + \Psi(a + \tau(a))) \quad (2.6)$$

on obtient :

$$\Psi'(a) = -e^{-(a+\tau(a))} \quad (2.7)$$

On en déduit que

$$\tau(a) = 1 - e^{-(a+\tau(a)+\tau(a+\tau(a)))} - \Psi(a + \tau(a))$$

d'où

$$\Psi(0) = e^{-\tau(0)}(\tau(0) + \Psi(\tau(0)))$$

et

$$\tau(0) = 1 - e^{-(\tau(0)+\tau(\tau(0)))} - \Psi(\tau(0)).$$

On en déduit que la performance optimale (en moyenne) s'exprime en fonction des valeurs des deux premières tentatives $\tau(0)$ et $\tau(\tau(0))$:

$$\Psi(0) = e^{-\tau(0)}(1 - e^{-\tau(0)-\tau(\tau(0))}).$$

Posons $z(a) = a + \tau(a)$ dans les équations (2.5), (2.6), (2.7). Le système composé de ces trois équations s'écrit :

$$\begin{cases} \psi(a) &= e^{-z(a)}(z(a) - a + \psi(z(a))) \\ z(a) &= 1 + a + \psi'(z(a)) - \psi(z(a)) \\ \psi'(a) &= -e^{-z(a)} \end{cases} \quad (2.8)$$

Voici quelques identités vérifiées par les solutions de (2.8) :

$$\begin{aligned} z(a) - a + \psi(z(a)) &= 1 + \psi'(z(a)) \\ &= \psi(a)e^{z(a)} \\ &= 1 - e^{-z(z(a))} \end{aligned}$$

D'où on déduit

$$\psi(x) = e^{-z(x)}(1 - e^{-z(z(x))}) \quad \text{et} \quad z(x) = x + 1 - e^{-z(z(x))} \left(2 - e^{-z(z(z(x)))} \right). \quad (2.9)$$

Ceci étant obtenu en remplaçant $\psi(x)$ par $e^{-z(x)}(1 - e^{-z(z(x))})$ dans la première équation de (2.8). Cette deuxième identité est particulièrement intéressante puisqu'elle permet de calculer la n^{eme} décision en connaissant les précédentes.

Pour justifier la dérivabilité de la solution Ψ de l'équation de Bellman, on va prouver les deux propositions suivantes :

Proposition 2.7 *Toute solution (ψ, z) au système (2.8) (avec ψ convexe décroissante et z croissante concave) est solution de l'équation de Bellman (2.4).*

Proposition 2.8 *Il existe une solution (ψ, z) au système (2.8) avec ψ convexe dérivable, décroissante et z croissante concave.*

Supposons ces deux points prouvés. On sait alors que la solution Ψ décroissante et convexe de l'équation de Bellman est dérivable. Ainsi, tous les calculs effectués en supposant la dérivabilité de Ψ sont justifiés.

2.4 Preuve de la Proposition 2.7

On veut montrer qu'une solution (ψ, z) du système (2.8) est une solution de l'équation de Bellman, c'est à dire quelle vérifie :

$$\begin{cases} \psi(a) &= \max \{ (x - a + \psi(x)) e^{-x} | x > a \} \\ z(a) &= \operatorname{argmax} \{ (x - a + \psi(x)) e^{-x} | x > a \} \end{cases}$$

On peut remarquer que l'infimum a disparu de la définition de z . Ceci est quand même cohérent avec la définition de τ puisque l'on va démontrer l'unicité du x qui maximise $(x - a + \psi(x)) e^{-x}$. Notons

$$G_a(y) = e^{-y}(y - a + \psi(y))$$

Proposition 2.9

$$z(a) = \operatorname{argmax} \{ G_a(y) | y > a \}$$

et

$$\psi(a) = G_a(z(a)) = \max \{ G_a(y) | y > a \}$$

On peut remarquer que cette proposition implique directement que (ψ, z) est solution de l'équation de Bellman. Il suffit donc de prouver la Proposition 2.9 et même seulement sa première assertion la deuxième étant équivalente. C'est une preuve en trois points.

(Puisque z n'est pas supposée dérivable mais concave, z' désignera dans la suite, la dérivée à droite de z .)

1) G_a possède un maximum local en $z(a)$.

On a

$$G'_a(y) = e^{-y}(1 + \psi'(y) - y + a - \psi(y)) \quad (2.10)$$

et donc

$$G'_a(z(a)) = 0.$$

$$G''_a(y) = e^{-y}(-1 + e^{-z(y)} + y - a + \psi(y) + z'(y)e^{-z(y)} - 1 - \psi'(y)).$$

Après quelques simplifications, on s'aperçoit que :

$$G''_a(z(a)) = -\frac{e^{-z(a)}}{z'(a)} < 0$$

2) Le seul maximum local de G_a dans $]a, +\infty[$ est $z(a)$. En effet,

$$e^y G'_a(y) = 1 + \psi'(y) - y + a - \psi(y)$$

est une fonction strictement concave; elle ne peut s'annuler que deux fois. L'autre point singulier, s'il existe est nécessairement un minimum local.

3) a n'est pas un maximum de G_a (sauf si $a=0$)

$$G_a(z(a)) - G_a(a) = \psi(a) - e^{-a}\psi(a)$$

D'autre part, le maximum de G_a n'est pas atteint en $+\infty$ puisque G_a est décroissante sur $[z(a), +\infty[$ d'après 2).

2.5 Preuve de la Proposition 2.8

C'est cette partie qui nous a posé problème jusque maintenant et elle n'est donc pas totalement achevée. Comme on va le voir, on ne parvient à démontrer l'existence d'une solution du système (2.8) avec les propriétés désirées que sur $[0.4, +\infty[$ alors que l'on a besoin de prouver son existence sur \mathbb{R}_+ .

Le but de cette partie est de prouver l'existence d'une solution au système (2.8). On notera par la suite $f^{[2]}$ et $f^{[3]}$ les fonctions $f \circ f$ et $f \circ f \circ f$. Soit (ψ, z) une solution de (2.8). Il est facile de voir que z' vérifie

$$z'(a) = 1 + z'(a)z'(z(a))e^{-z^{[2]}(a)} + z'(a)e^{-z^{[2]}(a)} \quad (2.11)$$

On a par ailleurs la condition "initiale" en $+\infty$:

$$z(a) = a + 1 - o(1). \quad (2.12)$$

Un système équivalent

Lemma 2.10 *Le système*

$$\begin{cases} z'(a) &= 1 + z'(a)z'(z(a))e^{-z^{[2]}(a)} + z'(a)e^{-z^{[2]}(a)} \\ \psi'(a) &= -e^{-z(a)} \end{cases} \quad (2.13)$$

est équivalent au système (2.8). Les deux systèmes ayant même condition "initiale".

Preuve : Il est clair qu'une solution à (2.8) sera également une solution à (2.13).

Montrons qu'une solution (ψ, z) de (2.13) vérifie

$$z(a) = 1 + a + \psi'(z(a)) - \psi(z(a))$$

On note

$$\begin{aligned} g(a) &= z(a) - 1 - a - \psi'(z(a)) + \psi(z(a)) \\ &= z(a) - 1 - a + e^{-z^{[2]}(a)} + \psi(z(a)) \end{aligned}$$

$$g'(a) = z'(a) - 1 - z'(z(a))z'(a)e^{-z^{[2]}(a)} + z'(a)\psi'(z(a)) = 0$$

donc g est constante et la condition initiale en $+\infty$ indique que $g \equiv 0$.

Montrons maintenant qu'on a aussi :

$$\psi(a) = e^{-z(a)}(z(a) - a + \psi(z(a))).$$

On note

$$h(a) = \psi(a) - e^{-z(a)}(z(a) - a - \psi(z(a)))$$

$$\begin{aligned} h'(a) &= \psi'(a) - e^{-z(a)}(-z'(a)(z(a) - a + \psi(z(a))) + z'(a) - 1 - z'(a)\psi'(z(a))) \\ &= \psi'(a) - e^{-z(a)}(-a) = 0 \end{aligned}$$

et puisque $\lim_{a \rightarrow +\infty} h(a) = 0$, $h \equiv 0$. \square

Il suffit donc de prouver que le système (2.13) a une solution.

2.6 Une “sous” solution

On va commencer par prouver directement qu’une équation fonctionnelle un peu plus simple que (2.9) possède des solutions. Elle nous servira de borne inférieure dans la sous-section suivante. Soit l’équation fonctionnelle

$$(E) \quad w(x) = 1 + x - 2e^{-w^{[2]}(x)}. \quad (2.14)$$

Alors,

Théorème 2.11 *Il existe une solution à l’équation (E) sur $[0.4, +\infty[$. De plus, cette solution est croissante et est asymptotique à la droite $y = x + 1$ en $+\infty$.*

Définissons la suite de fonctions $(w_k)_{k \geq 1}$ par

$$\begin{cases} w_0(x) &= x + 1 \\ w_{k+1}(x) &= 1 + x - 2 \exp(-w_k^{[2]}(x)) \end{cases} \quad (2.8)$$

alors

Lemma 2.12 *La suite $(w_k)_k$ est une suite décroissante de fonctions dérivables, croissantes et asymptotiques à la droite d’équation $y = x + 1$ en $+\infty$.*

Preuve : La dérivabilité ainsi que les propriétés asymptotiques de w_{k+1} s’héritent de façon simple de celles de w_k . La suite w_k est décroissante : On a évidemment $w_2(x) < w_1(x)$. On montre ensuite aisément la décroissance de la suite $(w_k(x))$ par récurrence ainsi que le fait qu’elles soient croissantes de pentes supérieures à 1. \square

Lemma 2.13 *Pour tout $x \geq 0.4$ et pour tout $k \geq 0$, on a*

$$w_k(x) \geq x + 0.45$$

Preuve : Ceci est vrai au rang 0 ; supposons que $w_k(x) \geq x + 0.45$ pour tout $x \geq 0.4$.

$$\begin{aligned} w_{k+1}(x) &\geq x + 1 - 2 \exp(-x - 0.9) \text{ pour } x \geq 0.4 \\ &\geq x + 1 - 2 \exp(-1.3) \\ &\geq x + 0.45. \quad \square \end{aligned}$$

En réinjectant cette borne dans l’équation définissant w_{k+1} , on remarque que la distance entre la droite d’équation $y = x + 1$ et w_k est un $O(e^{-x})$, uniformément en k .

2.7 Existence d’une solution z

On va prouver qu’il existe une solution à l’équation (2.11),

$$z'(x) = 1 + z'(x)z'(z(x))e^{-z^{[2]}(x)} + z'(x)e^{-z^{[2]}(x)}$$

avec z croissante, asymptotique à $x + 1$ en $+\infty$. On note $(z_k)_{k \geq 0}$ la suite de fonctions définies par :

$$\begin{cases} z_0(x) &= x + 1 \\ z'_{k+1}(x) &= 1 + z'_k(x)z'_k(z_k(x))e^{-z_k^{[2]}(x)} + z'_k(x)e^{-z_k^{[2]}(x)} \end{cases}$$

Pour définir correctement z_{k+1} , on prend pour $z_{k+1}(0.4)$ (puisque l'existence ne sera prouvée uniquement que jusqu'en 0.4) la valeur telle que $z_{k+1}(x)$ soit asymptotique à $x+1$ en $+\infty$. On va prouver qu'il est possible de choisir une telle valeur. Pour ceci, on compare les fonctions z_k avec les fonctions w_k définies dans la sous-section ci-dessus; une définition équivalente pour les fonctions w_k est la suivante :

$$\begin{cases} w_0(x) &= x + 1 \\ w'_{k+1}(x) &= 1 + 2w'_k(x)w'_k(w_k(x))e^{-w_k^{[2]}(x)} \end{cases}$$

$w_{k+1}(0.4)$ étant choisi de sorte que $w_{k+1}(x)$ soit asymptotique à $x+1$ en $+\infty$.

Lemma 2.14 *La suite $(z'_k(x))_k$ est une suite croissante et donc $(z_k(x))_k$ est une suite décroissante.*

Preuve : Tout d'abord, $z_0(x)$ est une fonction positive, dont les dérivées vérifient :

$$\text{signe}(z_0^{(n)}(x)) = (-1)^{n+1}$$

Cette propriété est conservée pour les fonctions z_k (voir Feller [3] p. 439 et suivantes). On a $z'_1(x) = 1 + 2e^{-x-2}$ d'où $z'_1 > z'_0$ et $z_1 < z_0$. Supposons que $z'_k > z'_{k-1}$ et que $z_k < z_{k-1}$ alors, en utilisant la décroissance de z'_k , il est clair que $z'_{k+1} > z'_k$.

Lemma 2.15 *Pour tout $k \geq 0$ et pour tout $x \geq 0.4$, on a :*

$$\begin{aligned} w'_k(x) &\geq z'_k(x) \\ w_k(x) &\leq z_k(x) \end{aligned}$$

Preuve : Immédiat par récurrence en utilisant le fait que $z'_k(z_k(x)) \geq 1$. \square

De tout ceci, on déduit que

Théorème 2.16 *Sur $[0.4, +\infty[$:*

$(z_k)_{k \geq 0}$ est une suite décroissante de fonctions croissantes et concaves. Elle est minorée; elle converge donc vers une fonction z croissante et concave. De plus, la suite des dérivées $(z'_k)_k$ est croissante; elle converge donc sur $[0.4, +\infty[$ vers z' solution de l'équation différentielle.

2.8 Approximation de z et de Ψ

On va rechercher une solution numérique approchée à l'équation fonctionnelle

$$z(x) - x - 1 + e^{-z(z(x))} \left(2 - e^{-z(z(z(x)))} \right) = 0$$

de la façon suivante : on prend

$$y_0(x) := x + 1$$

puis par itération pour $n \geq 1$, on calcule

$$y_n(x) := x + 1 - e^{-y_{n-1}(y_{n-1}(x))} \left(2 - e^{-y_{n-1}(y_{n-1}(y_{n-1}(x)))} \right).$$

Cette construction de y_n n'est pas équivalente à celle de z_n dont on a montré la convergence vers la solution z sur $[0.4, +\infty[$ plus haut. Ainsi, les résultats qui suivent ne sont donc pas rigoureusement démontrés.

La suite y_n ainsi définie peut être implémentée de façon assez simple et a le bon goût de converger relativement vite. Les résultats que l'on obtient sont tout à fait cohérents avec ce que l'on a démontré jusque maintenant sur z et également avec les valeurs qui maximisent

$$v_k(x) = x_1 e^{-w_1} + x_2 e^{-w_2} + \dots + x_k e^{-w_k}$$

pour les premières valeurs de k . On trouve donc les premières tentatives à effectuer, une valeur approchée de v , ainsi qu'une représentation graphique approchée de z et de Ψ .

$$\begin{aligned} z(0) &= 0.426575\dots \\ z(z(0)) &= 1.185596\dots \\ z(z(z(0))) &= 2.092408\dots \\ z(z(z(z(0)))) &= 3.057487\dots \end{aligned}$$

Ainsi, on tentera successivement d'exécuter les algorithmes de durée

$$\begin{aligned} &0.426575 \\ &1.185596 \\ &2.092408 \\ &3.057487 \\ &\vdots \end{aligned}$$

Grâce à l'équation fonctionnelle (2.9) on trouve également une valeur approchée de v :

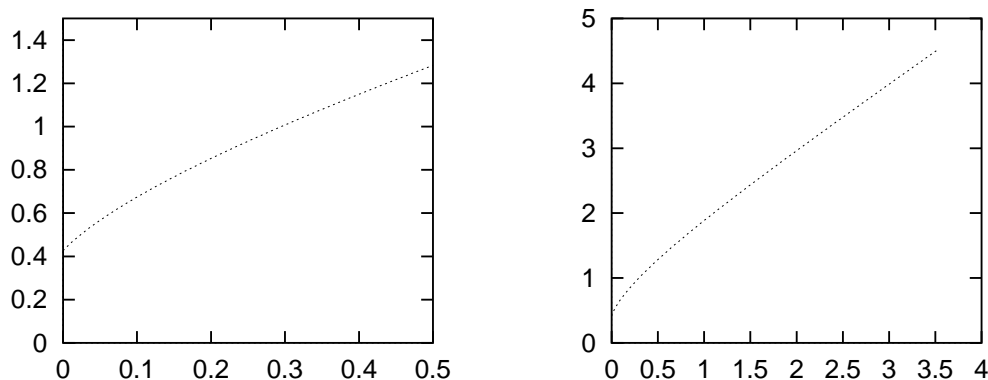


Figure 2.1 Fonction z

$$v = \Psi(0) = 0.453286525\dots$$

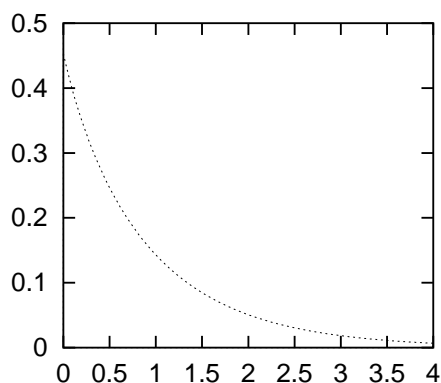


Figure 2.2 Fonction Ψ

Références

- [1] R. Baeza-Yates, J.C. Culberson, G.J.E. Rawlins, (1993) *Searching in the plane*, Inf. Comput. 106, No.2, 234-252.
- [2] R. Bellman, (1947) *Dynamic programming*, Princeton, New Jersey : Princeton University Press. XXV.
- [3] W. Feller, (1966) *An introduction to probability theory and its applications*, Vol. II, John Wiley and Sons, Inc. XVIII.
- [4] E. Horvitz, (1988) *Reasoning under varying and uncertain constraints*, Seventh National Conference on Artificial Intelligence, 11-116, 1988.
- [5] S. Zilberstein, (1995) *Optimizing decision quality with contract algorithms*, 14th international joint conference on artificial intelligence 49, p.1576-1582.
- [6] S. Zilberstein, F. Charpillet, P. Chassaing, (1999) *Real-Time Problem-Solving with Contract Algorithm*, 16th international joint conference on artificial intelligence, Stockholm, Aug. 1999.
- [7] S. Zilberstein, S.J. Russel, (1996) *Optimal composition of real-time systems*, Artificial Intelligence 82 (1-2), 181-213.

ANNEXES

Le but de ces annexes, n'est pas de donner une liste exhaustive des définitions, propriétés ou théorèmes utilisés dans cette thèse mais plutôt d'illustrer, d'expliquer un peu plus largement et simplement les principaux objets mis en oeuvre, ce que la rédaction sous forme d'article n'a pas toujours permis ; pour cette raison, on débordera parfois légèrement du cadre de cette thèse.

On s'attachera également à donner des références les plus accessibles possibles.

1 Processus browniens

1.1 Marches aléatoires et excursions discrètes

Une marche aléatoire est une suite $\omega = (S_k(\omega))_{k \in \mathbb{N}}$ définie par :

$S_0(\omega) = 0$ et $S_k(\omega) = X_1(\omega) + \dots + X_k(\omega)$ où les X_i désignent des variables aléatoires indépendantes et de même loi. On représente généralement les marches aléatoires par la ligne brisée liant les points (k, S_k) . Les *marches simples symétriques* sont les marches où les variables X_i sont uniformes sur $\{-1, 1\}$.

La loi sur ces marches (de longueur n) induite par la loi des X_i est alors l'équiprobabilité. Ainsi, calculer des probabilités sur ces marches, revient souvent à effectuer des dénombrements. On pourra se référer aux livres de Révész [40] et Spitzer [46].

Les excursions discrètes de longueur n sont des marches aléatoires conditionnées à respecter l'ensemble de contraintes suivant :

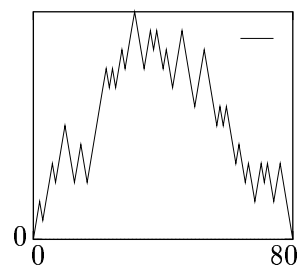
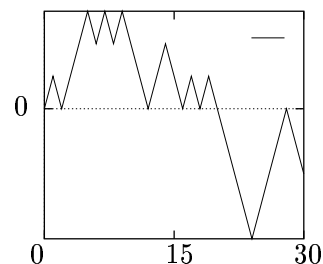
$$S_0 = 0, S_1 > 0, \dots, S_{n-1} > 0, S_n = 0. \quad (\star)$$

On appellera excursion de Bernoulli, les excursions discrètes dont les incréments valent $+1$ ou -1 . Elles sont obtenues en conditionnant l'ensemble des marches simples symétriques à respecter l'ensemble de contraintes (\star) . Pour cette raison, la loi induite sur l'ensemble des *excursions de Bernoulli* est également l'équiprobabilité. Le nombre d'excursions de Bernoulli de longueur $2n$ est donné par $C_{n-1} = \binom{2n-2}{n-1}/n$, le $n-1$ ème nombre de Catalan.

De manière assez générale, si ω est une marche aléatoire sur les entiers, dont les seuls incréments négatifs sont des -1 , et si on note $T_{-k} = \min\{j | S_j = -k\}$ le temps d'atteinte de $-k$ ($-k$ négatif) par la marche, on a d'après Kemperman [28], Otter [37] ou Dwass [19] :

$$\mathbb{P}(T_{-k} = n) = \frac{k}{n} \mathbb{P}(S_n = -k)$$

Ceci nous permet de calculer la probabilité qu'une marche aléatoire de longueur n soit une



excursion ; en effet,

$$\begin{aligned}
 \mathbb{P}(S_1 > 0, \dots, S_{n-1} > 0, S_n = 0) &= \sum_{k>0} \mathbb{P}(S_1 = k, S_2 > 0, \dots, S_{n-1} > 0, S_n = 0) \\
 &= \sum_{k>0} \mathbb{P}(X_1 = k) \mathbb{P}(\tilde{S}_1 > -k, \dots, \tilde{S}_{n-2} > -k, \tilde{S}_{n-1} = -k) \\
 &= \sum_{k>0} \mathbb{P}(X_1 = k) \frac{k}{n-1} \mathbb{P}(\tilde{S}_{n-1} = -k)
 \end{aligned}$$

où S et \tilde{S} ont même loi. Cette dernière expression n'est pas toujours simple à calculer mais, si les variables X sont centrées et possèdent une variance finie, en utilisant le théorème central limite local, on voit que cette probabilité est de l'ordre de $n^{-3/2}$. On retrouve donc facilement le fait que la longueur d'une excursion est p.s. finie mais de moyenne infinie.

Les processus décrits ci-dessous sont obtenus comme limites de marches aléatoires conditionnées ou non. Les lois des variables aléatoires liées à ces processus (maximum, temps d'atteinte...) sont obtenues soit via des propriétés connues sur le mouvement brownien, soit par des principes d'invariances, comme limite de lois analogues sur les marches aléatoires.

A ce propos, on pourra trouver un petit historique des lois associées aux processus browniens dans l'article de Csáki & Mohanty [15] ainsi qu'une idée sur les méthodes employées dans le papier de Durrett & Iglehart [18].

1.2 Mouvement brownien

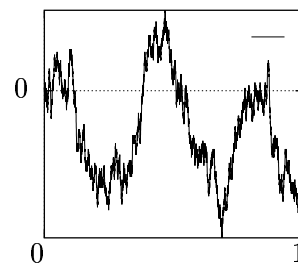
Le mouvement brownien, que l'on notera \mathcal{B} , est un processus défini sur \mathbb{R}^+ , continu, à accroissements indépendants tel que pour tout t , \mathcal{B}_t est une v.a. centrée, gaussienne et a pour variance t . On pourra se rapporter aux livres de Revuz & Yor [41] ou de Karatzas & Shreve [26] pour avoir tous les éléments nécessaires à la bonne compréhension de cet objet. Le livre de Borodin & Salminen [10] donne un ensemble très complet des formules relatives aux temps d'atteinte.

Le mouvement brownien est un objet central en probabilité puisqu'il apparaît, après renormalisation, comme limite d'une large classe de marches aléatoires, de la même façon que la loi normale apparaît comme limite de sommes de v.a. renormalisées. Plus précisément, soit ξ_1, ξ_2, \dots une suite de v.a. indépendantes, de même loi, centrées et de variance σ^2 finie. Pour tout $k \in \mathbb{N}$, notons $S_k = \xi_1 + \dots + \xi_k$ les sommes partielles. On définit alors pour tout $n \in \mathbb{N}$ et $t \geq 0$,

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega)$$

le processus renormalisé, linéaire par morceaux, interpolé de $(S_k)_{k \geq 0}$. Le théorème clé est donné par Donsker :

Théorème 1.1 $X_n \xrightarrow{\text{faiblement}} \mathcal{B}$ dans l'espace des fonctions continues munies de la topologie de la norme uniforme.



C'est à dire que pour toute fonction f continue bornée sur l'ensemble des fonctions continues de \mathbb{R}^+ dans \mathbb{R} ,

$$\lim_n \mathbb{E}(f(S_{[n.]} / \sqrt{n}) = \mathbb{E}(f(\mathcal{B})).$$

On démontre généralement la convergence faible d'une suite de mesures en prouvant la convergence des lois finies dimensionnelles plus le caractère relativement compact de cette suite dans un espace topologique approprié. Pour une démonstration du théorème de Donsker ainsi que pour tout problème relatif aux mesures et aux convergences des mesures, on pourra se référer aux livres de Billingsley [8], [9], Breiman [11] ou Pollard [38] (par exemple pour la convergence faible de $S[n.]/\sqrt{n\sigma^2}$ vers \mathcal{B} . dans l'espace des fonctions continues à droite et limitées à gauche munie de la topologie de Skohorod).

Les principales propriétés du mouvement brownien utilisées ici sont :

- Son caractère markovien et fortement markovien : Pour $T > 0$ déterministe donné, $(\tilde{\mathcal{B}}_s)_{s \geq 0}$ défini par

$$\tilde{\mathcal{B}}_s = \mathcal{B}_{T+s} - \mathcal{B}_T$$

est un mouvement brownien indépendant de ce qui s'est passé avant le temps T :

$$(\tilde{\mathcal{B}}_s)_{s \geq 0} \perp\!\!\!\perp \sigma\{\mathcal{B}_t | 0 \leq t \leq T\}.$$

Cette propriété est également vraie si T est une variable aléatoire positive indépendante du brownien ou si il est un temps d'arrêt pour la filtration brownienne (voir [41]), par exemple, un temps d'atteinte d'une hauteur donnée.

- Ses propriétés de changement d'échelles : Si a est un nombre réel strictement positif et \mathcal{B} un mouvement brownien,

$$\left(\frac{1}{\sqrt{a}} \mathcal{B}_{at}\right)_{t \geq 0} \text{ est un mouvement brownien}$$

- Retournement du temps : pour $T > 0$ déterministe donné,

$$(\mathcal{B}_s)_{0 \leq s \leq T} \stackrel{\text{loi}}{=} (\mathcal{B}_{T-s} - \mathcal{B}_T)_{0 \leq s \leq T}$$

- Soit

$$g = \sup\{t < 1 | \mathcal{B}_t = 0\}$$

le dernier temps avant 1 où le brownien s'annule et

$$\theta = \inf\{t | \mathcal{B}_t = M_1^{(\mathcal{B})}\}$$

la date où il atteint son maximum $M_1^{(\mathcal{B})}$ (cette date est unique p.s.). Alors g et θ suivent la loi de l'arcsinus sur $[0, 1]$, c'est à dire que

$$g \stackrel{\text{loi}}{=} \theta,$$

$$\mathbb{P}(\theta \leq s) = \frac{2}{\pi} \arcsin \sqrt{s}, \quad \mathbb{P}(\theta \in ds) = \frac{ds}{\pi \sqrt{s(1-s)}}$$

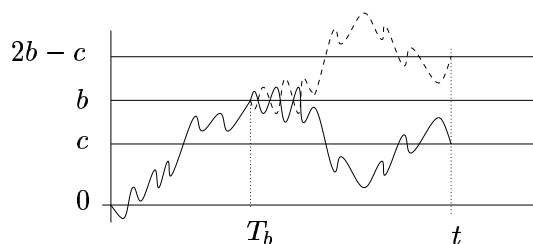
On va donner quelques propriétés de la v.a.

$$T_b = \inf\{t | \mathcal{B}_t = b\}$$

le temps d'atteinte de la hauteur b par le brownien et

$$M_t^{\mathcal{B}} = \max\{\mathcal{B}_s | s \in [0, t]\}$$

le maximum de la trajectoire brownienne sur l'intervalle $[0, t]$. Le principe de réflexion de Désiré André permet d'obtenir des identités entre trajectoires et ainsi d'en déduire simplement des lois de v.a. liées. Il repose uniquement sur le fait que le brownien est fortement markovien à accroissements de loi symétrique; il est donc également valable pour les marches simples symétriques. On peut l'énoncer ainsi : supposons que $T_b \leq t$. Alors, les 2 processus, $(\mathcal{B}_s)_{s \in [0, t]}$ et $\tilde{\mathcal{B}}_s = \begin{cases} \mathcal{B}_s & \text{pour } 0 \leq s \leq T_b, \\ 2b - \mathcal{B}_s & \text{pour } T_b \leq s \leq t, \end{cases}$ ont même loi en tant que processus (et ce sont donc des browniens!). Ainsi, avec cette identité, on montre facilement que :



$$\mathbb{P}(T_b < t \cap \mathcal{B}_t < b) = \mathbb{P}(T_b < t \cap \mathcal{B}_t > b) = \mathbb{P}(\mathcal{B}_t > b).$$

On en déduit par exemple que

$$\mathbb{P}(T_b < t) = \mathbb{P}(T_b < t \cap \mathcal{B}_t < b) + \mathbb{P}(T_b < t \cap \mathcal{B}_t > b) = 2\mathbb{P}(\mathcal{B}_t > b) = \sqrt{\frac{2}{\pi}} \int_{b/\sqrt{t}}^{+\infty} e^{-x^2/2} dx,$$

et donc que

$$\mathbb{P}(T_b \in dt) = |b|(2\pi t^3)^{-1/2} e^{-b^2/2t} dt \quad \text{pour } t > 0.$$

Le processus temps d'atteinte $(T_x)_{x>0}$ est le processus stable d'indice $1/2$. Il est purement discontinu à accroissements stationnaires et indépendants (cf. [41] p.102). On peut déduire ses propriétés de celles du mouvement brownien. Par exemple, les propriétés de changement d'échelles du brownien donnent :

$$T_b \stackrel{\text{loi}}{=} b^2 T_1.$$

\mathcal{B} fortement markovien entraîne que pour $b > a > 0$, on a :

$$T_b - T_a \perp T_a \quad \text{et} \quad T_b - T_a \stackrel{\text{loi}}{=} T_{b-a}.$$

L'espérance mathématique de T_1 est $+\infty$; sa transformée de Laplace est :

$$\mathbb{E}(e^{-a T_b}) = e^{-|b|\sqrt{2a}}, \quad \text{pour } a > 0$$

Dans le plongement de Skohorod, on s'intéresse également à la variable

$$\Sigma_1^\sharp = T_{-1} \wedge T_1,$$

le premier temps de sortie de l'intervalle $[-1, 1]$ par le brownien. Sa transformée de Laplace est :

$$\mathbb{E}(e^{-a(\Sigma_1^\sharp)}) = \frac{1}{\cosh(\sqrt{2a})}$$

et on a

$$\mathbb{P}(\Sigma_1^\sharp \in dt) = (2\pi t^3)^{-1/2} \sum_{n \in \mathbb{Z}} 2(4n+1)e^{-(4n+1)^2/(2t)} dt.$$

T_x et $M_t^{\mathcal{B}}$ sont reliés par l'identité évidente suivante :

$$\{T_b \leq t\} = \{M_t^{\mathcal{B}} \geq b\}.$$

Par le changement d'échelle du mouvement brownien, on a donc :

$$\mathbb{P}(T_b \leq t) = \mathbb{P}(\sqrt{t}M_1^{\mathcal{B}} \geq b) = \mathbb{P}\left(\left(\frac{b}{M_1^{\mathcal{B}}}\right)^2 \leq t\right)$$

et donc,

$$T_b \stackrel{\text{loi}}{=} \frac{b^2}{N^2}$$

où N est une normale centrée réduite. D'autre part, pour tout $t \geq 0$, on a :

$$M_t^{\mathcal{B}} \stackrel{\text{loi}}{=} |\mathcal{B}_t| \stackrel{\text{Lévy}}{=} M_t^{\mathcal{B}} - \mathcal{B}_t.$$

L'égalité de Lévy est également une égalité en loi entre les processus $(|\mathcal{B}_t|)_{t \geq 0}$ et $(M_t^{\mathcal{B}} - \mathcal{B}_t)_{t \geq 0}$. Par contre celle de gauche, n'en est pas une puisque le processus $(M_t)_{t \geq 0}$ est croissant alors que $(|\mathcal{B}_t|)_{t \geq 0}$ ne l'est pas.

1.3 Méandre brownien

Le méandre brownien $(m_t)_{t \in [0,1]}$ est un mouvement brownien conditionné à être strictement positif entre 0 et 1. On peut également le construire à partir du brownien, en renormalisant la portion du brownien comprise entre 1 et son dernier passage en 0 avant 1. Autrement dit,

$$\left(\frac{|\mathcal{B}_{g+(1-g)t}|}{\sqrt{1-g}} \right)_{t \in [0,1]}$$

est un méandre brownien, qui de plus est indépendant de g . Notons

$$M^+ = \sup\{m_t \mid 0 \leq t \leq 1\}$$

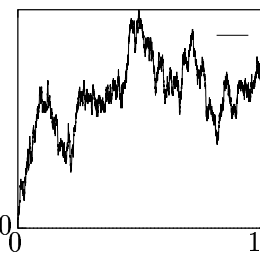
le maximum du méandre. Durrett et Iglehart [18] ont montré

que :

$$\mathbb{P}(M^+ \leq x, m_1 \leq y) = \sum_{k \in \mathbb{Z}} \left(e^{-(2kx)^2/2} - e^{-(2kx+y)^2/2} \right).$$

On peut en déduire que

$$\mathbb{P}(M^+ \leq x) = 1 + 2 \sum_{k \in \mathbb{N}^*} (-1)^k e^{-(kx)^2/2} \quad \text{et} \quad \mathbb{E}(M^+) = \sqrt{2\pi} \log 2.$$



1.4 Pont brownien

Le pont brownien $(b_t)_{t \in [0,1]}$ est un mouvement brownien conditionné à être nul en 0 et en 1. Il peut être construit de diverses manières à partir d'un mouvement brownien standard \mathcal{B} :

$$\begin{cases} b_t = \mathcal{B}_{gt}/\sqrt{g} & \text{pour } 0 \leq t \leq 1, \\ \tilde{b}_t = \mathcal{B}_t - t\mathcal{B}_1 & \text{pour } 0 \leq t \leq 1, \\ \tilde{\tilde{b}}_t = (1-t)\mathcal{B}_t/(1-t) & \text{pour } 0 \leq t \leq 1 \end{cases}$$

sont trois versions du pont brownien. De plus b est indépendant de g .

On peut également le construire comme solution d'une équation

$$\text{différentielle stochastique : } b_t = \begin{cases} \mathcal{B}_t - \int_0^t \frac{b_s}{1-s} ds, & \text{pour } t \in [0, 1[\\ = 0, & \text{pour } t = 1. \end{cases}$$

• La date où le pont atteint son maximum suit la loi uniforme sur l'intervalle $[0, 1]$.

• La loi du maximum est :

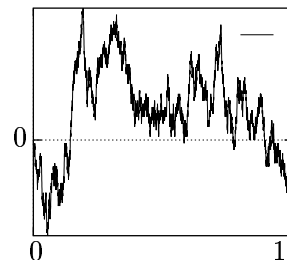
$$\mathbb{P}(\sup\{b_t \mid 0 \leq t \leq 1\} \leq x) = (1 - e^{-2x^2}) \mathbb{I}_{x \geq 0}.$$

Autrement dit, $\sup\{b_t \mid 0 \leq t \leq 1\}$ a même loi que $R/2$ si R suit la loi de Rayley.

• La loi du maximum de $|b_t|$, dont la distribution est dite de Kolmogorov-Smirnov est donnée par :

$$\mathbb{P}(\sup\{|b_t| \mid 0 \leq t \leq 1\} \leq x) = \sum_{n \in \mathbb{Z}} (-1)^n e^{-n^2 x} = \prod_{n \in \mathbb{N}^*} \frac{1 - e^{-nx}}{1 + e^{-nx}}.$$

L'importance de cette distribution (en statistique, en particulier) est évoquée dans le chapitre IV.



1.5 Excursion brownienne renormalisée

Elle peut être construite de diverses manières :

- C'est un mouvement brownien conditionné à vérifier $\mathcal{B}_0 = 0, \mathcal{B}_1 = 0$ et $\mathcal{B}_t > 0$ pour tout $t \in]0, 1[$.

- Une renormalisation de la portion de trajectoire brownienne qui "enjambe" 1. Autrement dit, considérons $d = \inf\{t \geq 1 | \mathcal{B}_t = 1\}$ le dernier moment avant 0 où le brownien est nul et g défini plus haut. Alors

$$\left(\frac{|\mathcal{B}_{g+(d-g)t}|}{\sqrt{d-g}} \right)_{t \in [0,1]}$$

est une excursion brownienne renormalisée.

- Un pont de Bessel de dimension 3 : soit $(\mathcal{B}_t^{(3)})_{t \geq 0}$ un mouvement brownien tridimensionnel conditionné à vérifier

$\mathcal{B}_0^{(3)} = 0$ et $\mathcal{B}_1^{(3)} = 0$; $e_t = |\mathcal{B}_t^{(3)}|$ est une excursion brownienne renormalisée.

- De la même manière que le mouvement brownien apparaît comme limite faible de marches aléatoires, l'excursion brownienne apparaît comme limite faible d'excursions discrètes (avec la même renormalisation).

- La loi du maximum de e . est donnée par :

$$\begin{aligned} \mathbb{P}(\sup\{e_s | 0 \leq s \leq 1\} \leq x) &= \sum_{j \in \mathbb{Z}} (1 - 4j^2 x^2) e^{-2j^2 x^2} \\ &= \sqrt{2\pi}^{5/2} x^{-3} \sum_{j=0}^{+\infty} j^2 e^{-j^2 \pi^2 / (2x^2)} \\ &= \left(\frac{\pi^3}{2}\right)^{1/2} \theta'\left(\frac{\pi}{2x^2}\right) \end{aligned}$$

où

$$\theta(x) = \sum_{n \in \mathbb{Z}} e^{-\pi n^2 x}$$

est la fonction *theta* de Jacobi (voir Chung [13]). Ceci permet de calculer les moments :

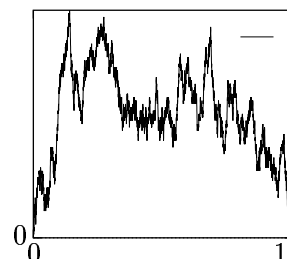
$$\mathbb{E}(\sup\{e_s | 0 \leq s \leq 1\}) = \sqrt{\pi/2} \quad \text{et}$$

$$\mathbb{E}(\sup\{e_s | 0 \leq s \leq 1\}^r) = 2^{-r/2} r(r-1)\Gamma(r/2)\zeta(r) \quad \text{pour } r > 1.$$

- La loi de l'aire sous l'excursion brownienne et celle de son temps local sont également connues. On pourra voir à ce sujet les articles de Louchard [33], [34], [35], Takács [48], [50].

1.6 Décompositions

Il existe de nombreuses décompositions des trajectoires des processus browniens comme le montre l'article de Bertoin et Pitman [6]. On va essayer ici de rendre plus visuelles les transformations utilisées ici.

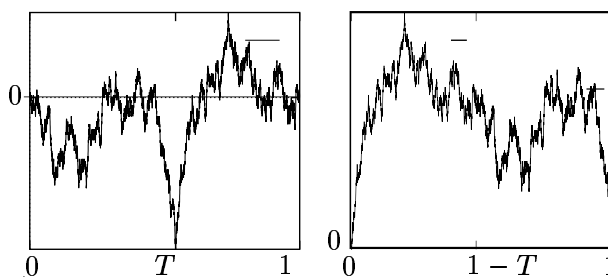


Pont brownien et excursion.

Transformation trajectorielle de Vervaat [51] : Soit b un pont brownien et U la date où il atteint son minimum. U suit une loi uniforme sur $[0, 1]$. Définissons le processus e sur $[0, 1]$:

$$e_t = \begin{cases} b_{U+t} - b_U, & \text{pour } 0 \leq t \leq 1 - U, \\ b_{t-(1-U)} - b_U, & \text{pour } 1 - U \leq t \leq 1. \end{cases}$$

Alors, d'après Vervaat (1979), U et e sont indépendants et e est une excursion brownienne.



Réciproquement, prenons U uniforme sur $[0, 1]$ et e une excursion brownienne normalisée indépendante de U ; alors le processus b défini par

$$b_t = \begin{cases} e_{U+t} - e_U, & \text{pour } 0 \leq t \leq 1 - U, \\ e_{t-(1-U)} - e_U, & \text{pour } 1 - U \leq t \leq 1. \end{cases}$$

est un pont brownien. Une conséquence directe de cette décomposition de trajectoires est :

$$\max\{b_t | 0 \leq t \leq 1\} - \min\{b_t | 0 \leq t \leq 1\} \stackrel{\text{loi}}{=} \max\{e_t | 0 \leq t \leq 1\}.$$

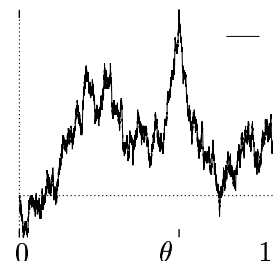
On pourra aussi trouver cette réciproque dans l'article de Biane [7].

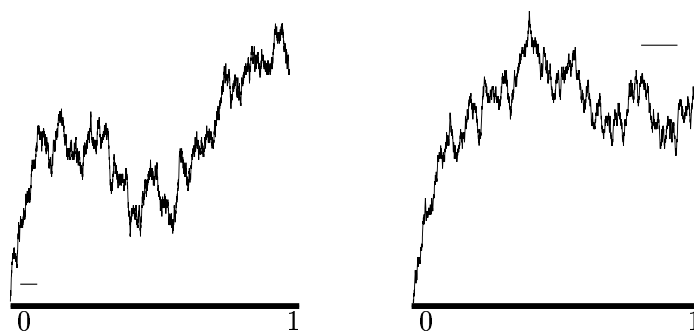
Décomposition du brownien en deux méandres (Denisov)

Soit \mathcal{B} un mouvement brownien sur l'intervalle $[0, 1]$. Notons θ l'abscisse où le maximum de \mathcal{B} est atteint. Les deux processus $(m_t^+)_{t \in [0,1]}$ et $(m_t^-)_{t \in [0,1]}$ définis par :

$$\begin{cases} m_t^+ = \frac{\mathcal{B}_\theta - \mathcal{B}_{\theta+t(1-g)}}{\sqrt{1-g}} & \text{pour } 0 \leq t \leq 1, \\ m_t^- = \frac{\mathcal{B}_\theta - \mathcal{B}_{\theta-tg}}{\sqrt{g}} & \text{pour } 0 \leq t \leq 1, \end{cases}$$

sont deux méandres browniens. Autrement dit, on décompose chaque trajectoire brownienne sur $[0, 1]$ en deux morceaux de part et d'autre du maximum. La partie à droite du maximum est ramenée sur $[0, 1]$ par une symétrie par rapport à l'axe des abscisses, par un changement d'origine et par le changement d'échelle habituel du brownien. La partie à gauche du maximum est ramenée sur $[0, 1]$ après avoir subi une rotation d'angle π , un changement d'origine et un changement d'échelle. Les deux méandres ci-dessous sont les méandres associés au brownien ci-contre. Le méandre de droite correspondant à la partie à droite du maximum du brownien.





2 Arbres - Processus de Galton-Watson

2.1 Processus de Galton-Watson

Harris [25], Athreya & Ney [5] et Asmussen & Hering [4] donnent un bon panorama des propriétés de ces objets.

Un processus de Galton-Watson est une chaîne de Markov $(Z_n)_{n=0,1,2,\dots}$ sur \mathbb{N} telle que $Z_0 = 1$ et telle que pour $k \geq 0$, Z_{k+1} s'écrit sous la forme :

$$Z_{k+1} = \sum_{l=1}^{Z_k} Y_l^{(k)}$$

où les variables $Y_l^{(k)}$ sont des copies indépendantes de Z_1 , de loi $(p_k)_{k=0,1,\dots}$ (i.e. pour tout $j \geq 0$, $\mathbb{P}(Z_1 = j) = p_j$). Les probabilités de transition de la chaîne de Markov $(Z_n)_{n \geq 0}$ s'exprime donc à partir de la suite $(p_k)_{k=0,1,\dots}$:

$$\mathbb{P}(i, j) = \mathbb{P}(Z_{n+1} = j | Z_n = i) = \begin{cases} p_j^{*i} & \text{si } i \geq 1, j \geq 0, \\ \delta_{0j} & \text{si } i = 0, j \geq 0 \end{cases}$$

où δ_{0j} est le symbole de Kronecker, $p_j^{*1} = p_j$ et p_j^{*i} est la probabilité que la somme de i copies indépendantes de Z_1 ait pour valeur j . En général, on donne la représentation concrète suivante de cette chaîne de Markov :

- Au temps $T = k$, Z_k individus sont en vie.
- Au temps $T = k + 1$, chacun des Z_k individus disparaît laissant chacun j descendants avec probabilité p_j indépendamment des autres individus. La loi $(p_k)_{k=0,1,\dots}$ de Z_1 est donc appelé loi de la *progéniture*. Le principal outil d'étude de ces processus est la fonction génératrice f de la loi de la progéniture $(p_k)_{k \geq 0}$:

$$f(s) = \sum_{k \geq 0} p_k s^k \text{ pour } |s| \leq 1.$$

En effet, par des considérations simples, il apparaît que la fonction génératrice f_n de Z_n s'écrit :

$$f_0(s) = s, f_1(s) = f(s), \dots, f_{n+1}(s) = f(f_n(s)).$$

Écartons les cas ($p_0 + p_1 = 1$) ainsi que le cas où il existe j tel que $p_j = 1$; les *moments* des v.a. Z_n peuvent s'exprimer en fonction des $(p_j)_{j \geq 0}$.

$$\begin{aligned}\mathbb{E}(Z_1) &= f'(1) \stackrel{\text{def}}{=} m \\ \mathbb{E}(Z_n) &= f'_n(1) = f'_{n-1}(1)f'(1) = (f'(1))^n = m^n\end{aligned}$$

et, en notant $\text{Var}(Z_1) = \sigma^2$ et en utilisant le fait que

$$f''_n(1) = f''(1)(f'_n(1))^2 + f'(1)f''_n(1)$$

on obtient que

$$\text{Var}(Z_n) = \begin{cases} \frac{\sigma^2 m^{n-1}(m^n-1)}{m-1} & \text{si } m \neq 1, \\ n\sigma^2 & \text{si } m = 1. \end{cases}$$

On appelle *date d'extinction* du processus $(Z_n)_{n \geq 0}$ le premier temps T tel que $Z_T = 0$ (selon les définitions ci-dessus, si $Z_T = 0$ alors $Z_{T+k} = 0, \forall k \geq 0$). Il est clair que si $p_0 > 0$, la *probabilité q d'extinction* du processus est strictement positive ($Z_1 = 0$ avec probabilité p_0). La probabilité d'extinction q vaut 1 si $m \leq 1$ (si $p_1 \neq 1$) et est la plus petite solution positive de l'équation $f(t) = t$ si $m > 1$. Le comportement du processus $(Z_k)_{k \geq 0}$ est donc fortement lié à la valeur de m . Le théorème suivant l'illustre une fois encore :

Théorème 2.1 notons $W_n = Z_n m^{-n}$ et F_n la σ -algèbre engendrée par Z_0, \dots, Z_n . Si $0 < m < +\infty$ alors la suite $\{W_n, F_n\}_{n \geq 0}$ est une martingale.

Puisque W_n est positive, il existe une v.a. W telle que :

$$\lim_{n \rightarrow \infty} W_n = W \quad \text{p.s..}$$

Dans le cas, où $m \leq 1$, puisque $q = 1$, il apparaît clairement que la v.a. W vaut 0 et donc que le facteur normalisateur m^n est trop grand. Dans le cas $m > 1$, cette variable n'est pas dégénérée et la croissance moyenne du processus $(Z_k)_{k \geq 0}$ est donc exponentielle.

On appelle *population totale* X la v.a. définie par :

$$X = Z_0 + Z_1 + Z_2 + \dots$$

Au temps 1, on a Z_1 individu(s). Chacun de ces Z_1 individus va avoir un nombre total de descendants de même loi que X . On a donc

$$X = 1 + \sum_{k=0}^{Z_1} X_k$$

où les X_k sont indépendants et ont même loi que X . Soit g la fonction génératrice de X ; g est solution de l'équation fonctionnelle :

$$g(x) = x f(g(x)).$$

Si $\mathbb{E}(Z_1) < 1$, on peut en déduire la population moyenne car :

$$\mathbb{E}(X) = g'(1-) = \frac{1}{1 - f'(1)} = \frac{1}{1 - \mathbb{E}(Z_1)}.$$

Si $\mathbb{E}(Z_1) \geq 1$, la population moyenne est infinie; on peut le voir aisément avec le théorème 2.1. Pour la distribution de X on pourra se référer aux articles de Dwass [20], Kennedy [29] ou Otter [37]; on a :

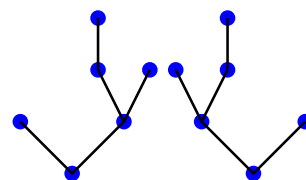
$$\mathbb{P}(X = k) = \frac{1}{k} \mathbb{P}(Z_1^{(1)} + \dots + Z_1^{(k)} = k - 1)$$

où les $Z_1^{(j)}$ sont des copies indépendantes de Z_1 . On peut également exprimer la loi de X en fonction de f et de g :

$$\begin{aligned} \mathbb{P}(X = k) &= \text{coefficient de } x^k \text{ dans } g(x) \\ &= \frac{1}{k} (\text{coefficient de } x^{k-1} \text{ dans } f(x)^k). \end{aligned}$$

2.2 Arbres de Galton-Watson

Le processus de G.W. défini ci-dessus peut, comme on l'a dit un peu plus haut, être vu comme l'évolution d'une population soumise à certaines règles de reproduction. L'arbre généalogique de cette population est appelé arbre de G.W.. Ce sont des arbres ordonnés. Sur le dessin ci-contre les deux arbres sont différents. Comme l'indique Aldous [1], on a tendance à considérer que le processus de G.W. et l'arbre ne sont qu'un seul objet, c'est à dire que l'on ne réduit pas le processus aux largeurs successives de l'arbre, mais que l'on s'intéresse aussi à qui est le fils de qui.



Dans cette thèse, on s'est souvent intéressé aux arbres de G.W. conditionnés par leur population totale X . Il est quelque peu surprenant, au premier abord, que la loi induite sur les arbres binaires (ou sur les arbres à progéniture géométrique) obtenues par le conditionnement $X = n$ soit la loi uniforme. L'explication qui suit est empruntée à Aldous [1].

Prenons le cas de l'ensemble des arbres T de G.W. non conditionnés de loi de progéniture $(p_i)_{i \geq 0}$. On a alors :

$$\begin{aligned} \mathbb{P}(T = t) &= \prod_{v \text{ noeuds}} p_{d(v,t)} \\ &= \prod_{i \geq 0} p_i^{D_i(t)} \\ &\stackrel{\text{def}}{=} w(t) \end{aligned} \tag{2.1}$$

où $d(v, t)$ désigne le nombre de descendants de v dans l'arbre t , $D_i(t)$ le nombre de noeuds ayant i descendants. Ainsi

$$\mathbb{P}(T = t \mid |t| = n) \text{ est proportionnel à } w(t) \text{ sur } \{t \mid |t| = n\},$$

où $|t|$ désigne la population totale de t .

- Dans le cas des *arbres binaires* ayant $2n + 1$ noeuds, une analyse simple montre que n de ces noeuds sont des noeuds internes et les $n + 1$ autres sont des feuilles. Ainsi, lorsque $X = 2n + 1$, $w(t)$ vaut $p_0^{n+1} p_2^n$ qui ne dépend pas de la structure de l'arbre t . Ainsi, tous les arbres binaires obtenus sont équiprobables et cette construction est donc équivalente à celle de la combinatoire.

- Prenons le cas des *arbres géométriques*, avec $X = n$. Il existe un réel $0 < p < 1$ tel que p_i s'écrive $p_i = p^i(1 - p)$ pour tout $i \in \mathbb{N}$; on a alors,

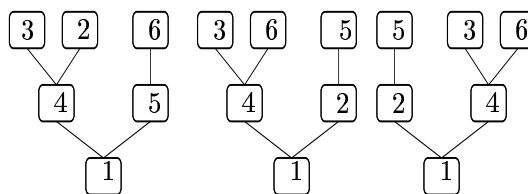
$$\begin{aligned} w(t) &= \prod_{i \geq 0} p_i^{D_i(t)} = (1 - p)^n p^{\sum_v \text{noeuds } d(v,t)} \\ &= (1 - p)^n p^{n-1} \end{aligned}$$

car chaque noeud sauf la racine est le fils d'un autre noeud. Ainsi, $w(t)$ ne dépend pas de l'arbre t et donc les arbres obtenus sont équiprobables; il est alors aisé de remarquer que l'on obtient avec cette loi tous les arbres ordonnés. Cette construction est donc équivalente à celle des arbres généraux de la combinatoire (c'est à dire des arbres planaires ordonnés).

Ces deux cas particuliers ne doivent pas faire oublier le cas général, c'est à dire, qu'en général, la loi induite sur les arbres de G.W. par le conditionnement $X = n$ n'est pas la loi uniforme, car il n'y a aucune raison pour que $\prod_{i \geq 0} p_i^{D_i(t)}$ ne dépende pas de t . On pourra se référer à Aldous [1] p. 26-29 pour comprendre comment construire les classes d'arbres standards de la combinatoire à partir d'arbres de G.W. conditionnés.

- Le cas des *arbres étiquetés* non ordonnés avec racine à n noeuds est un peu à part. On pourra voir Moon [36] pour tout ce qui concerne ces arbres. Un arbre ayant n noeuds étiqueté est un graphe connexe sur l'ensemble des sommets $\{1, \dots, n\}$ ayant $n - 1$ arêtes. Il y a n^{n-2} arbres de ce type d'après Cayley (1889). Il n'y a pas pour ces arbres de racines a priori. Si l'on choisit le noeud 1 pour être la racine, on obtient n^{n-2} arbres. Si l'on considère tous les arbres avec toutes les racines possibles, il y en a n^{n-1} .

Puisqu'il n'y a pas de structure d'ordre entre les fils de chaque noeuds, on peut représenter graphiquement le même arbre de plusieurs manières. Parfois, on choisit pour différencier plus facilement les arbres de les représenter en ordonnant les étiquettes des fils de chaque noeuds.



On obtient ainsi une représentation de chaque classe d'équivalence. Ainsi, sur le dessin ci-contre, les deux premiers arbres sont différents, et les arbres 2 et 3 sont les mêmes. Le 3^{eme} a ses étiquettes ordonnés et est donc le représentant dont on vient de parler. Un arbre étiqueté de racine 1, est entièrement défini par la donnée d'un n -uplet d'ensembles disjoints (A_1, A_2, \dots, A_n) où les A_i ont pour union l'ensemble $\{2, \dots, n\}$. Pour construire l'arbre, on choisit un sens de parcours sur celui-ci et on affecte au i^{eme} noeud parcouru l'ensemble de fils d'étiquettes A_i que l'on range par ordre croissant. Par exemple, choisissons le parcours en largeur (défini dans la section suivante); les deux arbres 2 et 3 sont définis par le 6-uplet :

$(\{2, 4\}, \{5\}, \{3, 6\}, \emptyset, \emptyset, \emptyset)$. Tout n -uplet d'ensembles (A_1, A_2, \dots, A_n) ayant les propriétés ci-dessus ne définit pas pour autant nécessairement un arbre. Il faut que jusqu'à la fin de la construction, il y ait toujours des noeuds "en attente" de fils. Par exemple, si $A_1 = \emptyset$ alors la construction est tout de suite finie. Pour construire un arbre avec n noeuds (de racine 1), en plus des conditions mentionnées ci-dessus il faut que les cardinaux des ensembles A_i satisfassent à l'ensemble de conditions suivantes :

$$\#A_1 > 0, \#A_1 + \#A_2 > 1, \dots, \#A_1 + \dots + \#A_{n-1} > n - 2, \#A_1 + \dots + \#A_n = n - 1.$$

Les quantités hauteur et largeur et toutes les autres considérées sur les arbres sans étiquettes ont un sens ici aussi. Pour faire des probabilités, on suppose que chaque arbre a pour probabilité n^{2-n} . Il s'avère que si l'on considère la loi de progéniture des individus d'un arbre étiqueté avec racine ayant n noeuds, c'est à dire uniquement le nombre d'enfants de chaque individu (et non pas leurs étiquettes), cette loi est celle d'un arbre de G.W. avec progéniture poissonnienne (conditionnée à avoir une population totale de n individus). Pour montrer ceci, on compte le nombre d'arbres dont les ensembles $(A_i)_{i=1, \dots, n}$ ont pour cardinaux $(a_i)_{i=1, \dots, n}$. Il est immédiat que le nombre de tels arbres est donné par :

$$\frac{(n-1)!}{a_1! \cdots a_n!};$$

la probabilité d'apparition d'un tel arbre est donc proportionnelle à ce nombre. Si on étudie les arbres de G.W. à progéniture poissonnienne $\mathcal{P}(\lambda)$, $\lambda > 0$, la probabilité qu'un arbre présente une suite $(a_i)_{i=1, \dots, n}$ (de descendants) est quant à elle proportionnelle à

$$\frac{e^{-n\lambda} \lambda^n}{a_1! \cdots a_n!}.$$

Ainsi, ces deux probabilités sont nécessairement les mêmes. Les lois des largeurs et hauteurs des arbres étiquetés sont donc les mêmes que celle des arbres de G.W. de loi de progéniture poissonnienne.

Une remarque de Kennedy

On a vu à plusieurs reprises l'importance que revêt la valeur de la moyenne de Z_1 dans le comportement général des processus de G.W.. Pourtant, on vient de voir que dans les trois cas ci-dessus la moyenne n'a joué aucun rôle. En effet, la loi sur les arbres (conditionnés par $X = n$) était la même pour toute valeur du paramètre dans la variable de Bernoulli, géométrique ou de Poisson. Ce phénomène se produit quel que soit la loi de la progéniture comme l'a montré Kennedy [29]. Plus précisément, soit $(Z_n)_{n \geq 0}$ un processus de G.W. de loi de progéniture $(p_k)_{k \geq 0}$ et de fonction génératrice f . Définissons une nouvelle loi de progéniture $(\tilde{p}_k)_{k \geq 0}$ par

$$\tilde{p}_k \stackrel{def}{=} \frac{a^k p_k}{f(a)}.$$

Alors le processus associé à cette loi, $(\tilde{Z}_n)_{n \geq 0}$ vérifie :

$$\mathbb{P}(\tilde{Z}_{k_1} = r_1, \dots, \tilde{Z}_{k_j} = r_j | \tilde{X} = n) = \mathbb{P}(Z_{k_1} = r_1, \dots, Z_{k_j} = r_j | X = n).$$

Les processus $(Z_n)_{n \geq 0}$ et $(\tilde{Z}_n)_{n \geq 0}$ ont donc même loi conditionnellement à leur population totale. Par un calcul simple, on remarque que :

- Si Z_1 est une variable de “Bernoulli” telle que $p_0 = 1 - p, p_2 = p$, alors $\tilde{p}_2 = a^2 p / (1 - p + a^2 p)$ et $\tilde{p}_0 = 1 - p_2$ donc \tilde{Z}_1 est une variable de “Bernoulli”.
- Si Z_1 est une variable géométrique de paramètre p , \tilde{Z}_1 est une variable géométrique de paramètre ap .
- Si Z_1 est une variable de Poisson $\mathcal{P}(\lambda)$, \tilde{Z}_1 est une variable de Poisson $\mathcal{P}(a\lambda)$.

En faisant varier a on obtient ainsi tous les paramètres de ces variables.

La (ou les) valeur de a telle que $af'(a) = f(a)$ fournit le cas *critique*, c’est à dire le cas ou $\tilde{m} = 1$. La variance de \tilde{Z}_1 , $\tilde{\sigma}^2$ associée à cette valeur de a s’écrit :

$$\tilde{\sigma}_2 = \frac{a^2 f''(a)}{f(a)}.$$

Elle joue un rôle normalisateur particulier comme on le verra dans les sections suivantes.

2.3 Hauteur et largeur des arbres

La *hauteur* d’un arbre de G.W. est définie comme la plus grande distance entre un noeud et la racine. Elle représente donc le nombre de générations non vides du processus moins un.

Dans le cas des arbres non conditionné, on peut donc calculer la loi de la hauteur en utilisant la fonction génératrice f . En notant H la hauteur, on a :

$$\mathbb{P}(H \leq h) = \mathbb{P}(Z_{h+1} = 0) = f_{h+1}(0).$$

Plaçons nous dans le cadre des arbres de G.W. conditionnés par leur population totale $X = n$. La loi et les moments de la hauteur des arbres ont été obtenus par Flajolet & Odlyzko [21] dans un cadre assez général (voir également Renyi et Szekeres [39], Riordan [42]). De façon assez générale, en notant H la hauteur, on obtient que :

$$\lim_n \mathcal{L}\left(\frac{H}{\sqrt{n}} \mid X = n\right) \xrightarrow{\text{loi}} c M_1^{(e)}$$

où $M_1^{(e)}$ est le maximum de l’excursion brownienne normalisée et c est une constante non nulle dépendant de la loi de reproduction du processus. Cette convergence est également une convergence pour tous les moments. c s’exprime en fonction de la valeur de $\tilde{\sigma}$ définie ci-dessus :

$$c = \frac{2}{\tilde{\sigma}}.$$

Ainsi, dans le cas des arbres binaires c vaut 2, dans le cas des arbres généraux c vaut $\sqrt{2}$, dans le cas des arbres étiquetés c vaut 2.

On trouvera de nombreux éléments sur la largeur des arbres dans le chapitre IV. On va donc se contenter de rappeler ici quelques résultats et définitions.

Le *profil* d’un arbre est la suite $(Z_k)_{k \geq 0}$. La *largeur* W d’un arbre est définie comme étant le maximum du profil, i.e. $W = \max_k Z_k$; W désigne donc le nombre maximum d’individus

présents dans une seule génération du processus, ou sur un niveau de l'arbre. On peut remarquer que contrairement à la loi de la hauteur, la loi de la largeur ne s'exprime pas simplement à partir des fonctions génératrices f_n .

Dans le cas des arbres de G.W. conditionnés par $X = n$, Drmota & Gittenberger [16] ont montré que le profil de l'arbre (divisé par \sqrt{n}) convergeait en tant que processus vers le temps local de l'excursion brownienne; c'est à dire que

$$\lim_n \mathcal{L}\left(\frac{Z_{\lfloor \sqrt{n} t \rfloor}}{\sqrt{n}} \middle| X = n\right) \xrightarrow{\text{faiblement}} \frac{\tilde{\sigma}}{2} l_1^{(e)}\left(\frac{\tilde{\sigma}}{2} t\right)$$

où $l_1^{(e)}(t)$ est le temps local au temps 1 à la hauteur t d'une excursion brownienne normalisée. $l_1^{(e)}(t)$ est défini par

$$l_1^{(e)}(t) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^1 \mathbb{I}_{[t, t+\varepsilon]}(e_s) ds.$$

Or, d'après Jeulin [24], ou [1], [12], on a :

$$\sup\{l_1^{(e)}(t) | t \geq 0\} \stackrel{\text{loi}}{=} 2 \sup\{e_t | t \geq 0\}$$

La loi de la largeur proprement dite se déduit donc de ce résultat ; il apparaît donc que

$$\lim_n \mathcal{L}\left(\frac{W}{\sqrt{n}} \middle| X = n\right) \xrightarrow{\text{loi}} \tilde{\sigma} M_1^{(e)}.$$

On pourra également voir les articles de Takács [49], d'Aldous [1] ainsi que le chapitre IV pour des approches probabilistes et pour des résultats sur la convergence des moments de W/\sqrt{n} .

2.4 Parcours d'arbres

Parcourir un arbre signifie suivre ses arêtes, selon un ordre prédéfini, afin de visiter chacun de ses sommets. On pourra se référer à Cormen et al. [14], chapitre 23 ou à Flajolet & Sedgewick [43] chapitre 5.

Comme on va le voir dans les exemples suivants, certains parcours sont tout à fait adaptés pour étudier la structure des arbres (et des graphes en général). L'idée est que l'on peut, grâce à des parcours particulier, associer (de manière bijective) à chaque arbre une excursion discrète de sorte que les valeurs associées à l'arbre (largeur ou hauteur ou longueur de cheminement total...) se traduisent par des valeurs associées à l'excursion (maximum ou temps local ou aire...). On utilise alors les résultats connus sur les excursions pour en déduire des propriétés équivalentes sur les arbres.

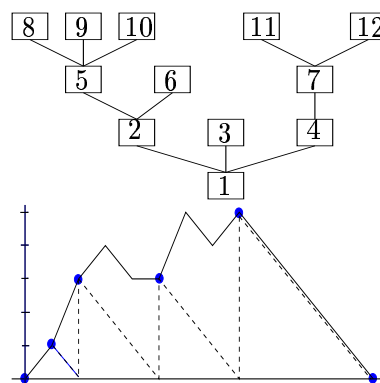
Parcours en largeur d'abord (Breadth first search)

Dans le parcours en largeur, on visite successivement les niveaux de l'arbre, de la racine vers le sommet, en visitant chaque noeud du plus à gauche au plus à droite. Sur l'arbre ci-dessous, on a indiqué l'ordre de visite des noeuds. On peut, grâce à ce parcours, associer à

chaque arbre une excursion discrète sur laquelle le profil de l'arbre apparaît être des points particuliers.

Pour cela, on maintient une file d'attente des fils des noeuds visités. Par exemple, ici, au début, la file est réduite à $\{1\}$. On visite 1, qui est effacé de la file, on ajoute $\{2, 3, 4\}$; on visite ensuite 2, que l'on efface de la file, on ajoute 5 et 6... de cette façon, on voit bien, qu'à chaque étape, que le nombre de noeuds dans la file est réduit de 1 et est augmenté de a_i , a_i étant le nombre d'enfants du i^{eme} individu visité. Puisque la file n'est vide qu'à la fin du parcours, on obtient une excursion discrète (voir chapitre IV).

Plaçons nous dans le cas des arbres de G.W. conditionnés par $X = n$. L'excursion associée est de longueur $n + 1$; notons la $(S_k)_{k=0, \dots, n+1}$. On a $S_0 = 0, S_1 = 1$ et les autres incréments ont pour loi, la loi de $\tilde{Z}_1 - 1$ et sont conditionnés à former une excursion de longueur $n + 1$ (voir la remarque de Kennedy). Ainsi, on a :



$$\frac{S_{[nt]}}{\sqrt{n}} \xrightarrow{\text{faiblement}} \tilde{\sigma} e_t$$

où $(e_t)_{t \in [0,1]}$ est une excursion brownienne normalisée et donc :

$$\frac{\max_k S_k}{\sqrt{n}} \xrightarrow{\text{loi}} \tilde{\sigma} M_1^{(e)}.$$

On voit alors sur le dessin (et on peut le prouver rigoureusement) que la largeur de l'arbre et la hauteur de l'excursion sont très proches l'une de l'autre et ont donc la même limite en loi (divisées par \sqrt{n}), $\tilde{\sigma} M_1^{(e)}$.

Grâce à ce parcours, dans le chapitre III, on montre également que la hauteur de l'arbre et la quantité

$$\sum_{k=1}^n \frac{1}{S_k + 1}$$

sont très proches dans le cas des arbres binaires (où $(S_k)_{k=1, \dots, n}$ est la suite des valeurs successives de l'excursion). On montre ensuite que cette quantité normalisée converge en loi vers

$$\int_0^1 \frac{1}{e_s} ds$$

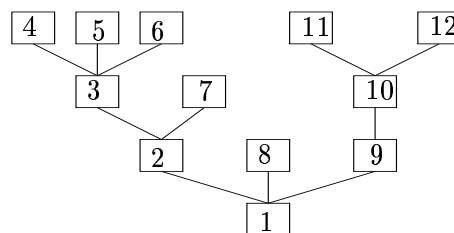
ce qui permet donc de retrouver les résultats de Flajolet & Odlyzko [21] dans ce cas particulier. En effet, selon Biane & Yor [12] (voir aussi Chapitre III), on a

$$\int_0^1 \frac{1}{e_s} ds \stackrel{\text{loi}}{=} 2 M_1^{(e)}.$$

On pourra également trouver des utilisations de ce parcours en largeur dans Aldous [1], [3] (pour l'étude des composantes géantes dans les graphes aléatoires), Takács [49] (file d'attente et largeur d'arbres), Spencer [45] (arbres étiquetés, graphes et loi de poisson), Schutzenberger [44], Foata & Riordan [22] et Françon [23] pour les relations entre fonctions acycliques, problèmes de parking, arbres étiquetés...

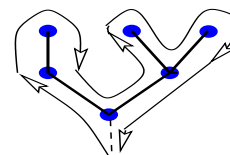
Parcours en profondeur d'abord (Depth first search)

On visite d'abord la racine. Plaçons nous au temps k après avoir visité k noeuds. Certains de ces noeuds ont des fils non visités. On prend le plus récemment visité parmi ces noeuds. Au temps $k + 1$, on visitera son fils le plus à gauche non encore visité.



On associe une excursion à chaque arbre en reportant l'ordonnée d'une particule allant du k^{eme} noeud visité au $k + 1^{eme}$ contrainte à rester sur l'arbre. (Pour être précis on fait comme si elle partait d'un noeud situé sous la racine). L'excursion discrète ainsi obtenue ne présente que des incréments $+1$ ou -1 et est de longueur $2n$ si l'arbre a n noeud.

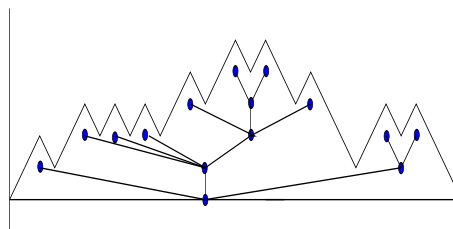
Dans le cas d'arbres de G.W. conditionnés par leur population totale $X = n$, cette excursion converge (renormalisée par \sqrt{n}) vers une excursion brownienne; plus précisément, notons $(S_k)_{k=0, \dots, 2n}$ les valeurs successives de l'excursion associée. Aldous [1], [2] a prouvé que :



$$\frac{S_{[2nt]}}{\sqrt{n}} \xrightarrow{\text{faiblement}} \frac{2e_t}{\tilde{\sigma}}$$

où $(e_t)_{t \in [0,1]}$ est une excursion brownienne normalisée et $\tilde{\sigma}$ a été défini plus haut.

On peut également, à partir d'une excursion de Bernoulli, construire un arbre. Si l'on considère les marches aléatoires simples symétriques, on peut associer à la première excursion positive de celles-ci un arbre. L'arbre obtenu a pour loi celle d'un arbre de G.W. avec progéniture de loi géométrique de paramètre $1/2$ non conditionné (voir Dwass [19]).



- On peut remarquer qu'avec cette correspondance, la différence entre la hauteur de l'arbre et le maximum de l'excursion associée vaut 1. On comprend donc aisément pourquoi, la hauteur d'un arbre de G.W. de taille n divisée par \sqrt{n} converge vers le maximum de l'excursion brownienne (au facteur $2/\tilde{\sigma}$ près).

- Ce parcours permet également d'étudier la *longueur de cheminement total* de l'arbre. Soit $V(T)$ l'ensemble des sommets de l'arbre T et pour tout $x \in V(T)$, on note $h(x)$ la distance de x à la racine. La longueur de cheminement total $LC(T)$ est la somme des distances des

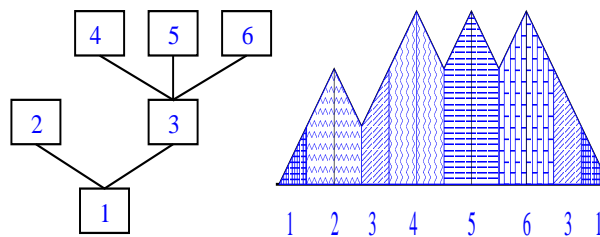
noeuds à la racine, i.e.

$$LC(T) = \sum_{x \in V(T)} h(x).$$

Comme on l'a dit dans la section 5 du Chapitre III, on peut associer à chaque noeud $x \in V(T)$ deux pas de l'excursion, $[S_{k_1}, S_{k_1} + 1]$ et $[S_{k_2}, S_{k_2} + 1]$ de sorte que

$$S_{k_1} \wedge S_{k_1} + 1 = S_{k_2} \wedge S_{k_2} + 1 = h(x).$$

La raison est que l'on parcourt chaque arête de l'arbre deux fois et qu'il y a donc deux pas de l'excursion associés à chaque arête et donc à chaque noeud. A un noeud à la hauteur k , on peut donc associer une partie de la surface sous l'excursion entre les droites d'équations $x = k_1$ et $x = k_1 + 1$ et entre les droites $x = k_2$ et $x = k_2 + 1$. L'aire de cette partie est alors $2h(x) + 1$. L'aire totale A sous l'excursion s'écrit donc :



$$\begin{aligned} A &= \sum_{x \in V(T)} (2 h(x) + 1) \\ &= 2LC(T) + \#E(T) \\ &= \sum_{k=0}^{2\#V(T)-1} \frac{S_k + S_{k+1}}{2}. \end{aligned}$$

Dans le cas des arbres de Galton-Watson conditionnés par $X = n$, on obtient, en utilisant le théorème d'Aldous rappelé ci-dessus :

$$\frac{A}{n^{3/2}} \xrightarrow{loi} \frac{4}{\tilde{\sigma}} A_e \quad \text{et} \quad \frac{LCT}{n^{3/2}} \xrightarrow{loi} \frac{2}{\tilde{\sigma}} A_e$$

où A_e est l'aire sous l'excursion brownienne normalisée (voir début des annexes),

$$A_e = \int_0^1 e_s ds.$$

On pourra également trouver des utilisations du parcours en profondeur dans les articles de Le Gall & Le Jan [32], Le Gall [31], Takàcs [47].

2.5 Largeur de bande (bandwidth)

La largeur de bande d'un arbre T est définie comme suit :

On note $E(T)$ (resp. $V(T)$) l'ensemble des arêtes (resp. des sommets) de l'arbre T . On suppose que $\#V(T) = n$ et on considère l'ensemble Q_n de toutes les bijections entre les ensembles

$\{1, \dots, n\}$ et $V(T)$, c'est à dire, l'ensemble de toutes les façons de numéroter de 1 à n les sommets de l'arbre. Pour une fonction f de Q_n , on peut considérer la valeur

$$\beta(T, f) = \max_{(u,v) \in E(T)} |f(u) - f(v)|$$

la plus grande différence entre deux valeurs affectées à deux sommets adjacents. La largeur de bande de T , $\beta(T)$ est définie par :

$$\beta(T) = \min_{f \in Q_n} \beta(T, f) = \min_{f \in Q_n} \max_{(u,v) \in E(T)} |f(u) - f(v)|$$

cette quantité apparaît dans des problèmes d'optimisation, de programmation dynamique. Notons $H(T)$ (resp. $W(T)$) la hauteur (resp. largeur) de l'arbre T ; on a :

$$\frac{n-1}{2H(T)} \leq \beta(T) \leq 2W(T).$$

Montrons d'abord la première inégalité. Prenons une fonction f pour laquelle la valeur $\beta(T)$ est réalisée et considérons les deux noeuds u et v tels que $f(u) = 1$ et $f(v) = n$. Ces deux sommets sont à une *distance* inférieure (ou égale) à $2H(T)$. Notons $a_0 = u, \dots, a_p = v$ les noeuds parcourus par ordre chronologique sur le chemin allant de u à v (sur T) ; on a donc $p \leq 2H(T)$. On a

$$\sum_{k=0}^{p-1} f(a_{k+1}) - f(a_k) = f(v) - f(u) = n - 1.$$

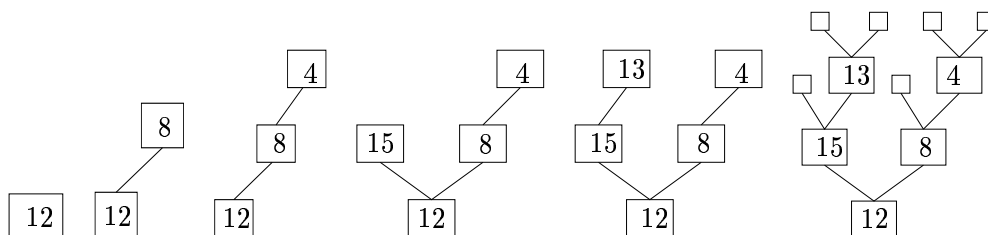
Cette somme comporte au plus $2H(T)$ termes et vaut $n - 1$; l'un de ses termes vaut donc au moins $(n - 1)/2H(T)$.

Pour la deuxième inégalité, on considère une fonction f qui numérote l'arbre de la racine vers son sommet, génération par génération (la racine prend le numéro 1, ses enfants 2, 3 ...). Les nombres affectés à un niveau de l'arbre appartiennent ainsi à un intervalle de longueur au plus $W(T)$; les nombres affectés à deux niveaux consécutifs appartiennent à un intervalle de longueur au plus $2W(T)$. Cette façon de numéroter garantit donc que $\beta(T, f) \leq 2W(T)$.

2.6 Arbres binaires de recherche

On trouve dans la littérature une classe d'arbres binaires appelés arbres binaires de recherche (ABR) utilisés en informatique par l'algorithme de tri rapide (quick sort algorithm) pour stocker des items de façon efficace. L'algorithme procède récursivement en insérant une séquence de nombres x_1, x_2, x_3, \dots comme étiquettes des noeuds d'un arbre binaire complet infini de la façon suivante : le nombre x_j que l'on insère est comparé à l'étiquette de la racine. S'il n'y a pas encore d'étiquette, on donne à la racine l'étiquette x_j . Si x_j est plus grand (resp. plus petit) que l'étiquette de la racine, on recommence cette procédure dans le sous arbre de gauche (resp. de droite).

En classant ainsi n nombres différents et en ne considérant que les noeuds étiquetés, on obtient un arbre unaire-binaire \tilde{T}_n . L'ABR est défini comme étant l'arbre binaire dont les

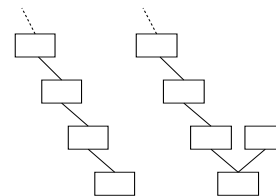


noeuds internes sont les noeuds de \tilde{T}_n . Par exemple, voici sur le dessin ci-dessus, comment est construit l'arbre associé aux nombres 12, 8, 4, 15, 13.

Il est clair que la forme de l'arbre ne dépend pas de la valeur des étiquettes, mais seulement de l'ordre des x_i . Supposons que les nombres x_i soient des v.a. uniformes indépendantes sur $[0, 1]$. Il est facile de voir que la permutation σ de S_n (le groupe symétrique), telle que $x_{\sigma(1)} < x_{\sigma(2)} < \dots < x_{\sigma(n)}$ suit également la loi uniforme sur S_n . L'ensemble des arbres binaires de recherche est classiquement muni de la loi induite par l'équiprobabilité sur S_n .

Muni de cette loi, les ABR ne sont pas des arbres de G.W..

Pour s'en convaincre, il suffit de se rappeler que ci-dessus on a prouvé que la loi induite sur les arbres de G.W. binaires par le conditionnement $X = n$ était la loi uniforme sur tous les arbres. Ce n'est pas le cas pour les ABR. En effet, considérons l'ABR dégénéré à gauche. Il correspond à la permutation unique $\sigma(1) = 1, \sigma(2) = 2, \dots, \sigma(n) = n$. (Voir 1^{er} dessin). Considérons maintenant l'arbre où seul la racine à un fils à droite.



L'étiquette de ce "fils" est inférieure à toutes les autres et il y a $n - 1$ façons de construire cet arbre; il suffit que tous les autres étiquettes soient rangées par ordre croissant.

Les ABR sont en fait très différents des arbres de G.W.. Par exemple, la hauteur d'un ABR est de l'ordre de $\log n$ contre \sqrt{n} pour les arbres de G.W.. On pourra se référer aux livres de Knuth [30], Sedgewick et Flajolet [43] ainsi qu'aux articles de L. Devroye, par exemple [17] pour plus d'éléments sur ces objets.

2.7 Arbres et combinatoire

En combinatoire, on définit les familles d'arbres *simples* non pas par une loi de reproduction mais par une équation fonctionnelle rendant compte des arbres possibles et des *poids* qu'on leur associe.

Un arbre binaire est par définition (combinatoire) ou un noeud externe (une feuille), ou un noeud interne connecté à un couple (ordonné) de deux arbres binaires. Ce genre de définition récursive s'applique à tous les types simples d'arbres. Par exemple, un arbre général est une feuille ou un noeud interne connecté à un nombre quelconque d'arbres généraux (ordonnés).

On décide alors de compter chaque arbre t avec un certain poids $\tilde{w}(t)$, compatible avec cette construction. On utilise alors ce poids comme une "probabilité" pour calculer des lois de hauteurs, largeurs, etc...

La remarque suivante, issue de Aldous [1], permet de passer d'une définition *combinatoire* des arbres à une définition *probabiliste* et réciproquement.

Soit $(c_i)_{i \geq 0}$ une suite de nombres positifs, avec $c_0 = 1$. On note

$$\phi(y) = \sum_k c_k y^k$$

la fonction génératrice associée. Supposons que l'on ait attribué à tout arbre t d'une certaine famille \mathcal{F} , un poids $\tilde{w}(t)$. On note y_n le poids total des arbres ayant n noeuds,

$$y_n = \sum_{t \mid |t|=n} \tilde{w}(t),$$

et Y la fonction génératrice associée. Alors, les deux égalités suivantes sont équivalentes :

$$\begin{aligned} Y(x) &= x\phi(Y(x)) \\ \tilde{w}(t) &= \prod_{i \geq 0} c_i^{D_i(t)} \end{aligned}$$

où $D_i(t)$ est le nombre de sommets de t ayant i fils. Une famille \mathcal{F} d'arbres est dite simple si la famille de poids vérifie ces égalités. La probabilité considérée sur les arbres de tailles n est la suivante :

$$\mathbb{P}(T = t \mid X = n) = \frac{\tilde{w}(t)}{y_n} \quad (2.2)$$

de sorte que la probabilité de choisir un arbre t de taille n parmi tous les arbres de taille n est proportionnelle à son poids.

- Pour les familles d'arbres ordonnés, on choisit $c_i = 1$ si le degré sortant i est autorisé, 0 sinon.

- Pour les arbres étiquetés, on prend $c_i = 1/i!$ si le degré sortant i est autorisé, 0 sinon. Avec ces choix des c_i , on obtient l'équiprobabilité sur l'ensemble des arbres de tailles n pour les familles que l'on a considérées plus haut (arbres binaires, généraux, étiquetés).

Le but maintenant est de construire une loi de progéniture $(p_i)_{i \geq 0}$ d'un processus de G.W. $(Z_i)_{i \geq 0}$ de sorte que la loi induite sur les arbres de taille n soit équivalente à (2.2). On pose

$$p_i = \frac{c_i a^i}{\phi(a)}, \text{ pour } i \geq 0,$$

a étant choisi de sorte que $\mathbb{E}(Z_1) = 1$ (comme on l'a déjà dit, a vérifie, $a\phi'(a) = \phi(a)$). En (2.1), on a donné la probabilité d'un arbre de Galton-Watson t , $w(t)$. Elle s'écrit ici

$$\begin{aligned} w(t) &= \prod_{i \geq 0} p_i^{D_i(t)} \\ &= \prod_{i \geq 0} \left(\frac{c_i a^i}{\phi(a)} \right)^{D_i(t)} \end{aligned}$$

Ainsi,

$$w(t) = \tilde{w}(t) \frac{a^{|t|-1}}{\phi^{|t|}(a)}$$

et donc, $w(t)$ et $\tilde{w}(t)$ sont proportionnels et donc définissent les mêmes probabilités et donc le même modèle.

Références

- [1] D.J. Aldous, (1991) *The continuum random tree. II*, Ann. Probab. 19, No.1, 1-28.
- [2] D.J. Aldous, (1993), *The continuum random tree. III*, Ann. Probab., **21**, 248-289.
- [3] D.J. Aldous, (1997) *Brownian excursions, critical random graphs and the multiplicative coalescent.*, Ann. Probab. 25, No.2, 812-854.
- [4] S. Asmussen, H. Hering, (1983) *Branching processes*, Progress in Prob. and Stat., Birkhaeuser, Vol. 3.
- [5] K.B. Athreya, P.E. Ney, (1972) *Branching processes*, Springer-Verlag.
- [6] J. Bertoin, J. Pitman, (1994) *Path transformations connecting Brownian bridge, excursion and meander*, Bull. Sci. Math., II. Ser. 118, No.2, 147-166.
- [7] P. Biane, (1986) *Relations entre pont et excursion du mouvement brownien reel*, Ann. Inst. Henri Poincare, Probab. Stat. 22, 1-7.
- [8] P. Billingsley, (1968) *Convergence of probability measures*, J. Wiley and Sons.
- [9] P. Billingsley, (1979) *Probability and measure*, Wiley Series in Prob. and math. Stat.
- [10] A.N. Borodin, P. Salminen, (1996) *Handbook of Brownian Motion - Facts and Formulae*, Birkäuser.
- [11] L. Breiman, (1968) *Probability*, Addison-Wesley.
- [12] P. Biane, M. Yor, (1987) *Valeurs principales associées aux temps locaux browniens*, Bull. Sci. Maths 111, 23-101.
- [13] K.L. Chung, (1976) *Excursions in Brownian motion*, Ark. för Math., **14**, 155-177.
- [14] T.H. Cormen, E. Leiserson, R.L. Rivest, (1990) *Introduction to Algorithm*, MIT press.
- [15] E. Csáki, S.G. Mohanty, (1986) *Some joint distributions for conditional random walks.*, Can. J. Stat. 14, 19-28.
- [16] M. Drmota, B. Gittenberger, (1997) *On the profile of random trees*, Random Struct. Algorithms 10, No.4, 421-451 .
- [17] L. Devroye, (1987) *Branching processes in the analysis of the heights of trees*. Acta Inf. 24, 277-298.
- [18] R.T. Durrett, D.L. Iglehart, (1977) *Functionals of Brownian meander and Brownian excursion*, Ann. Probab. 5, 130-135.
- [19] M. Dwass, (1962) *A fluctuation theorem for cyclic random variables*, Ann. Math. Stat., **33**, 1450-1453.
- [20] M. Dwass, (1969) *The total progeny in a branching process and a related random walk*, J. Appl. Probab. 6, 682-686.
- [21] P. Flajolet, A.M. Odlyzko, (1982) *The average height of binary trees and other simple trees*, J. Comput. Syst. Sci. 25, 171-213.
- [22] D. Foata, J. Riordan, (1974) *Mappings of acyclic and parking functions*. Aequationes math. Vol. **10** : 10-22.

- [23] J. Françon, (1975) *Acyclic and parking functions*. J. Combinat. Theory, Ser. A Vol. **18** : 27-35.
- [24] Th. Jeulin, (1980) *Semi-martingales et grossissement d'une filtration*. Lecture Notes in Mathematics, 833, Springer-Verlag.
- [25] T.E. Harris, (1963) *The theory of branching processes* , Springer-Verlag.
- [26] I. Karatzas, S.E. Shreve, (1988) *Brownian motion and stochastic calculus*, Springer-Verlag.
- [27] R. Kemp, (1979) *On the stack size of regularly distributed binary trees*, Proceedings, 6th ICALP Conf., Udine.
- [28] J.H.B. Kemperman, (1961) *The passage problem for a stationary Markov chain*, Univ. of Chicago Press.
- [29] D.P. Kennedy, (1975) *The Galton-Watson process conditioned on the total progeny*. J. appl. Probab. 12, 800-806.
- [30] D.E. Knuth, (1968) *The art of computer programming. Vol. 1, 2, 3*, Add-Wes pub. Comp., XXII.
- [31] J.F. Le Gall, (1986) *Une approche élémentaire des théorèmes de décomposition de Williams*, Lect. Notes Math. 1204, Sémin. Prob. XX, 447-464, Springer.
- [32] J.F. Le Gall, Y. Le Jan, (1996) *Branching processes in Lévy processes : The exploration process*, preprint Université de Paris-Sud, Orsay.
- [33] G. Louchard, (1984) *Kac's formula, Levy's local time and Brownian excursion*. J. Appl. Probab. 21, 479-499.
- [34] G. Louchard, (1984) *The Brownian excursion area : A numerical analysis*, Comput. Math. Appl. 10, 413-417.
- [35] G. Louchard, (1986) *Erratum : The Brownian excursion area : a numerical analysis*, Comput. Math. Appl., Part A 12, 375.
- [36] J.W. Moon, (1970) *Counting labelled trees*. Canadian Mathematical Monographs, No. 1, Canadian Mathematical Congress, Montreal.
- [37] R. Otter, (1949) *The multiplicative process*. , Ann. Math. Statist., Baltimore Md. 20, 206-224 (1949).
- [38] D. Pollard, (1984) *Convergence of stochastic processes*, Springer (1984).
- [39] A. Renyi, G. Szekeres, (1967) *On the height of trees*, Austral. J. Math. 7, 497-507.
- [40] P. Révész, (1990) *Random walk in random and non-random environments*, World Scientific xiv.
- [41] D. Revuz, M.Yor, (1990) *Continuous martingales and Brownian motion*, Springer-Verlag, 9-10.
- [42] J. Riordan, (1960) *The enumeration of trees by height and diameter*, IBM J. Res. Dev. **4** , 473-478.
- [43] R. Sedgewick, Ph. Flajolet, (1996) *An introduction to the analysis of algorithms*, Amsterdam : Addison-Wesley, xv.

- [44] M. P. Schutzenberger, (1968) *On an enumeration problem*. J. Combinatorial Theory Vol. **4**, 219–221.
- [45] J. Spencer, (1997). *Enumerating graphs and Brownian motion*. Commun. Pure Appl. Math. Vol. **50** (3), 291-294.
- [46] F. Spitzer, (1964) *Principes of Random Walk* , D. Van Nostrand Compagny, p. 61.
- [47] L. Takács, (1991) *A Bernoulli excursion and its various applications*, Adv. Appl. Prob, 23, 557-585.
- [48] L. Takács, (1992) *Random walk processes and their various applications*, Probability theory and applications, Essays to the Mem. of J. Mogyorodi, Math. Appl. 80, 1-32.
- [49] L. Takács, (1993) *Limit distributions for queues and random rooted trees*, J. Appl. Math. Stoch. Ana., **6** , No.3, 189 - 216
- [50] L. Takács, (1995) *Brownian local times*, J. Appl. Math. Stochastic Anal. 8, No.3, 209-232.
- [51] W. Vervaat, (1979) *A relation between Brownian bridge and Brownian excursion*, Ann. Probab. 7, 143-149.