



HAL
open science

Statistiques d'extrêmes du mouvement brownien et applications

Julien Randon-Furling

► **To cite this version:**

Julien Randon-Furling. Statistiques d'extrêmes du mouvement brownien et applications. Physique mathématique [math-ph]. Université Paris Sud - Paris XI, 2009. Français. NNT : . tel-00524212

HAL Id: tel-00524212

<https://theses.hal.science/tel-00524212>

Submitted on 12 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE de DOCTORAT de L'UNIVERSITÉ PARIS-SUD XI

Spécialité : Physique Théorique

présentée par

Julien RANDON-FURLING

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ PARIS-SUD XI

Sujet :

Statistiques d'extrêmes du mouvement brownien et applications

Soutenance le 13 novembre 2009

devant le jury composé de :

M. Philippe Biane	Examineur
M. Jean-Philippe Bouchaud	Président du jury
M. Alain Comtet	Co-directeur de thèse
M. Bertrand Duplantier	Rapporteur
M. Hendrik-Jan Hilhorst	Rapporteur
M. Thierry Huillet	Examineur
M. Satya N. Majumdar	Directeur de thèse

Table des matières

Introduction	1
I Au fil du mouvement brownien linéaire : extrêma et temps d'atteinte	3
1 Définitions et notations	7
1.1 Définitions	7
1.2 Notations	8
2 Temps d'atteinte du maximum sur une durée fixée	11
2.1 Deux exemples classiques : résolution d'une équation de Fokker-Planck	12
2.2 Traitement par intégrale de chemin et formule de Feynman-Kac : excursion et méandre browniens	20
3 Temps d'atteinte du maximum avant le premier passage	33
3.1 Motivations	34
3.2 Loi jointe et lois marginales	36
3.3 Graphiques, simulations numériques et asymptotes	39
3.4 Cas d'un mouvement brownien dérivant vers l'origine	42
4 Temps d'atteinte du maximum pour n browniens	47
4.1 Cas de n mouvements browniens libres indépendants	48
4.2 Cas de n ponts browniens indépendants	51
4.3 Cas de n excursions et de n méandres browniens indépendants	53
4.4 Mouvements non indépendants	54

II	Autour du mouvement brownien plan : enveloppes convexes aléatoires	57
5	Aperçu de travaux et résultats existants relatifs aux enveloppes convexes aléatoires	61
6	Du maximum unidimensionnel à la fonction support	79
6.1	Présentation du problème	79
6.2	Cas particulier des angles $\theta = \frac{\pi}{n}$	82
6.3	Cas général	85
6.4	Valeur moyenne de M	88
7	Fonction support et formules de Cauchy : une approche globale des enveloppes convexes aléatoires	95
7.1	Notion de fonction support	96
7.2	Les formules de Cauchy	101
7.3	Les formules de Cauchy appliquées à un échantillon aléatoire .	102
8	Points indépendants	105
8.1	Cas général	105
8.2	Cas isotropes	106
8.3	Comportement asymptotique du périmètre et de l'aire moyens	108
8.4	Un cas non isotrope	120
9	Points corrélés	125
9.1	Chemins browniens plans et <i>domaine vital</i>	125
9.2	Enveloppe convexe du mouvement brownien plan	128
9.3	Cas (particulier) d'une excursion brownienne	133
9.4	Enveloppe convexe de n chemins browniens plans	135
9.5	Simulations numériques et discussion	140

Conclusion	149
Annexes	155
A Note historique	155
B Somme faisant intervenir une fonction de Bessel modifiée	161
C Formules de Cauchy	163
D Chemins browniens multiples : analyse asymptotique	165
D.1 Chemins ouverts : périmètre moyen	165
D.2 Chemins ouverts : aire moyenne	167
D.3 Chemins fermés : périmètre moyen	169
D.4 Chemins fermés : aire moyenne	170
E Simulations numériques	173
F Note : enveloppes convexes aléatoires en 3 dimensions	177
F.1 Formule de Cauchy en trois dimensions	177
F.2 Application de la formule de Cauchy « tridimensionnelle » aux enveloppes convexes aléatoires	178
F.3 Résultats	179
G Articles	181
Time of maximum before first passage	181
Time to reach maximum for a variety of Brownian motions	195
Exact distribution of the maximal height of p vicious walkers	213
Convex hull of N planar Brownian motions	217
Bibliographie	237

Remerciements

Il est de coutume, au seuil d'un mémoire de thèse, de remercier ceux qui par leur soutien, leurs soins et leur attention ont rendu possible un travail qui, sans eux, n'aurait pu voir le jour. Si la langue veut que l'on *sacrifie* à une tradition, c'est, en pareille occasion, un grand bonheur : je tiens à exprimer ici à tous, humblement, ma reconnaissance.

Eu égard au lieu de ces remerciements, qu'il me soit permis de ne nommer que ceux qui, ayant inspiré, guidé et encadré mon travail au cours des second et troisième cycles de l'enseignement supérieur, l'ont fait aboutir à cette thèse : M. Kirone Mallick, Mme Nilanjana Datta et bien sûr MM. Alain Comtet et Satya Majumdar. Indécision ou vocation à (se) questionner, j'ai plus souvent qu'à mon tour interrogé les raisons du parcours universitaire qui a été le mien jusqu'ici ; ils ont non seulement su comprendre ces doutes mais ont surtout pris sans compter sur leur temps et sur leur énergie pour m'éclairer.

La modestie de ma position ne m'y autorise peut-être pas, mais je voudrais, avant de refermer ce chapitre, rendre tout particulièrement hommage à l'exceptionnelle qualité de mes directeurs de thèse qui, bien au-delà de ce que cette fonction implique habituellement, ont été et demeureront des maîtres dans le sens le plus noble du terme — alliant humanité et générosité à la plus enthousiasmante excellence intellectuelle et scientifique. Les remercier va de soi ; si je sais, dans la mesure de mes moyens, faire vivre leur enseignement, ce sera un honneur.

À l'École publique, et à ceux qui la défendent.

Introduction

LES statistiques de valeurs extrêmes, qui s'intéressent aux événements rares — ceux dont la probabilité se trouve dans les queues des distributions —, apparaissent naturellement dans des contextes physiques nombreux et variés, en particulier ceux de catastrophes naturelles : crues décennales ou centennales et hauteur des barrages et digues susceptibles de les contenir [1], tempêtes « du siècle » occasionnant d'importants dommages matériels et environnementaux [2], grands incendies [3], tremblements de terre [4, 5], mais aussi records sportifs ou autres [6, 7], risques financiers [8], ou encore résistance d'un matériau fibreux [9, 10].

Développée par les mathématiciens dès la première moitié du vingtième siècle [11–14], l'étude de la distribution du maximum d'un ensemble de variables aléatoires indépendantes est venue donner aux physiciens les outils théoriques adéquats pour traiter ce type de questions [15–17], la théorie pouvant être étendue à des variables faiblement corrélées¹. Notons cependant que, pour des variables fortement corrélées, il n'existe pas, pour l'heure, de théorie générale.

Autre domaine où se croisent physique, mathématiques et phénomènes naturels, le mouvement brownien est peut-être l'un des sujets les plus étudiés des cent dernières années, comme ont pu le mettre en évidence plusieurs des articles [18–22] qui lui ont été consacrés autour du centenaire, en 2005, de l'*annus mirabilis* d'Einstein. C'est en 1905 en effet [23], qu'Albert Einstein proposait une théorie physique rendant compte du mouvement observé par le botaniste britannique Robert Brown sur des particules contenues

¹Pour lesquelles il existe une longueur de corrélation finie.

dans des grains de pollen [24]. Cette théorie venait par ailleurs étayer l'hypothèse atomiste, dont les expériences menées peu après par Jean Perrin [25] confirmèrent la validité². Dès ses origines, le mouvement brownien montrait donc une versatilité que ne devaient pas démentir les cent années suivantes : trouvant régulièrement de nouvelles applications dans des domaines très divers (physique [18–20, 26, 27], astronomie [21, 28, 29], biologie [30–32], informatique [20, 33, 34], finance [8, 35, 36])³, applications générant elles-mêmes en retour de nouvelles questions, la théorie du mouvement brownien n'a cessé de se développer sans pour autant s'épuiser.

Diverses applications ont motivé, en particulier récemment [36–39], des questions portant sur le mouvement brownien linéaire, et notamment sur son maximum et sur le temps d'atteinte de ce maximum. Suscitant l'élaboration de nouvelles méthodes sur le plan théorique, ces questions, que nous aborderons dans la première partie de ce mémoire, contribuent au développement d'une statistique des valeurs extrêmes pour des variables aléatoires fortement corrélées — contexte jusqu'à présent moins étudié que celui de variables indépendantes, et dont le mouvement brownien constitue un exemple simple et fondamental⁴. Nous verrons ensuite dans une seconde partie que ces questions trouvent d'autres applications pratiques et théoriques à travers l'étude de l'enveloppe convexe du mouvement brownien plan, qui peut intervenir, entre autres, dans la modélisation de systèmes écologiques tels que le territoire exploré par un groupe d'animaux.

Les résultats que nous présentons dans ce mémoire constituent des exemples de résultats exacts nouveaux ayant trait aux statistiques d'extrêmes du mouvement brownien — ils se situent ainsi à la croisée de ces deux sujets d'études.

²Nous proposons en annexe une courte note historique (§ A).

³Sur un sujet renvoyant des dizaines de milliers de références dans les bases de données scientifiques en ligne, nous ne prétendons bien sûr pas, à travers les quelques références citées ici à titre d'exemple, refléter autre chose que la simple variété des domaines d'applications et la continuité de l'intérêt porté au mouvement brownien !

⁴Rappelons la fonction de corrélation à deux points d'un mouvement brownien linéaire $x(t) : \langle x(t)x(t') \rangle = 2D \min \{t, t'\}$, où D est la constante de diffusion du mouvement.

Première partie

Au fil du mouvement brownien linéaire : extrêma et temps d'atteinte

CETTE première partie est consacrée au mouvement brownien unidimensionnel, que l'on peut envisager comme le déplacement d'une particule sur une ligne, sa position obéissant dans les cas les plus simples à une équation de Langevin libre (1.1). Nous nous intéressons, dans ce contexte, aux extrêma du déplacement de la particule, notamment son écart maximum au point 0 de la droite dont, dans la plupart des cas, elle est issue. La question de la distribution de ce maximum — la probabilité qu'il soit inférieur ou égal à telle ou telle valeur étant généralement connue — ne nous intéressera pas tant que celle de la distribution du temps d'atteinte de ce maximum sur l'intervalle de temps considéré. En effet, si celle-ci est connue dans les cas les plus simples que constituent le mouvement brownien libre sur une durée T fixée et le pont brownien (qui n'est autre qu'un mouvement brownien contraint à revenir en son point de départ au temps T) [40, 41], il n'en va pas de même, à notre connaissance, pour d'autres types de mouvements browniens contraints tels que l'excursion (cas particulier de pont brownien avec la contrainte supplémentaire de demeurer du côté positif de la droite) ou le méandre (mouvement demeurant positif mais n'étant pas contraint à revenir en son point de départ au temps final T).

Physiquement, ces divers types de mouvement brownien linéaire ont pu être utilisés pour décrire l'interface entre deux milieux [42], le temps T correspondant alors à la longueur de cette interface et les fluctuations du mouvement aux irrégularités de cette surface de séparation. De même, dans les modèles représentant des chaînes polymériques par des mouvements browniens [43] — les contraintes de retour en un point origine correspondant alors à des points de fixation des extrémités de la chaîne, et celles de positivité du mouvement à une impossibilité pour le polymère de pénétrer un substrat sous-jacent — le déplacement maximal du mouvement donne l'extension transversale maximale de la molécule. Dans le premier exemple, la connaissance du temps d'atteinte du maximum (pour conserver la terminologie liée au contexte du mouvement) permet de connaître la position de la plus forte irrégularité (en relief) de l'interface ; dans le second, il permet de connaître la position où l'extension transversale est maximale.

Une première manière de modifier les données du problème consiste à envisager non plus une durée fixée pour le mouvement, mais au contraire un in-

tervalle de temps variable, dont la longueur changerait à chaque réalisation du mouvement. C'est ce que nous ferons ici dans un second temps, en considérant un mouvement brownien entre l'instant de son départ d'un point x_0 strictement positif et le moment auquel il passe en 0 pour la première fois. Ce temps, dit de premier passage, fixera la durée de l'intervalle sur lequel nous rechercherons le déplacement maximum de la particule. Nous verrons que ce type de mouvement peut correspondre à des problèmes concrets tels que l'évolution de la longueur d'une file d'attente ou celle du cours d'une action en bourse.

Une seconde manière d'élargir le champ de notre étude est de considérer simultanément plusieurs particules browniennes. Au lieu de rechercher le déplacement maximal d'une seule particule issue de l'origine et observée sur un temps T fixé, considérons ainsi n particules issues de l'origine, contraintes ou non à y retourner au temps T et demandons-nous comme précédemment quelles sont les distributions du maximum global et de son temps d'atteinte pour l'ensemble des n mouvements. Dans les cas les plus simples, les particules seront indépendantes, elles ne se « verront » pas et pourront se croiser — nous verrons dans la seconde partie de ce mémoire que de tels exemples peuvent correspondre à des systèmes bidimensionnels d'intérêt physique, vus en projection unidimensionnelle. En complexifiant un peu le problème, il est possible de retrouver des modèles d'interfaces : si l'on impose aux n mouvements considérés de ne pas se croiser, ils peuvent être vus comme des parois séparant autant de domaines, et forment des configurations dites de « marcheurs malveillants » [42] que nous évoquerons à la fin de cette première partie.

Chapitre 1

Définitions et notations

1.1 Définitions



FIG. 1.1 – Mouvement brownien standard libre sur l'intervalle $[0, T]$

Nous entendons par mouvement brownien unidimensionnel (que nous qualifierons parfois de simple, de standard et / ou de libre) un processus aléatoire réel $x(t)$ satisfaisant l'**équation de Langevin** suivante :

$$\boxed{\frac{dx}{dt} = \eta(t)} \quad (1.1)$$

où η est un bruit blanc, c'est à dire un processus stochastique gaussien de

moyenne nulle et de corrélation *delta* :

$$\langle \eta(t) \rangle = 0 \quad (1.2)$$

$$\langle \eta(t)\eta(t') \rangle = 2D\delta(t-t') \quad (1.3)$$

δ étant la fonction delta de Dirac et D une constante de diffusion que nous fixons égale à $\frac{1}{2}$ dans toute la suite.

Nous pouvons calculer le **propagateur du mouvement brownien libre**, c'est-à-dire la probabilité que, étant en x_0 au temps t_0 , la particule soit en x au temps t :

$$\text{Prob}(x, t|x_0, t_0) \equiv G_0(x, t) \quad (1.4)$$

Ce propagateur satisfait l'équation suivante (que l'on appellera « **équation de Fokker-Planck libre** ») :

$$\boxed{\frac{\partial G_0}{\partial t} = \frac{1}{2} \frac{\partial^2 G_0}{\partial x^2}}, \quad (1.5)$$

et s'écrit :

$$\boxed{G_0(x, t) = \frac{e^{-\frac{(x-x_0)^2}{2(t-t_0)}}}{\sqrt{2\pi(t-t_0)}}} \quad (1.6)$$

1.2 Notations

Autant que possible, nous noterons toujours par la suite $x(t)$ ou $y(t)$ (parfois z ou h dans la seconde partie) un mouvement brownien unidimensionnel, satisfaisant éventuellement un certain nombre de contraintes que nous nous efforcerons de bien préciser. De même, G désignera de manière général un propagateur. Les minuscules de la fin de l'alphabet grec (exceptée τ , correspondant, comme t , au temps), par exemple σ ou ρ , seront principalement réservées à des densités de probabilité, tandis que F sera le choix préférentiel pour les fonctions de répartition. D'autres majuscules latines telles que P ou S seront également employées pour désigner des probabilités. Nous réserverons \mathcal{B} à des mouvements browniens plans.

Nous utiliserons le symbole \equiv pour indiquer une égalité correspondant à une définition, et les crochets $\langle \dots \rangle$ pour désigner une valeur moyenne (sauf bien entendu dans le contexte de la notation *bra-ket*); il nous arrivera, en particulier lorsque nous rendrons compte de résultats obtenus par des mathématiciens, de parler d'espérance plutôt que de moyenne et d'utiliser la notation \mathbb{E} .

Les résultats importants, *per se* ou pour la suite du texte, seront en règle générale encadrés, un cadre double signalant en particulier un résultat nouveau (à notre connaissance). Enfin, par abus de langage, nous réserverons l'expression de *marche aléatoire* aux marches aléatoires en temps discret (pouvant être vues comme une succession dénombrable de pas), n'employant jamais celle-ci pour désigner un mouvement brownien, bien qu'il nous arrivera de parler de marcheur brownien, en particulier dans les contextes de mouvements sans intersection.

Chapitre 2

Temps d'atteinte du maximum sur une durée fixée

Nous abordons ici la question du calcul de la distribution du temps d'atteinte du maximum d'un mouvement brownien unidimensionnel, libre ou satisfaisant un certain nombre de contraintes relatives à sa position à un ou des instants donnés. Nous nous proposons d'utiliser une méthode exposée dans [37] qui repose sur l'introduction d'un paramètre « de coupure » ε appelé à tendre vers 0 à l'issue du calcul du poids probabiliste d'un chemin brownien quelconque. Ce paramètre de coupure permet de juguler le problème posé par une propriété caractéristique du mouvement brownien : celle de repasser un nombre infini de fois en un point qu'il vient de quitter. Nous présenterons ici d'abord un calcul classique du temps d'atteinte du maximum pour un mouvement brownien standard issu de 0 et observé sur un intervalle de temps de longueur T fixée. Nous utiliserons ensuite l'approche par intégrales de chemins développée par Comtet et Majumdar [37] pour étudier d'autres types de mouvements browniens unidimensionnels, toujours sur un intervalle de durée T fixée.

2.1 Deux exemples classiques traités par la résolution d'une équation de Fokker-Planck

2.1.1 Mouvement brownien libre



FIG. 2.1 – Maximum M et temps t_m d'atteinte de ce maximum pour un mouvement brownien libre sur l'intervalle $[0, T]$

Soit $x(t)$ un mouvement brownien sur $[0, T]$, issu de 0 (c'est-à-dire $x(0) = 0$). Notons :

$$M = \max_{0 \leq t \leq T} \{x(t)\}$$

et

$$t_m = \inf \{t \in [0, T] \text{ tel que } x(t) = M\}.$$

Posons

$$y(t) = M - x(t).$$

y satisfait l'équation (1.1) ci-dessus et est donc lui-même un mouvement brownien, mais avec les particularités suivantes : il s'annule en t_m et il est strictement positif sur le reste de l'intervalle $[0, T]$ (*cf* fig. 2.2).

Bien qu'il soit possible de procéder autrement, nous introduisons ici un paramètre de coupure ε en imposant non pas $y(t_m) = 0$ mais $y(t_m) = \varepsilon^1$.

¹Nous pourrions également poser $y = M + \varepsilon - x$.

L'idée d'introduire un paramètre de coupure vient du fait qu'il peut s'avérer délicat d'imposer à un mouvement brownien de demeurer strictement positif juste avant ou juste après un passage en 0 (nous reviendrons sur ce point au paragraphe § 2.2.1). Une manière de contourner ce problème dans le contexte présent est d'utiliser des statistiques de temps de premier passage, une autre est d'introduire une coupure, comme nous le faisons ici : cette coupure permet d'effectuer les calculs et « disparaît » ensuite par normalisation lorsque l'on prend la limite $\varepsilon \rightarrow 0$.

Pour calculer la probabilité d'un chemin du mouvement y , nous cherchons d'abord à calculer le propagateur G_+ d'un mouvement brownien libre sur \mathbb{R}^+ , qui diffère de G_0 , le propagateur libre sur \mathbb{R} tout entier donné par l'éq. (1.6). On considère donc l'équation de Fokker-Planck libre sur \mathbb{R}^+ , avec une condition d'absorption en 0 :

$$\frac{\partial G_+}{\partial t} = \frac{1}{2} \frac{\partial^2 G_+}{\partial x^2}, \quad (2.1)$$

avec

$$G_+(0, t) = 0 \quad (2.2)$$

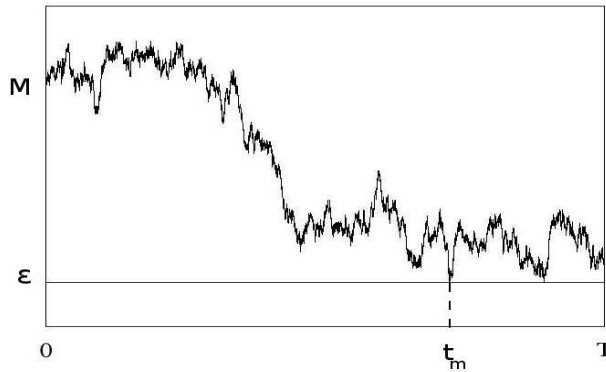


FIG. 2.2 – Mouvement brownien obtenu par soustraction de $x(t)$ à son maximum : $y(t) = M - x(t)$, avec un paramètre de coupure ε .

Une méthode classique pour calculer le propagateur G_+ est la méthode dite « des images », qui consiste, à partir du propagateur libre G_0 (éq. 1.6), à

retrancher au poids des trajectoires issues de M celui des trajectoires issues de $-M$:

$$G_+(x, t) = \frac{1}{\sqrt{2\pi t}} \left[e^{-\frac{(x-M)^2}{2t}} - e^{-\frac{(x+M)^2}{2t}} \right] \quad (2.3)$$

Notons que :

$$\int_0^\infty G_+(x, t) dx \leq 1, \quad (2.4)$$

l'inégalité s'expliquant par le fait qu'il y a absorption et donc « fuite de probabilité » en 0.

Désignons par $S_M(t)$ la probabilité que le mouvement y de propagateur G_+ ait survécu jusqu'au temps t , autrement dit qu'il n'ait pas visité 0 pendant $[0, t]$:

$$S_M(t) = \int_0^\infty G_+(y, t) dy \quad (2.5)$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi t}} \left[e^{-\frac{(y-M)^2}{2t}} - e^{-\frac{(y+M)^2}{2t}} \right] dy \quad (2.6)$$

$$= \operatorname{erf} \left(\frac{M}{\sqrt{2t}} \right) \quad (2.7)$$

où nous notons erf la fonction erreur :

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du \quad (2.8)$$

Remarquons que l'on peut écrire² :

$$S_M(t) = \operatorname{Prob}(\forall \tau \in [0, t], x(\tau) \leq M),$$

c'est à dire que la probabilité de survie du mouvement y jusqu'au temps t est égale à la probabilité que le maximum du mouvement x sur $[0, t]$ soit inférieur ou égal à M , puisque par construction $y(\tau) = M - x(\tau)$. De cette

²Signalons ici que $S_M(T)$ permet également d'accéder au temps de premier passage en 0 du mouvement y . En effet, si l'on note σ sa densité de probabilité, on peut écrire : $S_M(t) = \int_t^\infty \sigma(\tau) d\tau$. D'où il vient :

$$\sigma(t) = \frac{M}{\sqrt{2\pi t^{\frac{3}{2}}}} e^{-\frac{M^2}{2t}}$$

observation nous pouvons tirer la **densité de probabilité du maximum de $x(t)$ sur $[0, T]$** , en dérivant $S_M(T)$ par rapport à M :

$$\boxed{\sigma_1(M) = \sqrt{\frac{2}{\pi T}} e^{-\frac{M^2}{2T}}}. \quad (2.9)$$

Pour obtenir la loi jointe $\rho(M, t_m)$ de M et de t_m pour le mouvement x , nous utilisons le caractère markovien du mouvement brownien : le poids de ce chemin peut être calculé en multipliant le poids de la partie du chemin à gauche de t_m et celui de la partie à droite. Ceci est vrai également pour le mouvement y dont les trajectoires sont en bijection avec celles de x . Pour le mouvement y , le poids probabiliste de la partie du chemin à gauche de t_m est donné directement par $G_+(\varepsilon, t_m)$, le propagateur pris entre M en $t = 0$ et ε en $t = t_m$:

$$G_+(\varepsilon, t_m) = \frac{1}{\sqrt{2\pi t_m}} \left[e^{-\frac{(\varepsilon-M)^2}{2t_m}} - e^{-\frac{(\varepsilon+M)^2}{2t_m}} \right]. \quad (2.10)$$

En ce qui concerne la seconde partie du mouvement, nous pouvons exprimer son poids directement à partir de la probabilité de survie $S_\bullet(t)$ calculée ci-dessus. En effet, en remplaçant, dans (2.7), M par ε , nous accédons directement à la probabilité qu'un mouvement issu de ε survive jusqu'à un temps t . Or, la seconde partie d'un chemin du mouvement y n'est autre qu'un mouvement issu de ε survivant jusqu'à la fin de l'intervalle, c'est-à-dire pendant un temps $T - t_m$. Son poids est donc donné par :

$$S_\varepsilon(T - t_m) = \operatorname{erf} \left(\frac{\varepsilon}{\sqrt{2(T - t_m)}} \right) \quad (2.11)$$

Il nous suffit ensuite de combiner (2.10) et (2.11) en prenant la limite $\varepsilon \rightarrow 0$:

$$\begin{aligned} \rho(M, t_m) &= \mathcal{N} \sqrt{\frac{2}{\pi}} \frac{\varepsilon M e^{-\frac{M^2}{2t_m}}}{t_m^{\frac{3}{2}}} \cdot \varepsilon \sqrt{\frac{2}{\pi(T - t_m)}} \\ &= \mathcal{N} 2\varepsilon^2 \frac{M}{\pi t_m^{\frac{3}{2}} \sqrt{T - t_m}} e^{-\frac{M^2}{2t_m}}, \end{aligned} \quad (2.12)$$

\mathcal{N} étant une constante qui assure la normalisation de la distribution :

$$\iint \rho(M, t_m) dM dt_m = 1$$

Un rapide calcul permet de vérifier que $\mathcal{N} = \frac{1}{2\varepsilon^2}$, et donc **pour un mouvement brownien libre** :

$$\boxed{\rho(M, t_m) = \frac{M}{\pi t_m^{\frac{3}{2}} \sqrt{T - t_m}} e^{-\frac{M^2}{2t_m}}} \quad (2.13)$$

Nous pouvons bien sûr retrouver à partir de l'expression de la loi jointe, la loi marginale du maximum (éq. (2.9)) :

$$\sigma_1(M) = \sqrt{\frac{2}{\pi T}} e^{-\frac{M^2}{2T}}, \quad (2.14)$$

et sa fonction de répartition (éq. (2.7)) :

$$F_1(M) = \operatorname{erf}\left(\frac{M}{\sqrt{2T}}\right). \quad (2.15)$$

Nous pouvons également calculer la **loi marginale du temps d'atteinte du maximum** :

$$\boxed{\sigma_2(t_m) = \frac{1}{\pi \sqrt{t_m(T - t_m)}}}, \quad (2.16)$$

et sa fonction de répartition :

$$\boxed{F_2(t_m) = \frac{2}{\pi} \arcsin\left(\frac{t_m}{T}\right)} \quad (2.17)$$

La forme de cette dernière expression fait que l'on se réfère généralement à ρ_2 et/ou F_2 sous le vocable de « loi de l'arcsinus » — et parfois également sous celui de « loi de Lévy » car elle fut établie pour la première fois par le mathématicien français Paul Lévy [40]. Comme on peut le voir sur la représentation graphique (fig. 2.3) de ρ_2 , une des caractéristiques de cette distribution est que sa valeur moyenne, atteinte en $T/2$, non seulement ne correspond pas à un maximum, mais est même un minimum global de la fonction sur l'intervalle $[0, T]$. On pourrait dire que « le moyen est le moins

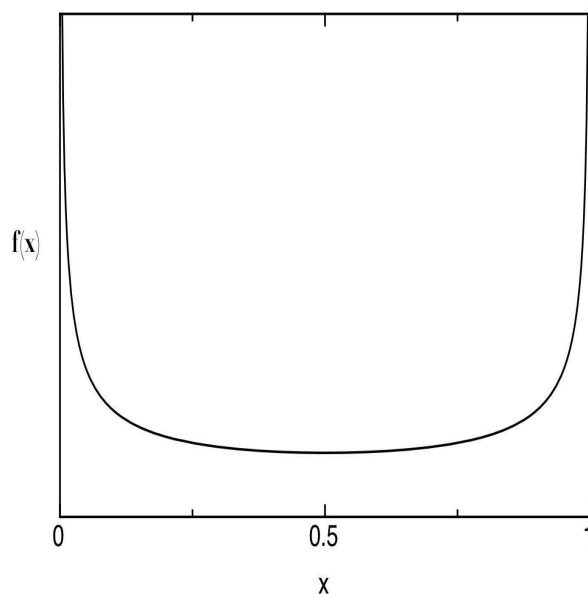


FIG. 2.3 – La loi de Lévy, dite de l’arcsinus : distribution $\sigma_2(t_m) = \frac{1}{T} f\left(\frac{t_m}{T}\right)$ du temps d’atteinte de son maximum par un mouvement brownien standard. (éq. (2.16))

probable », ou que « le typique n’est pas le moyen » — en d’autres termes, une trajectoire brownienne typique atteindra son maximum plutôt vers le début ou plutôt vers la fin de l’intervalle considéré, ce qui revient à dire que, typiquement, le mouvement brownien « s’éloigne » dans une direction (vers le bas dans le premier cas, vers le haut dans le second) plutôt qu’il ne languit autour de son point de départ, comme on peut d’ailleurs le voir sur la figure 1.1.

2.1.2 Pont brownien

Un deuxième exemple classique que l’on peut traiter de façon immédiate dans le prolongement de l’exemple précédent est celui du pont brownien. Un pont brownien est un mouvement brownien conditionné à revenir à son point de départ au bout de l’intervalle de temps T fixé (*cf.* figure 2.4). Si l’on désigne par x_t un mouvement brownien libre issu de 0 sur $[0, T]$, on peut

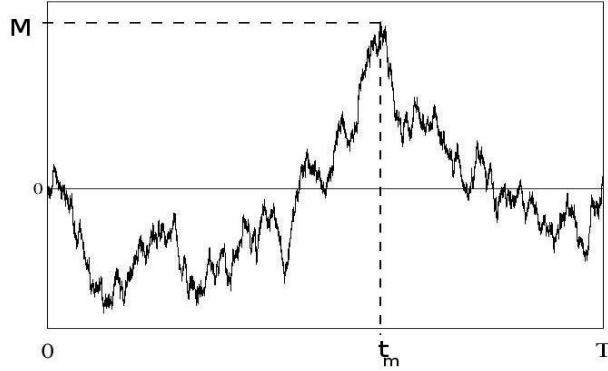


FIG. 2.4 – Pont brownien sur $[0, T]$ avec son maximum et le temps d'atteinte de ce maximum

obtenir à partir de x_l un pont brownien x_p sur $[0, T]$ de la manière suivante :

$$x_p(t) = x_l(t) - \frac{t}{T} x_l(T).$$

Nous reprenons les notations du premier exemple, et notons $x(t)$ un pont brownien sur $[0, T]$, M son maximum et y le processus défini sur $[0, T]$ par $y(t) = M - x(t)$ avec un paramètre de coupure $\varepsilon = y(t_m)$ (cf. fig. 2.5). Nous avons :

$$x(0) = x(T) = 0$$

et :

$$y(0) = y(T) = M \quad \text{et} \quad y(t_m) = \varepsilon$$

Le raisonnement sera tout à fait similaire à celui de l'exemple précédent, à ceci près que le point d'arrivée des deux mouvements étant fixé (en 0 pour x , en M pour y), les deux parties d'un chemin de y , avant et après t_m , seront données par le propagateur de y (cf. éq.(2.3))

$$\begin{aligned} \rho(M, t_m) &\propto G_+(\varepsilon, t_m) \cdot G_+(\varepsilon, T - t_m) \\ &\simeq \mathcal{N} \sqrt{\frac{2}{\pi}} \frac{\varepsilon M e^{-\frac{M^2}{2t_m}}}{t_m^{\frac{3}{2}}} \cdot \sqrt{\frac{2}{\pi}} \frac{\varepsilon M e^{-\frac{M^2}{2(T-t_m)}}}{(T-t_m)^{\frac{3}{2}}} \\ &\simeq \mathcal{N} \frac{2\varepsilon^2 M^2 e^{-\frac{M^2}{2} \frac{T}{t_m(T-t_m)}}}{\pi [t_m(T-t_m)]^{\frac{3}{2}}} \end{aligned} \quad (2.18)$$

La condition de normalisation donne cette fois $\mathcal{N} = \sqrt{\frac{\pi T}{2\varepsilon^4}}$, et l'on obtient donc finalement **pour un pont brownien** :

$$\rho(M, t_m) = \sqrt{\frac{2T}{\pi}} \frac{M^2}{[t_m(T - t_m)]^{\frac{3}{2}}} e^{-\frac{M^2}{2} \frac{T}{t_m(T-t_m)}} \quad (2.19)$$



FIG. 2.5 – Mouvement brownien obtenu en retranchant un pont brownien $x(t)$ à son maximum : $y(t) = M - x(t)$, avec un paramètre de coupure ε .

Comme précédemment, on peut calculer ensuite les lois marginales et les fonctions de répartition correspondantes. Pour le maximum M , on obtient :

$$\sigma_1(M) = \frac{4M}{T} e^{-\frac{2M^2}{T}}, \quad (2.20)$$

et sa fonction de répartition :

$$F_1(M) = 1 - e^{-\frac{2M^2}{T}}. \quad (2.21)$$

Et pour le **temps t_m d'atteinte du maximum d'un pont brownien sur $[0, T]$** :

$$\sigma_2(t_m) = \frac{1}{T}, \quad (2.22)$$

et sa fonction de répartition :

$$F_2(t_m) = \frac{t_m}{T}. \quad (2.23)$$

Remarquons en particulier le caractère uniforme sur l'intervalle $[0, T]$ de la loi du temps d'atteinte du maximum : la symétrie du pont brownien fait que la survenue du maximum est « équiprobable » sur tout l'intervalle de temps considéré.

2.2 Traitement par intégrale de chemin et formule de Feynman-Kac : exemples de l'excursion et du méandre

Nous reprenons ici une méthode exposée par A. Comtet et S. Majumdar [20, 37] dont l'idée centrale est que la mesure de probabilité $P[x(t)]$ associée à un mouvement brownien libre $x(t)$ sur l'intervalle $[0, T]$ peut s'écrire :

$$P[\{x(t')\}] \propto \exp \left[-\frac{1}{2} \int_0^T \left(\frac{dx}{dt} \right)^2 dt' \right] \quad (2.24)$$

La formule de Feynman-Kac permet de traduire ce formalisme d'intégrale de chemin en un formalisme équivalent de propagateur hamiltonien en temps imaginaire :

$$P[\{x(t')\}] = \langle x(0) | e^{-\hat{H}T} | x(T) \rangle, \quad (2.25)$$

où \hat{H} est le hamiltonien d'une particule libre :

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2}$$

L'avantage de ce formalisme est que les contraintes que nous imposerons au mouvement s'y intègrent assez naturellement, en ajoutant un potentiel au hamiltonien, comme nous allons le voir dans les deux exemples suivants.

2.2.1 Excursion brownienne

On appelle excursion brownienne un mouvement brownien $x(t)$ qui, comme le pont, revient à son point de départ au bout d'un temps T fixé, mais qui en outre demeure positif entre 0 et T , comme illustré sur la figure 2.6. W. Verwaat [44] a établi une élégante relation entre pont brownien et excursion brownienne, permettant de construire celle-ci à partir de celui-là par des transformations simples, comme nous l'illustrons sur la figure 2.7.

Nous notons comme précédemment M le maximum atteint par x sur l'intervalle $[0, T]$ et t_m le temps d'atteinte de ce maximum. Nous souhaitons à nouveau calculer la loi jointe $\rho(M, t_m)$

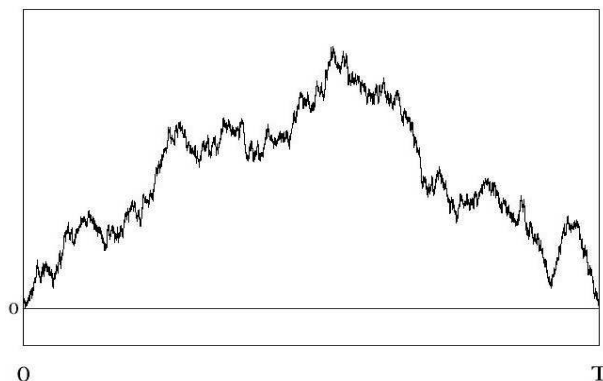


FIG. 2.6 – Excursion brownienne

Revenons d'abord sur l'utilisation d'un paramètre de coupure. Un mouvement brownien quittant 0 à l'instant $t = 0$ franchira l'origine un nombre infini de fois sur l'intervalle $[0, \delta]$ pour tout $\delta > 0$. La contrainte $x(0) = 0$ et $x(t) > 0$ pour tout $t > 0$ résultera dès lors en une probabilité nulle de tels chemins dans une approche « naïve » du problème. Il en va de même pour une contrainte imposant au mouvement de ne pas dépasser une certaine valeur M (en l'occurrence son maximum). Pour pallier cette difficulté, nous considérons (*cf.* fig. 2.8) un mouvement issu de $x_0 > 0$, retournant en x_0 au temps T et contraint d'une part à demeurer à l'intérieur de l'intervalle $[0, M]$ et d'autre part à passer en $M - \varepsilon$ au temps t_m . L'idée est de calculer la mesure de probabilité d'un tel chemin avant de faire tendre x_0 et ε vers 0 pour obtenir la loi correspondant à une véritable excursion brownienne.

Le caractère markovien du mouvement brownien nous permet de considérer séparément les portions $[0, t_m]$ et $[t_m, T]$, puis de multiplier entre elles les mesures correspondantes. Il s'agit donc dans un premier temps de calculer le propagateur :

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle, \quad (2.26)$$

où :

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x),$$

V représentant un puit de potentiel avec des barrières infinies en 0 et en M — traduction dans le formalisme hamiltonien des contraintes d'absorption

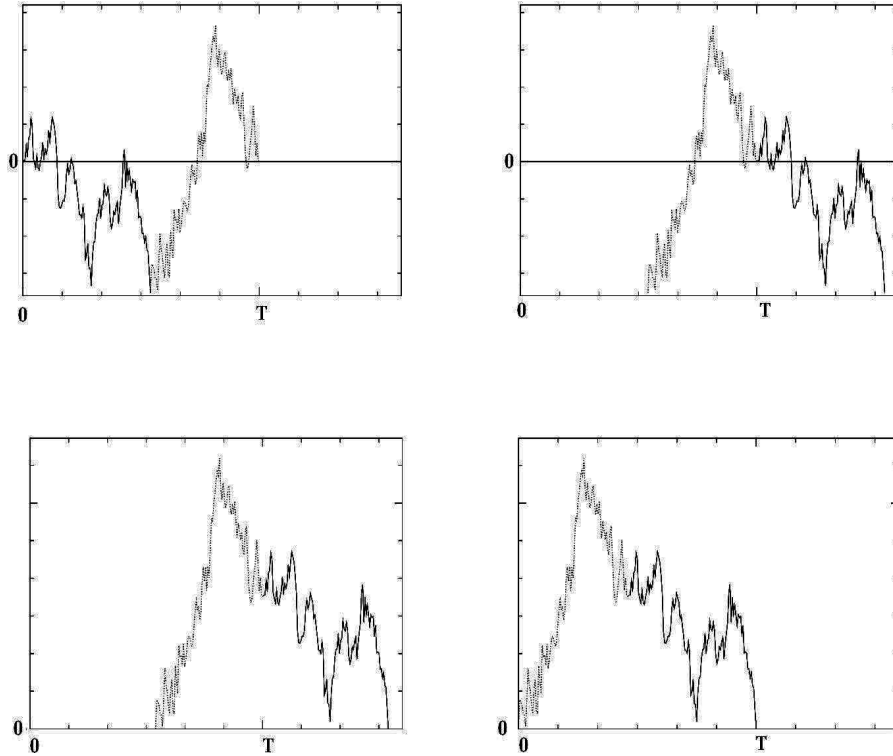


FIG. 2.7 – La construction de Vervaat : du pont brownien à l’excursion brownienne. La trajectoire du pont est séparée au niveau de son minimum, puis les parties à gauche et à droite sont permutées ; la trajectoire ainsi obtenue est continue, grâce au fait que le pont revient en son point de départ. Deux translations permettent ensuite d’obtenir une trajectoire d’excursion.

imposées au mouvement considéré. Les fonctions propres, d’indices $n \in \mathbb{N}$, de ce hamiltonien sont connues et données par :

$$\psi_n(x) = \sqrt{\frac{2}{M}} \sin\left(\frac{n\pi x}{M}\right), \quad (2.27)$$

avec les valeurs propres correspondantes :

$$E_n = \frac{n^2\pi^2}{2M^2}. \quad (2.28)$$

Il suffit alors de décomposer le propagateur (2.26) suivant cette base de vec-



FIG. 2.8 – Mouvement brownien contraint à passer en $M - \varepsilon$ au temps t_m et à retourner à son origine x_0 sans quitter l'intervalle $[0, M]$

teurs propres :

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle = \sum_n \langle x_0 | e^{-\hat{H}t_m} | \psi_n \rangle \langle \psi_n | M - \varepsilon \rangle. \quad (2.29)$$

Nous obtenons :

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle = \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \varepsilon)}{M}\right) e^{-\frac{n^2\pi^2}{2M^2}t_m}. \quad (2.30)$$

La mesure de la partie droite du chemin correspond quant à elle au propagateur suivant :

$$\langle M - \varepsilon | e^{-\hat{H}(T-t_m)} | x_0 \rangle. \quad (2.31)$$

Par analogie avec le calcul précédent, on peut écrire directement :

$$\langle M - \varepsilon | e^{-\hat{H}(T-t_m)} | x_0 \rangle = \frac{2}{M} \sum_{k=1}^{\infty} \sin\left(\frac{k\pi x_0}{M}\right) \sin\left(\frac{k\pi(M - \varepsilon)}{M}\right) e^{-\frac{k^2\pi^2}{2M^2}(T-t_m)}. \quad (2.32)$$

En combinant ensuite les équations (2.30) et (2.32) et en prenant les limites $\varepsilon, x_0 \ll 1$, nous obtenons alors :

$$\rho(M, t_m) = \mathcal{N} \frac{4\pi^4 \varepsilon^2 x_0^2}{M^6} \sum_{n,k=1}^{\infty} (-1)^{n+k} n^2 k^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}(T-t_m)} + o(\varepsilon^2 x_0^2), \quad (2.33)$$

\mathcal{N} désignant comme précédemment une constante de normalisation visant à assurer que :

$$\int_0^\infty \int_0^T \rho(M, t_m) dM dt_m = 1.$$

Le calcul de \mathcal{N} fait disparaître ε et x_0 et nous conduit finalement à l'expression suivante de la **loi jointe du maximum et de son temps d'atteinte pour une excursion brownienne sur l'intervalle $[0, T]$** :

$$\rho(M, t_m) = \sqrt{2} \frac{\pi^{\frac{9}{2}} T^{\frac{3}{2}}}{M^6} \sum_{n,k=1}^{\infty} (-1)^{n+k} n^2 k^2 e^{-\frac{n^2 \pi^2}{2M^2} t_m - \frac{k^2 \pi^2}{2M^2} (T-t_m)} \quad (2.34)$$

Nous pouvons ensuite par intégration sur l'une ou l'autre des variables obtenir les distributions marginales de cette loi jointe. L'intégration sur t_m est la plus simple, elle donne :

$$\sigma_1(M) = \frac{2^{\frac{3}{2}} \pi^{\frac{5}{2}} T^{\frac{3}{2}}}{M^4} \sum_{n,k=1}^{\infty} (-1)^{n+k} \frac{n^2 k^2}{k^2 - n^2} \left[e^{\frac{n^2 \pi^2}{2M^2} T} - e^{\frac{k^2 \pi^2}{2M^2} T} \right]. \quad (2.35)$$

Il existe dans la littérature mathématique une expression pour $\rho_1(M)$ [45–47], dont on peut prouver qu'elle est bien identique à la nôtre, même si ceci n'a rien d'évident à première vue. Il est particulièrement intéressant de comparer les formules que donne chacune de ces deux expressions pour les moments de M : il en résulte en particulier une identité remarquable faisant intervenir la constante d'Apéry, $\zeta(3)$ (ζ étant la fonction de Riemann) dont il ne nous semble pas immédiat d'obtenir une preuve directe :

$$\lim_{\alpha \rightarrow -1} \sum_{n,k=1}^{\infty} \alpha^{n+k} \frac{k^2 n^2}{k+n} \left(\frac{\ln k - \ln n}{k-n} \right) = \frac{3\zeta(3)}{8\pi^2}. \quad (2.36)$$

Nous ne développons pas ce point plus avant ici, mais reproduisons en annexe (page 195) l'article détaillé consacré aux résultats que nous reprenons ici [48].

L'intégration de (2.34) sur M conduit quant à elle à la distribution du **temps d'atteinte de son maximum par une excursion brownienne** :

$$\sigma_2(t_m) = 3T^{\frac{3}{2}} \sum_{n,k=1}^{\infty} (-1)^{n+k} \frac{n^2 k^2}{[n^2 t_m + k^2 (T - t_m)]^{\frac{5}{2}}}, \quad (2.37)$$

que nous pouvons écrire :

$$\sigma_2(t_m) = \frac{1}{T} f\left(\frac{t_m}{T}\right), \quad (2.38)$$

avec :

$$f(x) = 3 \sum_{n,k=1}^{\infty} (-1)^{n+k} \frac{n^2 k^2}{[n^2 x + k^2(1-x)]^{\frac{5}{2}}}. \quad (2.39)$$

Ce résultat est, à notre connaissance, nouveau, de même que celui donnant la loi jointe $\rho(M, t_m)$. Rappelons pour comparaison que, dans le cas d'un mouvement brownien libre, $f(x) = 1/\pi\sqrt{x(1-x)}$ et, dans le cas d'un pont brownien, $f(x) = 1$. Remarquons que f telle que donnée par (2.39) est symétrique autour de $x = \frac{1}{2}$, comme l'on peut s'y attendre, de par la symétrie des contraintes pesant sur l'excursion brownienne.

Nous examinerons le comportement asymptotique de f lorsque $x \rightarrow 0 / x \rightarrow 1$ un peu plus loin (§ 2.2.3), de même que nous donnerons une représentation graphique de f (figure 2.10).

2.2.2 Méandre brownien

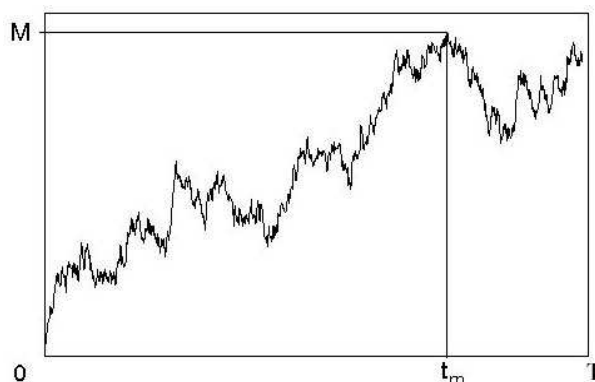


FIG. 2.9 – Méandre brownien

Comme une excursion brownienne, un méandre sur $[0, T]$ est issu de 0 et contraint à demeurer positif; en revanche son point d'arrivée est libre. Nous

repreons le même raisonnement qu'au paragraphe précédent, en notant $x(t)$ un méandre brownien sur $[0, T]$, M son maximum et t_m le temps d'atteinte de ce maximum. Tout se passe comme dans l'exemple précédent, en particulier avec l'introduction d'une coupure ε et d'un décalage x_0 appelés à tendre vers 0 au moment opportun. La seule différence est qu'il nous faut cette fois intégrer sur le point final $x(T)$, qui sera compris entre 0 et M . L'expression pour la portion à gauche de t_m est donc inchangée (éq. 2.30) :

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle = \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \varepsilon)}{M}\right) e^{-\frac{n^2\pi^2}{2M^2}t_m}, \quad (2.40)$$

tandis que l'expression à calculer pour la portion à droite de t_m , qui était donnée par (2.31), devient :

$$\int_0^M \langle M - \varepsilon | e^{-\hat{H}(T-t_m)} | u \rangle du, \quad (2.41)$$

puisque le point final est libre mais que, par définition, il ne peut excéder M et doit être positif. Ceci revient (éq. 2.32) à :

$$\frac{2}{M} \sum_{k=1}^{\infty} \int_0^M \sin\left(\frac{k\pi u}{M}\right) du \sin\left(\frac{k\pi(M - \varepsilon)}{M}\right) e^{-\frac{k^2\pi^2}{2M^2}(T-t_m)}, \quad (2.42)$$

où nous avons interverti la somme infinie et l'intégration, et utilisé la linéarité de cette dernière. L'intégrale se calcule aisément :

$$\int_0^M \sin\left(\frac{k\pi u}{M}\right) du = \frac{M}{k\pi} [1 - (-1)^k].$$

Il suffit alors de combiner (2.40) et (2.42) pour obtenir, dans la limite $\varepsilon, x_0 \ll 1$:

$$\begin{aligned} \rho(M, t_m) = \\ \mathcal{N} \frac{4\pi^2 \varepsilon^2 x_0}{M^4} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] n^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}(T-t_m)} + o(\varepsilon^2 x_0^2), \end{aligned} \quad (2.43)$$

\mathcal{N} désignant comme précédemment une constante de normalisation visant à assurer que :

$$\int_0^{\infty} \int_0^T \rho(M, t_m) dM dt_m = 1.$$

Le calcul de cette constante « fait disparaître » ε et x_0 et nous obtenons finalement **la loi jointe du maximum et du temps d'atteinte du maximum pour un méandre brownien** :

$$\boxed{\rho(M, t_m) = \frac{\sqrt{2}\pi^{\frac{5}{2}}T^{\frac{1}{2}}}{M^4} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] n^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}(T-t_m)}} \quad (2.44)$$

À l'instar de ce qui se passe dans le cas de l'excursion brownienne, le calcul de la loi marginale du maximum M d'un méandre brownien à partir de (2.44) est aisé et aboutit à une expression distincte de celle existant dans la littérature [49, 50], mais dont il est possible de prouver l'identité (*cf.* annexe de [48], p. 195) :

$$\sigma_1(M) = \frac{2^{\frac{3}{2}}\pi^{\frac{1}{2}}T^{\frac{1}{2}}}{M^2} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] \frac{n^2}{k^2 - n^2} e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}T} \quad (2.45)$$

De même que précédemment, la comparaison de cette distribution et de celle trouvée dans la littérature conduit à une identité remarquable que nous n'avons pas su prouver directement, mais que nos tests numériques ont semblé confirmer :

$$\lim_{\alpha \rightarrow -1} \sum_{n,k=1}^{\infty} [\alpha^{n+k} - \alpha^n] \frac{n^2}{k+n} \left(\frac{\ln k - \ln n}{k-n} \right) = \frac{1}{2} \ln 2. \quad (2.46)$$

Si nous nous tournons ensuite vers la marginale selon t_m en intégrant (2.44) sur M , il vient :

$$\sigma_2(t_m) = T^{\frac{1}{2}} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] \frac{n^2}{[n^2t_m + k^2(T-t_m)]^{\frac{3}{2}}}, \quad (2.47)$$

que nous pouvons écrire :

$$\boxed{\sigma_2(t_m) = \frac{1}{T} g\left(\frac{t_m}{T}\right)}, \quad (2.48)$$

avec :

$$g(x) = \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] \frac{n^2}{[n^2x + k^2(1-x)]^{\frac{3}{2}}}. \quad (2.49)$$

Nous pouvons en fait simplifier cette dernière expression en notant que seuls les termes où k est impair contribuent à la somme :

$$g(x) = 2 \sum_{k=0, n=1}^{\infty} (-1)^{n+1} \frac{n^2}{[n^2x + (2k+1)^2(1-x)]^{\frac{3}{2}}}. \quad (2.50)$$

Il s'agit à notre connaissance d'un résultat nouveau. Il est intéressant de noter que, contrairement à son analogue $f(x)$ pour une excursion brownienne, qui est symétrique autour de $x = \frac{1}{2}$, la fonction $g(x)$ est asymétrique et diverge quand $x \rightarrow 1$ — ce que l'on peut comprendre physiquement en combinant le fait que le point final soit libre et le mouvement contraint à être positif, avec la tendance du mouvement brownien à « partir » d'un côté ou de l'autre de la droite réelle.

Nous donnons une représentation graphique de g sur la figure 2.11.

2.2.3 Analyse asymptotique et simulations numériques

Nous analysons ici plus avant deux des résultats obtenus plus haut : la densité de probabilité $f(x)$ du temps d'atteinte du maximum pour une excursion brownienne sur l'intervalle $[0, 1]$, donnée par (2.39), et son analogue $g(x)$ pour un méandre, donnée par (2.50).

Nous commençons par l'excursion brownienne, pour laquelle nous représentons sur la figure 2.10 la densité $f(x)$, accompagnée des résultats de simulations numériques (dont l'accord avec la fonction obtenue par le calcul analytique est très bon) et d'asymptotes dont nous détaillons le calcul ci-dessous.

Nous partons de l'identité :

$$\frac{1}{a^{\frac{5}{2}}} = \frac{1}{\Gamma\left(\frac{5}{2}\right)} \int_0^{\infty} v^{\frac{3}{2}} e^{-av} dv, \quad (2.51)$$

et récrivons (2.39) comme suit :

$$f(x) = \frac{4}{\sqrt{\pi}} \int_0^{\infty} v^{\frac{3}{2}} \left\{ \sum_{n=1}^{\infty} (-1)^n n^2 e^{-n^2 xv} \right\} \left\{ \sum_{k=1}^{\infty} (-1)^k k^2 e^{-k^2(1-x)v} \right\} dv. \quad (2.52)$$

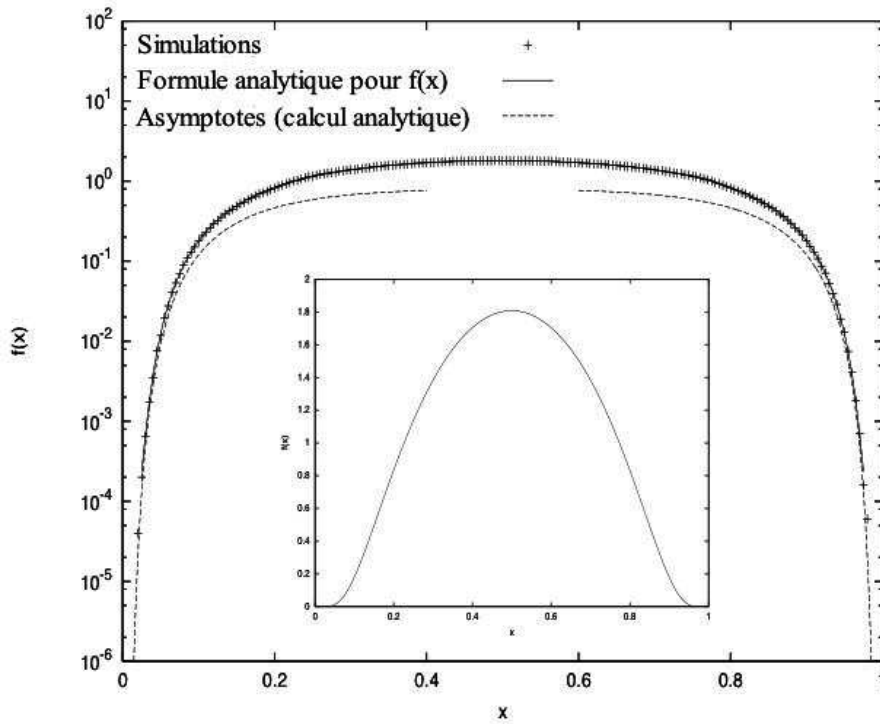


FIG. 2.10 – Distribution $f(x)$ du temps d'atteinte du maximum pour une excursion brownienne sur l'intervalle $[0, 1]$ (éq. (2.39)) : comparaison avec des simulations numériques (1 000 000 de réalisations) et les calculs asymptotiques (en insert, $f(x)$ en échelle linéaire).

Cette ré-écriture nous permet de séparer les sommes, mais bien sûr au prix d'une intégration à effectuer sur la variable muette v .

Nous considérons le cas limite $x \rightarrow 0$ et posons $v = \frac{w}{\sqrt{x}}$:

$$f(x) = \frac{4}{\sqrt{\pi}} \frac{1}{x^{\frac{5}{4}}} \int_0^\infty w^{\frac{3}{2}} \left\{ \sum_{n=1}^{\infty} (-1)^n n^2 e^{-n^2 \sqrt{x} w} \right\} \left\{ \sum_{k=1}^{\infty} (-1)^k k^2 e^{-k^2 (1-x) \frac{w}{\sqrt{x}}} \right\} dw. \quad (2.53)$$

Dans la limite $x \rightarrow 0$, la seconde somme sera dominée par le terme $k = 1$:

$$\sum_{k=1}^{\infty} (-1)^k k^2 e^{-k^2 (1-x) \frac{w}{\sqrt{x}}} \sim -e^{-\frac{w}{\sqrt{x}}}. \quad (2.54)$$

Pour la première somme apparaissant dans (2.53), nous partons d'une identité empruntée à la théorie des fonctions thêta de Jacobi [51] :

$$1 + 2 \sum_{n=1}^{\infty} (-1)^n e^{-n^2 w} = 2 \left(\frac{\pi}{w} \right)^{\frac{1}{2}} \sum_{n=0}^{\infty} e^{-\frac{\pi^2 (n+1/2)^2}{w}}. \quad (2.55)$$

En dérivant chacun des membres de (2.55) par rapport à w , en remplaçant ensuite w par $\sqrt{x}z$ et en ne conservant que le terme dominant dans le membre de droite, il vient :

$$\sum_{n=1}^{\infty} (-1)^n n^2 e^{-n^2 \sqrt{x}z} \sim -\frac{\pi^{\frac{5}{2}}}{4x^{\frac{5}{4}} z^{\frac{5}{2}}} e^{-\pi^2/4\sqrt{x}z}. \quad (2.56)$$

En insérant (2.56) et (2.54) dans (2.53), nous obtenons alors :

$$f(x) \sim \frac{\pi^2}{x^{\frac{5}{2}}} \int_0^{\infty} \frac{1}{z} \exp \left[-\frac{1}{\sqrt{x}} \left(\frac{\pi^2}{4z} + z \right) \right] dz. \quad (2.57)$$

La méthode de Laplace nous permet ensuite d'obtenir le terme asymptotique dominant :

$$\boxed{f(x) \underset{x \rightarrow 0}{\sim} \sqrt{2} \frac{\pi^2}{x^{\frac{9}{4}}} e^{-\frac{\pi}{\sqrt{x}}}.} \quad (2.58)$$

Notons que le comportement de $f(x)$ en $x \rightarrow 1$ est, par symétrie de f , donné par (2.58) en y remplaçant x par $1 - x$.

Nous nous tournons maintenant vers le méandre brownien et la distribution $g(x)$ du temps d'atteinte de son maximum, telle que donnée par (2.50). Cette distribution est représentée graphiquement sur la figure 2.11. Faisant appel initialement à la stratégie que nous venons de mettre en œuvre dans le cas de l'excursion brownienne et de la distribution f , nous ré-écrivons d'abord :

$$g(x) = \frac{4}{\sqrt{\pi}} \int_0^{\infty} u^{\frac{1}{2}} \left\{ \sum_{n=1}^{\infty} (-1)^{n+1} n^2 e^{-n^2 x u} \right\} \left\{ \sum_{k=0}^{\infty} e^{-(2k+1)^2 (1-x)u} \right\} du. \quad (2.59)$$

Comme précédemment, nous remplaçons u par z/\sqrt{x} et observons que dans la seconde somme constituant (2.59), le terme dominant dans la limite $x \rightarrow 0$ correspondra simplement au terme $k = 0$, tandis que la première somme est

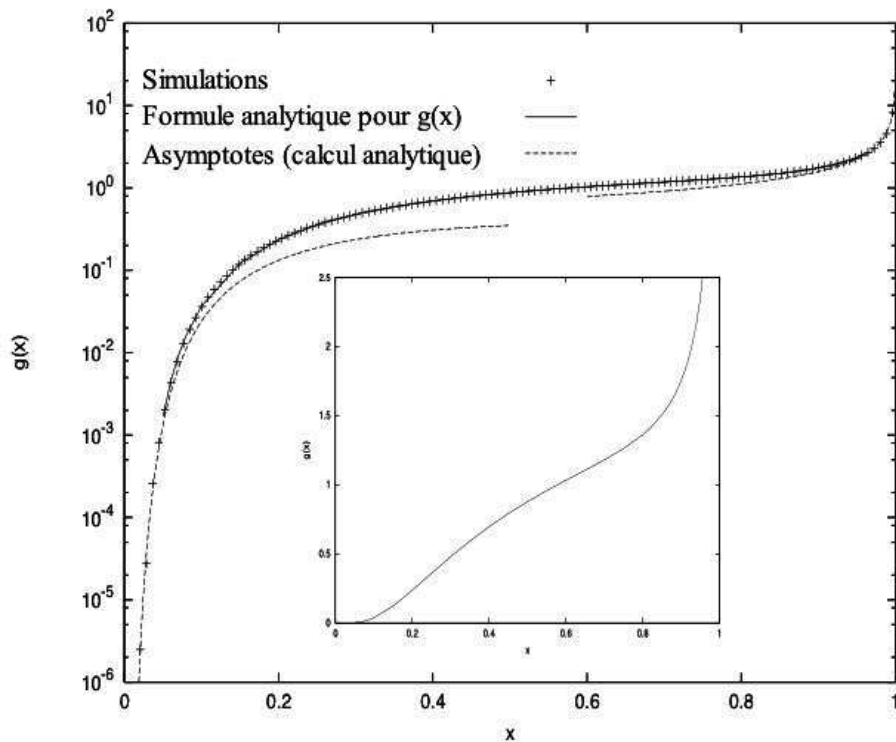


FIG. 2.11 – Distribution $g(x)$ du temps d'atteinte du maximum pour un méandre brownien sur l'intervalle $[0, 1]$ (éq. (2.50)) : comparaison avec des simulations numériques et les calculs asymptotiques (en insert, $g(x)$ en échelle linéaire).

identique à celle qui apparaissait déjà dans (2.53), dont nous avons calculé un équivalent (éq. 2.56). Ceci nous conduit au résultat suivant :

$$g(x) \underset{x \rightarrow 0}{\sim} 2^{\frac{2}{3}} \frac{\pi}{x^{\frac{7}{4}}} e^{-\frac{\pi}{\sqrt{x}}}. \quad (2.60)$$

Dans la limite $x \rightarrow 1$, nous ne pouvons cette fois-ci invoquer d'argument de symétrie à l'instar de ce que nous avons fait pour l'excursion brownienne.

Nous repartons donc de (2.59) en posant $u = z/(1 - \sqrt{x})$:

$$g(x) = \frac{4}{\sqrt{\pi}(1-x)^{\frac{3}{4}}} \int_0^\infty z^{\frac{1}{2}} \left\{ \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 x z / \sqrt{1-x}} \right\} \left\{ \sum_{k=0}^\infty e^{-(2k+1)^2 \sqrt{1-x} z} \right\} dz. \quad (2.61)$$

La seconde somme peut se traiter en faisant à nouveau appel à l'identité (2.55) relative aux fonctions theta de Jacobi ; ce qui aboutit à :

$$\sum_{k=0}^\infty e^{-(2k+1)^2 \sqrt{1-x} z} \underset{x \rightarrow 1}{\sim} \frac{\sqrt{\pi}}{4} \left(\frac{1}{1-x} \right)^{\frac{1}{4}} \frac{1}{z^{\frac{1}{2}}}. \quad (2.62)$$

Quant à la première somme de (2.61), il se trouve que tous ses termes contribuent au même ordre dans la limite $x \rightarrow 1$. Combinant (2.61) et (2.62), nous avons donc :

$$g(x) \underset{x \rightarrow 1}{\sim} \frac{1}{1-x} \int_0^\infty \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 z / \sqrt{1-x}} dz = \frac{1}{\sqrt{1-x}} \sum_{n=1}^\infty (-1)^{n+1}; \quad (2.63)$$

le dernier terme n'a pas de sens tel qu'écrit, il faut se souvenir de la régularisation qui lui est associée pour l'interpréter correctement :

$$\sum_{n=1}^\infty (-1)^{n+1} = \lim_{\alpha \rightarrow -1} \sum_{n=1}^\infty \alpha^{n+1} = \lim_{\alpha \rightarrow -1} \frac{\alpha^2}{1-\alpha} = \frac{1}{2}. \quad (2.64)$$

Le comportement asymptotique de $g(x)$ dans la limite $x \rightarrow 1$ se révèle donc *in fine* très simple :

$$\boxed{g(x) \underset{x \rightarrow 1}{\sim} \frac{1}{2\sqrt{1-x}}}. \quad (2.65)$$

Il est intéressant de noter que, à un facteur constant près, la divergence est la même que pour un mouvement brownien libre (éq. (2.16)), en racine carrée, ce qui confirme l'interprétation de cette divergence donnée plus haut en termes de la tendance à « partir » du mouvement brownien.

Chapitre 3

Temps d'atteinte du maximum avant le premier passage

Les divers types de mouvement browniens linéaires que nous avons considérés jusqu'ici présentent tous la caractéristique d'être d'une durée T fixée. Les contraintes que nous leur avons imposées portaient uniquement sur leur position à certains instants, mais ne modifiaient pas la durée totale du mouvement.

Dans ce chapitre, nous allons examiner le cas d'un mouvement brownien libre, issu d'un point $x_0 > 0$, que l'on observe jusqu'à son temps de premier passage en 0. En notant t_f cet instant, cela signifie que nous considérons un mouvement brownien $x(t)$ sur un intervalle $[0, t_f]$ dont la longueur elle-même varie d'une réalisation à une autre. Les temps de premiers passages en une valeur donnée (ici en 0) ont fait l'objet de nombreuses études et l'on pourra consulter par exemple [52] pour approfondir ce sujet. La distribution du maximum d'un mouvement brownien avant son premier passage en 0 a été calculée par Kearney et Majumdar [53], par la résolution d'une équation de Fokker-Planck « à rebours ». Nous nous proposons ici de calculer la loi jointe du maximum M et de son temps d'atteinte t_m , dont la marginale sur M redonne le résultat de Kearney et Majumdar, et dont la marginale sur t_m donne la loi du temps d'atteinte du maximum avant son premier passage pour un mouvement brownien libre issu d'un point strictement positif — ce résultat étant à notre connaissance nouveau [54].

3.1 Motivations

Outre l'intérêt théorique de la question que nous posons dans ce chapitre, deux exemples « concrets » peuvent venir motiver le calcul que nous entreprenons.

Exemple 1 : période d'occupation d'une file attente

Considérons une file d'attente à serveur unique, en temps discret, modélisée par une marche aléatoire [39, 55] :

$$l_n = l_{n-1} + \xi_n \tag{3.1}$$

où l_n est la longueur de la file au temps n et les ξ_n sont des variables aléatoires indépendantes et identiquement distribuées prenant chacune pour valeur $+1$ avec probabilité p (cas correspondant à l'arrivée d'un individu dans la file), -1 avec probabilité q (passage d'un individu au serveur ou départ d'un individu) ou 0 avec probabilité $(1 - p - q)$. C'est ce que l'on appelle une file Geo/Geo/1 [55].

Étant donnée une longueur initiale $l_0 > 0$, on appelle *période d'occupation* la durée au-delà de laquelle la file redevient vide pour la première fois, tel qu'illustré sur la figure 3.1. Il est assez naturel de s'interroger sur le moment auquel la file est la plus longue au cours d'une telle période d'occupation, et ce moment est précisément, dans le modèle de la marche aléatoire, le temps m auquel la déviation l_m de la marche (issue de $l_0 > 0$) est maximale avant son premier passage en 0.

Exemple 2 : cours d'une action cotée en bourse

En finance, un modèle classique du prix S_n d'une action au temps n consiste à exprimer S_n comme l'exponentielle de la position d'une marche aléatoire [8] :

$$S_n = R e^{l_n},$$

où l_n est une marche aléatoire définie comme en (3.1) et R une valeur de référence.

En partant d'une certaine valeur initiale S_0 , le cours de l'action connaît une évolution stochastique. Une question naturelle pour le détenteur de l'action consiste à savoir quand la revendre. Il peut par exemple considérer que

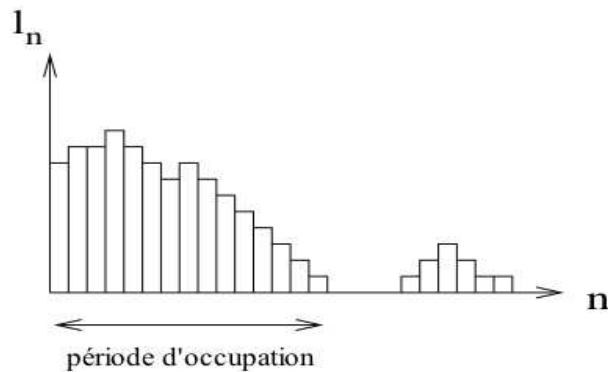


FIG. 3.1 – Période d’occupation d’une file d’attente en temps discret.

si le cours passe sous un certain seuil, disons R , il devient trop risqué pour lui financièrement de conserver cette action plus longtemps — autrement dit il souhaite revendre cette action avant que le rapport S_n/R ne devienne inférieur à 1. S’il souhaite la revendre au meilleur prix, la question à laquelle il doit être capable de répondre est alors : à quel moment le rapport S_n/R atteint-il son maximum (en partant de $S_0/R > 1$), avant d’atteindre la valeur 1 ? En supposant que le processus aléatoire défini par $l_n = \ln(S_n/R)$ soit bien une marche aléatoire issue de $l_0 = \ln(S_0/R) > 0$, $S_n/R = 1$ correspond alors à $l_n = 0$ et nous retrouvons la même question que dans l’exemple 1 : celle du temps d’atteinte du maximum d’une marche aléatoire avant son premier passage en 0¹.

Nous traiterons ici non pas exactement le problème présenté dans ces deux exemples, mais un problème proche, puisque nous ne considérons pas une marche aléatoire en temps discret mais un mouvement brownien en temps continu. Il semble en effet plus aisé d’obtenir dans ce contexte les résultats désirés, dont nous pouvons espérer qu’ils correspondent aux comportements asymptotiques des cas en temps discret.

¹En utilisant la même modélisation du cours d’une action, la distribution du temps d’atteinte du maximum pour un mouvement libre sur un intervalle de temps donné permet au porteur s’étant fixé une date limite de revente de connaître (statistiquement parlant) le meilleur moment pour céder ses parts. Le cas général avec dérive a d’ailleurs été traité récemment [36].

3.2 Loi jointe et lois marginales

Nous calculerons la loi jointe du maximum et du temps d'atteinte du maximum $\rho(M, t_m)$ puis la marginale sur t_m , $\sigma_2(t_m)$, qui est à notre connaissance un résultat nouveau et qui nous intéresse ici plus particulièrement. Nous présentons d'abord le calcul direct de la loi de M tel que donné dans [53].

Soit $q(x)$ la probabilité pour qu'un mouvement brownien issu de $x \in [0, M]$ quitte cet intervalle *via* 0 — autrement dit, que le maximum du mouvement brownien avant son premier passage en 0 soit inférieur ou égal à M . En notant $\phi_{\Delta t}(\Delta x)$ la distribution d'un déplacement brownien sur l'intervalle temporel Δt , nous avons :

$$q(x) = \int q(x + \Delta x) \phi_{\Delta t}(\Delta x) d\Delta x.$$

En développant ensuite $q(x + \Delta x)$ pour de petites valeurs de Δx , et en utilisant le fait que pour un mouvement libre la valeur moyenne de Δx est 0, nous voyons que q satisfait :

$$\frac{d^2 q}{dx^2} = 0, \quad q(0) = 1, \quad q(M) = 0, \quad (3.2)$$

dont la solution est :

$$q(x) = 1 - \frac{x}{M}. \quad (3.3)$$

Il suffit alors de dériver par rapport à M pour obtenir la densité de probabilité de M :

$$\sigma_1(M) = \frac{x}{M^2}. \quad (3.4)$$

Pour le calcul de la loi jointe $\rho(M, t_m)$, nous procédons suivant la même méthode que celle appliquée plus haut à l'excursion et au méandre (§ 2.2) : nous supposons que le maximum est atteint au temps t_m et séparons le chemin brownien en deux parties (avant et après t_m), que nous pouvons traiter indépendamment grâce à la nature markovienne du mouvement. Nous introduisons comme plus haut un paramètre de coupure ε pour contourner la

difficulté posée par le fait qu'un mouvement brownien repasse un nombre infini de fois en un point qu'il vient de visiter et ne peut donc être contraint à atteindre son maximum M en t_m puis à quitter l'intervalle $[0, M]$ par 0 sans repasser par M . La figure 3.2 illustre ces points.

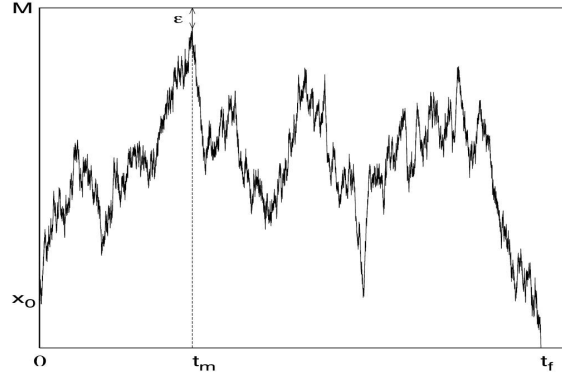


FIG. 3.2 – Étude de la loi jointe du maximum M et de son temps d'atteinte t_m pour un mouvement brownien libre issu de $x_0 > 0$ et considéré jusqu'à son temps de premier passage en 0. Nous introduisons une coupure ε et séparons les portions à gauche et à droite de t_m , qui sont indépendantes.

À droite de t_m , nous avons un processus allant de $M - \varepsilon$ en $t = t_m$ à 0 en $t = t_f$ à nouveau sans franchir M ni 0 entre temps. Sa mesure est donnée par l'éq. (3.3 :) :

$$q(M - \varepsilon) = \frac{\varepsilon}{M} \quad (3.5)$$

À gauche de t_m , nous avons un processus qui se propage de x_0 en $t = 0$ à $M - \varepsilon$ en $t = t_m$ sans franchir le niveau M ni l'origine 0 sur l'intervalle $[0, t_m)$. Nous utilisons à nouveau la méthode par intégrale de chemin introduite dans [37] qui permet d'exprimer la mesure d'un chemin en termes du propagateur :

$$\left\langle x_0 \left| e^{-\hat{H}t_m} \right| M - \varepsilon \right\rangle,$$

avec

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x),$$

$V(x)$ étant ici simplement un puits de potentiel avec des barrières infinies en $x = 0$ et $x = M$ ($V(x) = 0$ pour $0 < x < M$), comme dans les cas de l'excursion et du méandre browniens.

Les fonctions propres normalisées de \hat{H} , indexées par les entiers $n = 1, 2, 3, \dots$ s'annulent aux extrémités, $x = 0$ et $x = M$, de l'intervalle et sont données par :

$$\psi_n(x) = \sqrt{\frac{2}{M}} \sin\left(\frac{n\pi x}{M}\right).$$

Les valeurs propres associées sont :

$$E_n = \frac{n^2\pi^2}{2M^2}.$$

Il est alors aisé d'évaluer le propagateur qui nous intéresse dans cette base propre :

$$\begin{aligned} \langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle &= \sum_{n=1}^{\infty} \psi_n(x_0) \psi_n(M - \varepsilon) e^{-E_n t_m} \\ &= \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \varepsilon)}{M}\right) e^{-\frac{n^2\pi^2}{2M^2} t_m}. \end{aligned} \quad (3.6)$$

Dans la limite $\varepsilon \rightarrow 0$, nous avons donc à l'ordre dominant :

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle = \frac{2\pi}{M^2} \varepsilon \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) e^{-\frac{n^2\pi^2}{2M^2} t_m} + \mathcal{O}(\varepsilon^3). \quad (3.7)$$

Le produit des éqs. (3.5) and (3.7) nous donne ensuite la densité recherchée, à l'ordre dominant en ε :

$$\rho(M, t_m; \varepsilon) = \mathcal{N} \varepsilon^2 \frac{\pi}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) e^{-\frac{n^2\pi^2}{2M^2} t_m}. \quad (3.8)$$

La constante de proportionnalité \mathcal{N} est déterminée par la condition de normalisation de la densité de probabilité : $\int_{x_0}^{\infty} dM \int_0^{\infty} dt_m \rho(M, t_m; \varepsilon) = 1$. Il apparaît en effectuant le calcul que $\mathcal{N} = \varepsilon^{-2}$. Enfin, donc, en prenant la limite $\varepsilon \rightarrow 0$, nous obtenons :

$$\boxed{\rho(M, t_m) = \frac{\pi}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) e^{-\frac{n^2\pi^2}{2M^2} t_m}} \quad (3.9)$$

En intégrant sur t_m , nous retrouvons :

$$\sigma_1(M) = \frac{2}{\pi M} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sin\left(\frac{n\pi x_0}{M}\right) = \frac{x_0}{M^2} \quad (3.10)$$

Cette série est égale, comme prévu, au résultat de Kearney et Majumdar [53].

L'intégration de (3.9) sur M (qui varie de x_0 à l'infini) conduit à **la loi du temps d'atteinte du maximum d'un mouvement brownien avant son premier passage** :

$$\begin{aligned} \sigma_2(t_m) &= \pi \int_{x_0}^{\infty} \frac{dM}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) e^{-\frac{n^2 \pi^2}{2M^2} t_m} \\ &= \frac{1}{\pi t_m} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \int_0^{n\pi} du \cos(u) e^{-\frac{u^2}{2x_0^2} t_m}. \end{aligned} \quad (3.11)$$

La série qui apparaît dans (3.11) peut s'écrire en termes d'une fonction spéciale, la quatrième fonction theta de Jacobi ϑ_4 [51]; d'où la forme suivante :

$$\boxed{\sigma_2(t_m) = \frac{1}{2\pi t_m} \left[\pi - \int_0^{\pi} \vartheta_4\left(\frac{y}{2}, e^{-y^2 \frac{t_m}{2x_0^2}}\right) dy \right]}. \quad (3.12)$$

Il s'agit d'un résultat à notre connaissance nouveau, que nous étudions plus en détail au paragraphe suivant.

3.3 Graphiques, simulations numériques et asymptotes

La figure 3.3 donne une représentation graphique de f définie par :

$$\sigma_2(t_m) = \frac{1}{x_0^2} f\left(\frac{t_m}{x_0^2}\right),$$

ce qui revient à normaliser le mouvement par sa distance initiale à l'origine. Il est intéressant d'examiner le comportement asymptotique de ρ_2 lorsque t_m

devient grand devant x_0^2 (en se souvenant de la relation d'échelle du mouvement brownien, $x(t) \sim \sqrt{t}$, on peut dire que les temps $t \gg x_0^2$ sont suffisamment grands pour que le mouvement ait « oublié » sa condition initiale) et lorsque t_m est très petit devant x_0^2 (en reprenant la même image, ces temps sont tels que le mouvement n'a pas encore eu le temps d'« oublier » sa position initiale).

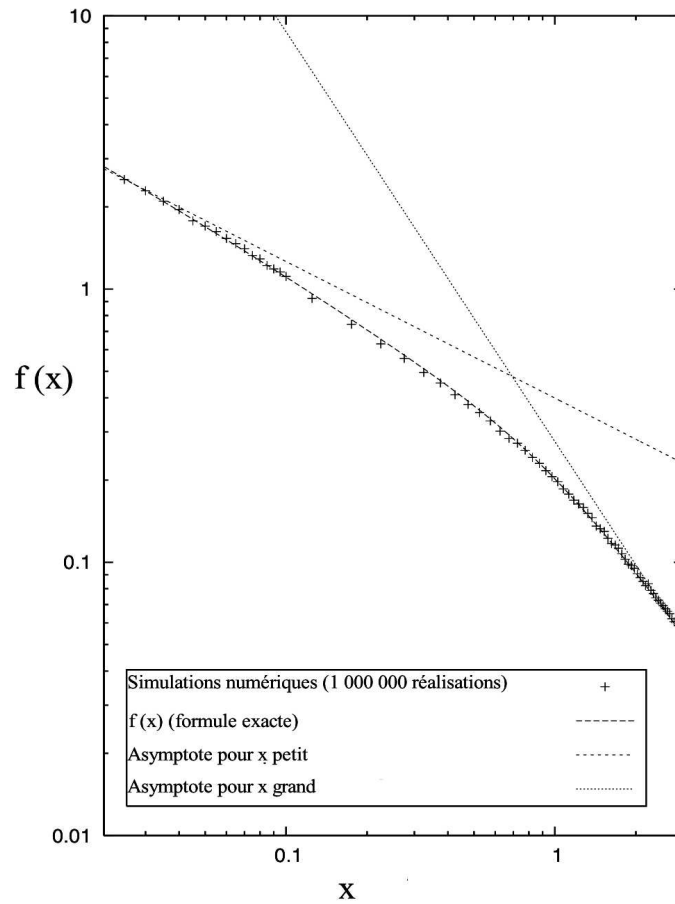


FIG. 3.3 – Distribution du temps d’atteinte du maximum pour un mouvement brownien libre avant son premier passage en zéro (éq. (3.12) pour la distribution et éqs. (3.15), (3.13) pour ses asymptotes). (Échelle logarithmique sur les deux axes.)

Asymptote pour t_m grand :

Nous considérons d'abord les cas où $t_m \gg x_0^2$. En changeant de variable dans (3.11) :

$$z = \sqrt{\frac{t_m}{2x_0^2}} u$$

et en prenant la limite $z \rightarrow 0$ nous obtenons, pour $t_m \gg x_0^2$:

$$\boxed{\sigma_2(t_m) \approx \frac{x_0 \ln 2}{t_m^{3/2}} \sqrt{\frac{1}{2\pi}}} \quad (3.13)$$

Asymptote pour t_m petit :

Dans le cas opposé où $t_m \ll x_0^2$, nous partons de l'éq. (3.9) et effectuons une transformée de Laplace :

$$\begin{aligned} \int_0^\infty dt_m e^{-st_m} \rho(M, t_m) &= \frac{2}{\pi M} \sum_{n=1}^\infty (-1)^{n+1} \frac{n}{(n^2 + \frac{2M^2s}{\pi^2})} \sin\left(\frac{n\pi x_0}{M}\right) \\ &= \frac{\sinh(x_0\sqrt{2s})}{M \sinh(M\sqrt{2s})}, \end{aligned}$$

(la somme de la série se trouve par exemple dans [56].) Si s devient très grand devant x_0^{-2} et M^{-2} , nous obtenons alors :

$$\int_0^\infty dt_m e^{-st_m} \rho(M, t_m) \approx \frac{e^{-\sqrt{2s}(M-x_0)}}{M},$$

qui donne, après inversion de la transformée de Laplace ([57]) :

$$\rho(M, t_m) \approx \frac{t_m^{-3/2}}{\sqrt{2\pi}} \frac{(M-x_0)}{M} e^{-\frac{(M-x_0)^2}{2t_m}}. \quad (3.14)$$

Enfin, une intégration sur M donne, pour $t_m \ll x_0^2$:

$$\boxed{\sigma_2(t_m) \approx \frac{1}{x_0 \sqrt{2\pi t_m}}} \quad (3.15)$$

Ainsi, $\rho_2(t_m)$ a des queues en loi de puissance des deux côtés. Pour t_m grand, la densité de probabilité décroît comme $\rho_2(t_m) \sim t_m^{-3/2}$, alors que pour t_m petit, elle diverge en $\rho_2(t_m) \sim t_m^{-1/2}$. La forme analytique exacte de $\rho_2(t_m)$ (éq. (3.12)) et ses asymptotes (éqs. (3.15), (3.13)) sont représentées sur la figure 3.3 ainsi que les points obtenus par simulation numérique (1 000 000 réalisations).

3.4 Cas d'un mouvement brownien dérivant vers l'origine

Nous pouvons affiner notre traitement du problème en considérant des mouvements non plus libres mais présentant une dérive vers l'origine. Ces mouvements peuvent être décrits par une équation de Langevin du type :

$$\frac{dx}{dt} = -|\mu| + \eta(t), \quad (3.16)$$

où η est un bruit blanc et μ est la dérive du mouvement. Le cas libre étudié ci-dessus correspond à la limite $\mu \rightarrow 0$, et le cas d'un mouvement avec dérive sur un intervalle de temps fixé à été traité récemment par Bouchaud et Majumdar [36].

En procédant exactement comme en l'absence de dérive, c'est-à-dire en introduisant une coupure ε et en séparant le mouvement autour de t_m , puis en résolvant une équation de Fokker-Planck « à rebours » pour la partie à droite de t_m et en utilisant la méthode basée sur l'intégrale de chemin et la formule de Feynman-Kac à gauche de t_m , nous obtenons la loi jointe du maximum et du temps d'atteinte du maximum :

$$\boxed{\rho^{(\mu)}(M, t_m) = \frac{|\mu| M e^{|\mu|x_0 - \frac{|\mu|^2}{2} t_m}}{\sinh(|\mu|M)} \rho(M, t_m)} \quad (3.17)$$

où $\rho(M, t_m)$ est la densité jointe pour le mouvement libre. Les détails des calculs sont donnés dans [54] (reproduit p. 181).

Nous ne sommes pas parvenus à dégager une expression compacte de la marginale $\sigma_2^{(\mu)}(t_m) = \int \rho^{(\mu)}(M, t_m) dm$ mais l'équation (3.17) permet un calcul des asymptotes de $\sigma_2^{(\mu)}(t_m)$, comme nous le montrons ci-dessous.

Asymptote pour t_m petit :

En insérant dans l'équation (3.17) le comportement asymptotique de $\sigma_2(t_m)$ pour $t_m \ll x_0^2$ donné par l'équation (3.14) et en intégrant sur M , il vient :

$$\boxed{\sigma_2^{(\mu)}(t_m) \sim \frac{|\mu|e^{|\mu|x_0 - \frac{|\mu|^2}{2}t_m}}{\sinh(|\mu|x_0)\sqrt{2\pi t_m}}} \quad (3.18)$$

Ainsi pour t_m petit, $\sigma_2^{(\mu)}(t_m)$ diverge en $t_m^{-1/2}$, à l'instar de la distribution correspondante du mouvement libre (2.16).

Asymptote pour t_m grand :

Pour étudier le comportement asymptotique de $\sigma_2^{(\mu)}(t_m)$ lorsque $t_m \gg x_0^2$, nous partons de l'expression suivante de $\sigma_2^{(\mu)}(t_m)$:

$$\begin{aligned} \sigma_2^{(\mu)}(t_m) &= \int_{x_0}^{\infty} dM \rho^{(\mu)}(M, t_m) \\ &= \int_{x_0}^{\infty} dM \frac{|\mu|M e^{|\mu|x_0 - \frac{|\mu|^2}{2}t_m}}{\sinh(\mu M)} \rho(M, t_m) \\ &= \int_{x_0}^{\infty} dM \frac{|\mu|\pi e^{|\mu|x_0 - \frac{|\mu|^2}{2}t_m}}{\sinh(|\mu|M)M^2} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) e^{-\frac{n^2\pi^2}{2M^2}t_m}. \end{aligned}$$

La série qui apparaît dans (3.19) est dominée par son premier terme ($n = 1$) lorsque t_m devient très grand devant M^2 . Dès lors, en ne retenant que ce terme $n = 1$ et en effectuant le changement de variable $y = 1/M$ dans l'intégrale, nous obtenons :

$$\pi|\mu|e^{|\mu|x_0 - \frac{|\mu|^2}{2}t_m} \int_0^{\frac{1}{x_0}} dy \frac{\sin(\pi x_0 y)}{\sinh(|\mu|/y)} e^{-\frac{\pi^2 y^2}{2}t_m}.$$

Pour t_m grand, la contribution la plus importante à l'intégrale vient des y petits. En développant les fonctions sin et sinh et en ne conservant que les termes dominants, l'intégrale se réduit à :

$$2\pi^2 x_0 |\mu| e^{|\mu|x_0 - \frac{|\mu|^2}{2}t_m} \int_0^{\frac{1}{x_0}} dy y e^{-t_m \left(\frac{\pi^2}{2} y^2 + \frac{|\mu|}{y t_m} \right)}.$$

Nous posons ensuite

$$h(y) = \frac{\pi^2}{2} y^2 + \frac{|\mu|}{y t_m},$$

puis nous utilisons la méthode de Laplace pour extraire le terme dominant de l'intégrale, en minimisant la fonction h . Nous obtenons finalement :

$$\sigma_2^{(\mu)}(t_m) \sim \left[2\sqrt{\frac{2}{3}}\pi^{\frac{5}{6}}x_0|\mu|^{\frac{4}{3}}e^{|\mu|x_0} \right] t_m^{-5/6} e^{-\frac{|\mu|^2}{2}t_m - \frac{3}{2}(|\mu|\pi)^{2/3}t_m^{1/3}} \quad (3.19)$$

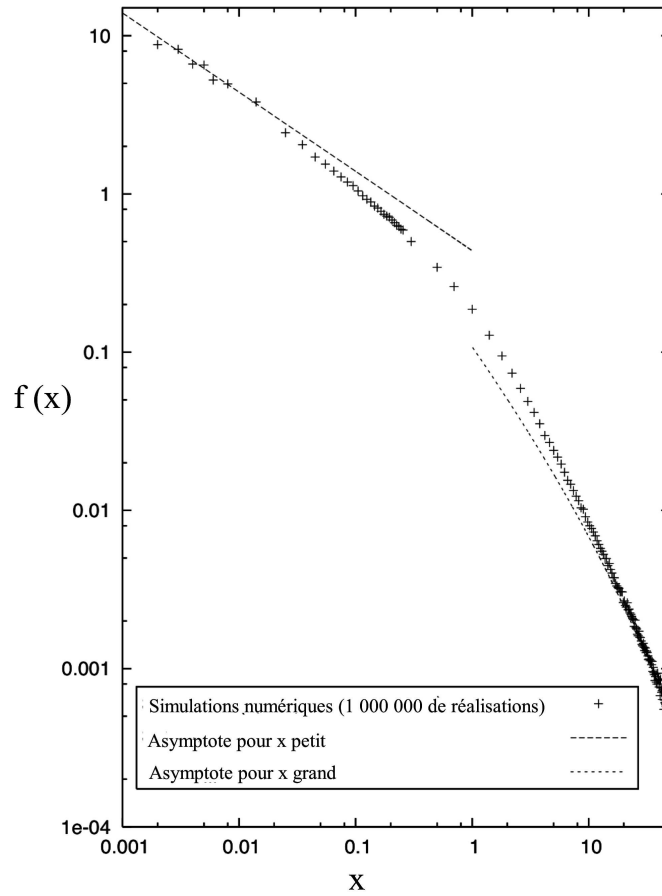


FIG. 3.4 – Temps d'atteinte du maximum avant le premier passage en 0 pour un mouvement dérivant vers l'origine ($|\mu| = 0.1$ et $x_0 = 1$). (Asymptotes données par les éq. (3.18) et (3.19).) (Échelle logarithmique sur les deux axes.)

Ainsi, comme l'on pouvait s'y attendre pour un mouvement dérivant vers l'origine, $\sigma_2^{(\mu)}(t_m)$ décroît exponentiellement lorsque t_m devient grand.

La figure 3.4 donne une représentation graphique des asymptotes (éq. 3.18 et 3.19) ainsi que les points issus d'une simulation numérique (1 000 000 de réalisations avec $|\mu| = 0.1$ et $x_0 = 1$).

Chapitre 4

Temps d'atteinte du maximum pour n mouvements browniens linéaires

Suivant le fil des trois premiers chapitres de ce mémoire, nous généralisons ici le problème que nous y avons traité — à savoir celui du temps d'atteinte du maximum pour différents types de mouvements browniens linéaires « isolés » — au cas de plusieurs mouvements browniens linéaires considérés simultanément sur le même intervalle de temps. Il peut s'agir par exemple de n mouvements browniens libres et indépendants, issus de 0, que l'on examine sur un intervalle de durée T fixée, ou bien de n ponts browniens, issus de 0 et retournant en 0 au temps T .

Cette généralisation, des plus naturelles sur le plan théorique, trouvera une application physique concrète dans la seconde partie de ce mémoire (*via* la décomposition en mouvements unidimensionnels de mouvements browniens plans).

4.1 Cas de n mouvements browniens libres indépendants

Commençons par étudier le cas de n mouvements browniens libres indépendants (fig. 4.1) :

$$\begin{aligned} & x_1(t), x_2(t), \dots, x_n(t) \\ & t \in [0, T], \quad x_i(0) = 0 \\ & \frac{dx_i}{dt} = \eta_i(t), \quad \eta_i \text{ bruits blancs gaussiens indépendants} \end{aligned}$$

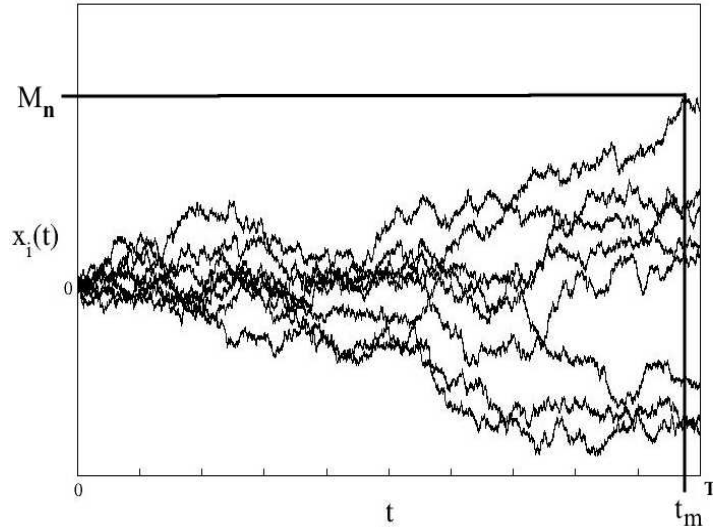


FIG. 4.1 – Mouvements browniens libres indépendants sur l'intervalle $[0, 1]$. (Maximum global M et temps d'atteinte t_m .)

Nous désignons par M la variable aléatoire correspondant au maximum global des n mouvements sur l'intervalle $[0, T]$:

$$M = \max_{t,i} \{x_i(t)\}, \quad (4.1)$$

et par t_m le temps d'atteinte de ce maximum. Il existe donc un indice i^* tel que :

$$x_{i^*}(t_m) = M. \quad (4.2)$$

Nous notons $\rho_n(M, t_m)$ la loi jointe de M et de t_m , $F_n(M)$ la fonction de répartition de M et σ_n la densité de probabilité de t_m .

Comme dans le cas d'un mouvement « isolé », la loi du maximum peut se calculer assez aisément directement, sans passer par la loi jointe $\rho_n(M, t_m)$. En effet, les mouvements étant indépendants, nous pouvons écrire :

$$\boxed{F_n(M) = [F(M)]^n}, \quad (4.3)$$

où

$$F(M) = F_1(M) = \operatorname{erf}\left(\frac{M}{\sqrt{2T}}\right)$$

comme nous l'avons établi au chapitre 2 (éq. (2.15)). Nous pouvons ensuite, si nous le désirons, calculer les moments de M .

Quant à la densité de t_m , elle peut se calculer en observant que la probabilité pour que le maximum global des n mouvements soit M et qu'il soit atteint au temps t_m peut s'écrire comme la probabilité que l'un des n mouvements atteigne son maximum M au temps t_m et que les $n - 1$ autres mouvements ne dépassent jamais M sur l'intervalle $[0, T]$:

$$\boxed{\rho_n(M, t_m) = n\rho_1(M, t_m)[F(M)]^{n-1}}. \quad (4.4)$$

Nous avons montré au chapitre 2 comment calculer la loi jointe $\rho_1(M, t_m)$ (éq. (2.13)) :

$$\rho_1(M, t_m) = \frac{M}{\pi t_m^{\frac{3}{2}} \sqrt{T - t_m}} e^{-\frac{M^2}{2t_m}}.$$

Une intégration sur M conduit ensuite à la densité marginale souhaitée :

$$\sigma_n(t_m) = \int_0^\infty n \frac{M}{\pi t_m^{\frac{3}{2}} \sqrt{T - t_m}} e^{-\frac{M^2}{2t_m}} \left[\operatorname{erf}\left(\frac{M}{\sqrt{2T}}\right) \right]^{n-1} dM \quad (4.5)$$

$$= \frac{2n}{T\pi\sqrt{z_{t_m}(1-z_{t_m})}} \int_0^\infty u e^{-u^2} [\operatorname{erf}(u\sqrt{z_{t_m}})]^{n-1} du, \quad (4.6)$$

où nous avons posé $z_{t_m} = \frac{t_m}{T}$.

Nous pouvons écrire

$$\boxed{\sigma_n(t_m) = \frac{1}{T} f_n\left(\frac{t_m}{T}\right)},$$

avec :

$$f_n(z) = \frac{2n}{\pi\sqrt{z(1-z)}} \int_0^\infty u e^{-u^2} [\operatorname{erf}(u\sqrt{z})]^{n-1} du. \quad (4.7)$$

Si pour $n = 1$, nous retrouvons bien la loi de l'arcsinus (éq. (2.16)), il est intéressant de noter que la symétrie de cette loi est brisée dès $n = 2$, comme le montre la figure 4.2. Ceci peut se comprendre, en apparence peut-être paradoxalement, en invoquant le même argument que celui donné pour éclairer la forme de la loi de l'arcsinus (page 17). En effet, le mouvement brownien ayant tendance à « partir » d'un côté ou de l'autre de son origine, il suffit qu'un seul des n mouvements considérés « parte » du côté positif pour que le temps d'atteinte du maximum global se trouve décalé vers la droite de l'intervalle de temps considéré, comme on peut d'ailleurs l'observer sur la figure 4.1. Ceci a d'autant plus de chances de se produire que le nombre n de mouvements considérés est grand — d'où la dissymétrie de la densité de probabilité de t_m , un phénomène de concentration vers la droite courant en statistique des valeurs extrêmes.

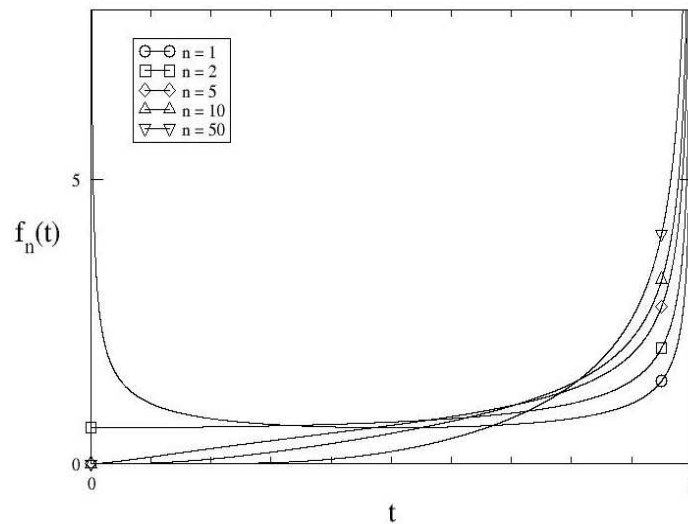


FIG. 4.2 – Distribution du temps d'atteinte du maximum global pour n mouvements browniens indépendants.

4.2 Cas de n ponts browniens indépendants

Dans le cas de n ponts browniens indépendants (fig. 4.3), le raisonnement est exactement identique à celui du paragraphe précédent, mais bien sûr $F(M)$, la fonction de répartition du maximum pour $n = 1$, et ρ_1 , la densité jointe du maximum M et de son temps d'atteinte t_m , diffèrent. Nous les avons toutes deux calculées au chapitre 2 (éq. (2.21) et (2.19)) :

$$F(M) = 1 - e^{-\frac{2M^2}{T}} \quad (4.8)$$

$$\rho_1(M, t_m) = \frac{2T}{\pi} \frac{M^2}{[t_m(T - t_m)]^{\frac{3}{2}}} e^{-\frac{M^2 T}{2t_m(T - t_m)}}, \quad (4.9)$$

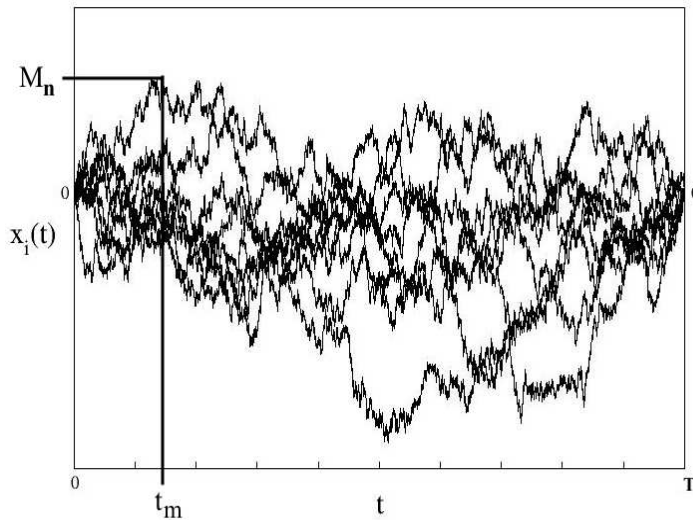


FIG. 4.3 – Ponts browniens indépendants sur l'intervalle $[0, 1]$. (Maximum global M et temps d'atteinte t_m .)

Nous obtenons alors pour n ponts browniens indépendants, à partir de l'éq. (4.4) :

$$\sigma_n(t_m) = \frac{4n}{T\sqrt{\pi}} \int_0^\infty u^2 e^{-u^2} \left[1 - e^{-4u^2 z_{t_m}(1-z_{t_m})}\right]^{n-1} du \quad (4.10)$$

où nous avons posé comme au paragraphe précédent $z_{t_m} = \frac{t_m}{T}$.

Nous pouvons écrire

$$\sigma_n(t_m) = \frac{1}{T} g_n \left(\frac{t_m}{T} \right),$$

avec :

$$g_n(z) = n \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{(-1)^k}{[1 + 4kz(1-z)]^{\frac{3}{2}}}. \quad (4.11)$$

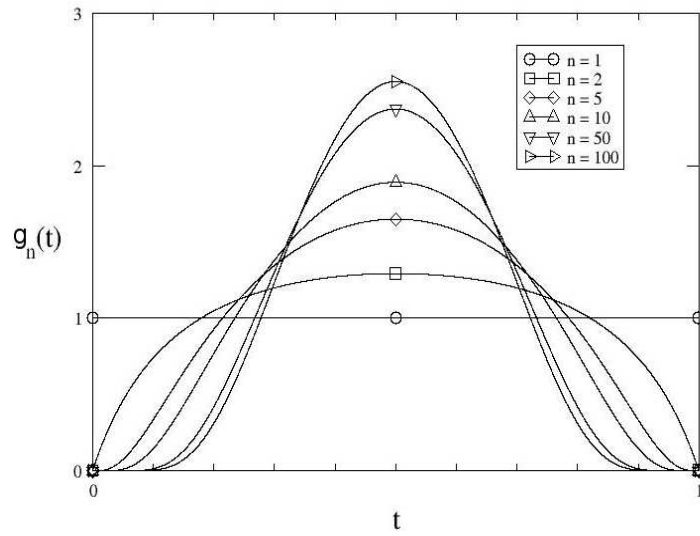


FIG. 4.4 – Distribution du temps d’atteinte du maximum global pour n ponts browniens indépendants.

Remarquons ici aussi que pour $n = 1$ nous retrouvons aisément la distribution uniforme du temps d’atteinte de son maximum par un pont brownien. Toutefois, contrairement à ce qui se passe dans le cas de mouvements libres, la symétrie de la distribution du temps d’atteinte du maximum n’est pas brisée lorsque l’on passe de $n = 1$ à $n > 1$ ponts browniens (fig. 4.4). Notons cependant qu’elle apparaît d’autant plus piquée autour du centre de l’intervalle que n devient grand.

Le premier point (la persistance de la symétrie) se comprend aisément de par la symétrie des contraintes imposées au mouvement ($x(0) = x(T) = 0$). Notons parallèlement qu’un artefact du cas $n = 1$, à savoir le fait que les

extrémités de l'intervalle ($t_m = 0$ et $t_m = T$) n'apparaissent pas comme des points singuliers de la distribution du temps d'atteinte du maximum, disparaît pour $n > 1$: comme si cet artefact, qui correspond au cas d'une excursion dans \mathbb{R}^- (dont le maximum est alors atteint en $t_m = 0$), cédait le pas à la probabilité de rencontrer une sorte de trajet « optimal » en terme de maximum — un pont brownien utilisant la première moitié de l'intervalle pour « monter » vers son maximum et la seconde moitié pour « redescendre » en 0, d'où un temps d'atteinte du maximum au milieu de l'intervalle.

4.3 Cas de n excursions et de n méandres browniens indépendants

En principe, nous pouvons suivre le même raisonnement pour obtenir la distribution du temps d'atteinte du maximum global de n excursions browniennes indépendantes ou de n méandres browniens indépendants. Il suffit pour cela d'insérer dans l'équation (4.4) les résultats obtenus au chapitre 2 pour $F(M)$, la fonction de répartition du maximum quand $n = 1$, et pour ρ_1 , la densité jointe du maximum M et de son temps d'atteinte t_m quand $n = 1$, à savoir les éq. (2.35) et (2.34) pour une excursion brownienne :

$$F(M) = \frac{2^{\frac{3}{2}}\pi^{\frac{5}{2}}T^{\frac{3}{2}}}{M^4} \sum_{n,k=1}^{\infty} (-1)^{n+k} \frac{n^2k^2}{k^2 - n^2} \left[e^{\frac{n^2\pi^2}{2M^2}T} - e^{\frac{k^2\pi^2}{2M^2}T} \right] \quad (4.12)$$

$$\rho_1(M, t_m) = \sqrt{2} \frac{\pi^{\frac{9}{2}}T^{\frac{3}{2}}}{M^6} \sum_{n,k=1}^{\infty} (-1)^{n+k} n^2k^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}(T-t_m)}. \quad (4.13)$$

et les éq. (2.45) et (2.44) pour un méandre brownien :

$$F(M) = \frac{2^{\frac{3}{2}}\pi^{\frac{1}{2}}T^{\frac{1}{2}}}{M^2} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] \frac{n^2}{k^2 - n^2} e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}T} \quad (4.14)$$

$$\rho_1(M, t_m) = \frac{\sqrt{2}\pi^{\frac{5}{2}}T^{\frac{1}{2}}}{M^4} \sum_{n,k=1}^{\infty} [(-1)^{n+k} - (-1)^n] n^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{k^2\pi^2}{2M^2}(T-t_m)}. \quad (4.15)$$

La relative complexité de ces expressions ne laissant pas entrevoir de forme simple pour $\sigma_n(t_m)$, et la motivation physique sous-tendant l'étude des deux cas précédents étant moins prégnante dans les cas présents, nous n'avons pas poussé plus avant nos calculs.

4.4 Mouvements non indépendants

Il paraît assez naturel d'envisager également le cas où les n mouvements browniens considérés ne sont plus indépendants. En particulier, on peut s'intéresser à des chemins sans intersection, autrement dit qui conservent le même ordre de superposition du début à la fin de l'intervalle (*cf.* fig. 4.5). Ce type d'ensemble de chemins est parfois dénommé « marcheurs malveillants », nom que leur donne M. E. Fisher par analogie avec des marcheurs ivres dont la rencontre tournerait immédiatement à l'homicide (!).

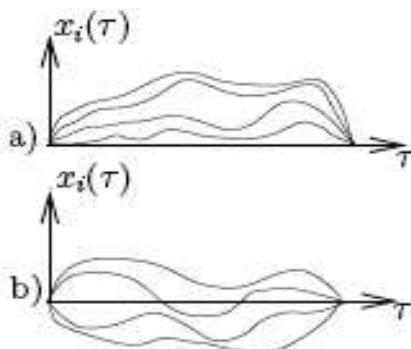


FIG. 4.5 – Marcheurs « malveillants » : avec (a) ou sans (b) barrière en 0.

P.-G. de Gennes fut semble-t-il le premier à introduire une configuration de ce type, dès les années 1960, pour fournir un modèle soluble de structure fibreuse [43], inspiré notamment par les solutions aqueuses lipidiques dans lesquelles les chaînes de lipides peuvent former des structures lamellaires. M. E. Fisher reprit ce type de modèle pour traiter de problèmes de mouillage et de fusion [42]. On le rencontre également dans des contextes de réseaux de polymères [58, 59], de persistance pour des systèmes hors d'équilibre [60] ou des modèles de croissance stochastique [61, 62].

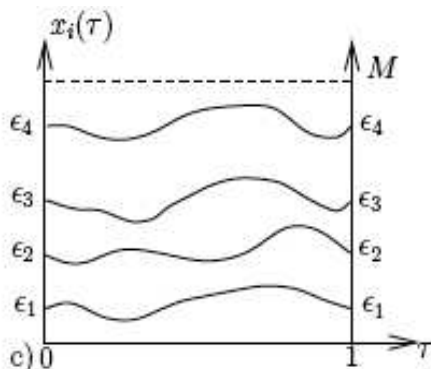


FIG. 4.6 – Marcheurs « malveillants » avec une barrière en 0 et des paramètres de coupure ε_i appelés à tendre vers 0.

Autre domaine d'utilisation du modèle des marcheurs malveillants, l'étude des interfaces [63, 64]. C'est dans ce contexte, d'ailleurs, mais avec un seul marcheur, qu'avait été présenté l'essentiel de la méthode que nous utilisons ici, en particulier les notions de coupure et le recours à l'intégrale de chemin [20, 37, 38]. On voit assez bien comment un mouvement brownien unique peut représenter une interface simple. Les marcheurs malveillants, eux, interviennent par exemple dans la description de transitions de phase faisant apparaître plusieurs domaines séparés par des parois émergeant en un point et se terminant en un autre — exactement comme des ponts browniens ou des excursions browniennes sans intersection. La taille de ces dislocations peut alors être caractérisée par la hauteur maximale des chemins sans intersection, hauteur dont Schehr *et al.* [63] ont calculé la distribution exacte en appliquant essentiellement la même méthode que celle présentée dans ce chapitre : les n chemins considérés partent de et arrivent en des points ε_i (*cf.* fig. 4.6), ce qui permet d'éviter que la contrainte $x_1(0) = x_2(0) = \dots = x_n(0)$ et $x_1(t) < x_2(t) < \dots < x_n(t)$ ne résulte en une probabilité nulle du fait que les mouvements ainsi définis se recouperaient un nombre infini de fois immédiatement après leur point de départ. Comme précédemment, pour calculer la fonction de répartition du maximum M , il suffit d'introduire dans le hamiltonien correspondant à la traduction de l'intégrale de chemin *via* la formule de Feynman-Kac, un potentiel présentant une barrière infinie en M .

Cependant, pour tenir compte de la condition de non-intersection, les fonctions propres que l'on retiendra pour l'expression du propagateur seront celle de type « fermionique », c'est-à-dire s'annulant dès que deux de leurs arguments (correspondant aux positions des marcheurs browniens) sont égaux ¹.

Ces remarques closent ce chapitre et la première partie de ce mémoire, dont nombre des résultats seront utilisés dans la seconde partie, conduisant à l'étude de l'enveloppe convexe du mouvement brownien plan.

¹L'étude complète est disponible dans [63], article que nous reproduisons page 213. Cette étude se situe comme on le voit dans un prolongement naturel du sujet de cette thèse. Ma contribution personnelle à ce travail n'ayant pas été centrale, je n'ai pas souhaité inclure ces travaux dans ce mémoire.

Deuxième partie

Autour du
mouvement brownien plan :
enveloppes convexes aléatoires

LES ensembles convexes sont définis par la propriété que le segment joignant deux points quelconques de l'ensemble est lui-même entièrement contenu dans l'ensemble. Dans le monde physique, il existe de nombreux exemples de formes convexes, depuis certains éléments acoustiques rencontrés dans les salles de concert [65], jusqu'à la cristallographie, où la construction dite de Wulff conduit, dans le cas général, à un polyèdre convexe dont les faces correspondent aux plans cristallins minimisant l'énergie de surface [66]. Peuvent également être notés des travaux récents en neurosciences, dont les résultats indiquent que le cerveau humain semble plus à même de distinguer deux formes différentes lorsque celles-ci sont convexes [67], ce qui est particulièrement intéressant lorsque l'on connaît par ailleurs la grande utilisation des propriétés de convexité dans le traitement informatique des images, notamment en matière de détection de motifs [68]. De telles applications seraient bien sûr limitées si elles ne s'adressaient qu'à des formes naturellement, intrinsèquement convexes ; mais l'on peut en fait élargir leur champ d'application en utilisant une sorte d'« approximation » des objets non-convexes par un objet convexe défini de manière univoque : parmi tous les ensembles convexes susceptibles de contenir un objet donné, il suffit de choisir le plus petit en termes de volume. C'est ce qu'on appelle l'enveloppe convexe de l'objet, et celle-ci peut être utilisée à des fins de comparaisons de formes complexes, par exemple dans le cas de protéines et de leurs sites de fixation [69]. Les enveloppes convexes suscitent donc, de par leur généralité et leurs propriétés, beaucoup d'intérêt, tant pour les défis que leur calcul présente en algorithmique [70–79] que pour leurs applications [68, 69, 80–82].

Les enveloppes convexes aléatoires sont les enveloppes convexes d'objets aléatoires, un exemple simple étant le triangle formé par trois points choisis au hasard dans le plan. Il est bien sûr possible de prendre plus de trois points, et l'enveloppe convexe changera à chaque tirage : son nombre d'arêtes, son périmètre, son aire varieront à chaque réalisation. De surcroît, les points peuvent être choisis de manière indépendante ou, au contraire, avec un certain degré de corrélation : ils peuvent être les positions d'une marche aléatoire ou d'un mouvement brownien, par exemple. Que ce soit dans ce cas-ci ou dans celui-là, on peut s'enquérir de la statistique des propriétés géométriques de l'enveloppe convexe de l'échantillon : comment varient-elles

selon la distribution des points ? Cette question a suscité beaucoup d'intérêt parmi les mathématiciens au cours des cinquante dernières années, et a généré un corpus conséquent dont nous essaierons de dégager quelques lignes et résultats principaux dans le premier chapitre de cette partie. Par contraste, les références relatives à cette question dans la littérature physique semblent rares, alors même que l'approche d'un physicien peut conduire à des résultats nouveaux, en particulier dans des cas correspondants à des situations physiques [83], comme nous le montrerons dans les deux derniers chapitres, à l'aide d'une méthode tirant parti des résultats obtenus dans la première partie de ce mémoire.

Chapitre 5

Aperçu de travaux et résultats existants relatifs aux enveloppes convexes aléatoires

Nous passons ici en revue quelques articles relatifs à l'enveloppe convexe de points aléatoires. Toute prétention d'exhaustivité serait bien sûr déplacée : le fait qu'un article donné figure dans la liste qui suit résulte de la conjonction de plusieurs facteurs, dont sa pertinence et son accessibilité pour un lecteur physicien, ses liens éventuels avec les résultats nouveaux que nous donnerons dans cette seconde partie et... le hasard, comme dans toute recherche bibliographique.

Nous adoptons une présentation chronologique et résumerons dans un tableau en fin de chapitre ceux des résultats de cette liste qui sont particulièrement pertinents pour notre étude.

- Dans son ouvrage consacré aux processus stochastiques et au mouvement brownien (paru en 1948) [84], P. Lévy aborde, en quelques paragraphes et essentiellement de manière heuristique, la question de l'enveloppe convexe du mouvement brownien plan : « Ce contour [celui de l'enveloppe convexe du mouvement brownien plan] est, à un ensemble de mesure nulle près, constitué par des parties rectilignes. »
- Plus de dix ans plus tard, en 1959, J. Geffroy est, semble-t-il, le premier à publier des résultats relatifs à l'enveloppe convexe d'un nuage de

points tirés au hasard selon une distribution donnée [85], en l'espèce n points choisis dans \mathbb{R}^2 suivant une distribution normale f . Il montre que si l'on note :

- ∂C_n la frontière de l'enveloppe convexe du nuage,
- Σ_n l'ellipsoïde d'équation¹ $f(x, y) = \frac{1}{n}$,
- Δ_n la distance entre ∂C_n et Σ_n ,
- et Λ_n le rayon maximal d'un disque ouvert ayant son centre à l'intérieur de l'enveloppe convexe et ne contenant aucun des points du nuage,

alors presque sûrement :

$$\Delta_n \text{ et } \Lambda_n \xrightarrow[n \rightarrow \infty]{} 0 \quad (5.1)$$

En d'autres termes, l'enveloppe convexe du nuage « tend » vers l'ellipsoïde d'équation $f(x, y) = \frac{1}{n}$. Il généralisera ce résultat à \mathbb{R}^d , $d \geq 1$, en 1961 [86].

- En 1961 également, F. Spitzer et H. Widom publient leurs résultats concernant le périmètre L_n de l'enveloppe convexe d'une marche aléatoire représentée par des sommes de nombres complexes aléatoires Z_i . En combinant une identité due à M. Kac et une formule d'A. Cauchy (dont ils sont les premiers à faire mention dans ce contexte), ils établissent l'identité suivante [87] :

$$\mathbb{E}(L_n) = 2 \sum_{k=1}^n \frac{\mathbb{E}(|S_k|)}{k}, \quad (5.2)$$

¹Par exemple, pour le cas gaussien :

$$f(x, y) = \exp \left[-\frac{x^2 + y^2}{2} \right],$$

et l'ellipsoïde en question est tout simplement le cercle de centre l'origine et de rayon $\sqrt{2 \ln n}$. Nous verrons plus loin comment obtenir directement le comportement asymptotique du périmètre moyen de l'enveloppe convexe dans le cas de n points choisis dans le plan selon une distribution normale. Il est donné par :

$$\langle L_n \rangle \sim 2\pi \sqrt{2 \ln n},$$

en parfait accord avec le résultat de Geffroy.

avec $S_k = Z_1 + Z_2 + \cdots + Z_k$, la somme partielle de rang k des Z_i .

Ils donnent deux exemples :

1. En écrivant $Z_k = X_k + iY_k$ et en prenant $\mathbb{E}(X_k) = \mathbb{E}(Y_k) = 0$, $\mathbb{E}(X_k^2) = a^2$, $\mathbb{E}(Y_k^2) = b^2$ et $\mathbb{E}(X_k Y_k) = \rho ab$, on a :

$$\mathbb{E}(L_n) \underset{n \rightarrow \infty}{\sim} 4c\sqrt{n}, \quad (5.3)$$

où $c(a, b, \rho)$ ne dépend pas de n .

2. Si l'on prend $Z_k = X_k + i$ avec $\mathbb{E}(X_k) = \mu$ et $\mathbb{E}((X_k - \mu)^2) = \sigma^2$, on a :

$$\mathbb{E}(L_n) \underset{n \rightarrow \infty}{\sim} 2n\sqrt{1 + \mu^2} + \frac{\sigma^2}{(1 + \mu^2)^{\frac{3}{2}}} \ln n, \quad (5.4)$$

à comparer avec le double de la longueur du vecteur $n(\mu + i)$.

- Toujours en 1961, et pour une marche aléatoire très générale, représentée par une somme de n vecteurs dans le plan complexe ($Z_k = X_k + iY_k$ avec X_k et Y_k des variables aléatoires réelles données par une distribution jointe), G. Baxter [88] démontre trois formules relatives, respectivement :
 - au nombre F_n de côtés de l'enveloppe convexe de la marche aléatoire,
 - au nombre K_n de pas de la marche appartenant au contour de l'enveloppe convexe,
 - à la longueur L_n du contour de l'enveloppe convexe (c'est-à-dire son périmètre).

Dans ce dernier cas, la formule est la même que celle démontrée par Spitzer et Widom [87], mais cette fois-ci par un raisonnement purement combinatoire, qui ne fait pas appel à la formule de Cauchy mais repose sur le décompte du nombre de permutations de l'ordre des pas de la marche dans lesquelles une somme partielle donnée appartient au contour de l'enveloppe convexe.

Les formules sont :

$$\mathbb{E}(F_n) = 2 \sum_{m=1}^n \frac{1}{m} \sim 2 \ln n \quad (5.5)$$

$$\mathbb{E}(K_n) = 2 \quad (5.6)$$

$$\mathbb{E}(L_n) = 2 \sum_{k=1}^n \frac{\mathbb{E}(|S_k|)}{k} \quad (5.7)$$

avec $S_k = Z_1 + Z_2 + \dots + Z_k$, la somme partielle de rang k des Z_i .

- En 1963, A. Rényi et R. Sulanke publient le premier [89] de leurs deux articles fondateurs, consacrés à l'enveloppe convexe de n points aléatoires indépendants et identiquement distribués P_i , $i = 1..n$, en dimension 2. En notant F_n le nombre de côtés de cette enveloppe convexe, ils montrent que :

1. Si les P_i sont distribués uniformément dans un polygone convexe K possédant r côtés :

$$\mathbb{E}(F_n) = \frac{2}{3}r(\ln n + \gamma) + T(K) + o(1) \quad (5.8)$$

où γ est la constante d'Euler et T une constante ne dépendant que de K et qui est maximale pour les polygones réguliers à r côtés.

2. Si les P_i sont distribués uniformément à l'intérieur d'un convexe K à bord lisse :

$$\mathbb{E}(F_n) \underset{n \rightarrow \infty}{\sim} \alpha(K)n^{\frac{3}{2}} \quad (5.9)$$

où $\alpha(K)$ est une constante dépendant de K

3. Si les P_i suivent une distribution normale dans tout l'espace :

$$\mathbb{E}(F_n) \underset{n \rightarrow \infty}{\sim} 2\sqrt{2\pi \ln n} \quad (5.10)$$

- L'année suivante, en 1964, le second article de Rényi et Sulanke [90] étend leurs résultats, en s'intéressant au comportement asymptotique (dans la limite où le nombre de points n devient grand) du périmètre L_n et de l'aire A_n de l'enveloppe convexe de n points aléatoires P_i tirés uniformément et indépendamment à l'intérieur d'un convexe K de périmètre L et d'aire A :

1. Si K est à bord lisse :

$$\mathbb{E}(L_n) = L - \mathcal{O}\left(n^{-\frac{2}{3}}\right) \quad (5.11)$$

$$\mathbb{E}(A_n) = A - \mathcal{O}\left(n^{-\frac{2}{3}}\right) \quad (5.12)$$

2. Si K est un carré de côté a :

$$\mathbb{E}(L_n) = 4a - \mathcal{O}\left(n^{-\frac{1}{2}}\right) \quad (5.13)$$

$$\mathbb{E}(A_n) = a^2 - \mathcal{O}\left(\frac{\ln n}{n}\right) \quad (5.14)$$

- B. Efron s'inspire, en 1965, des travaux de Rényi et Sulanke pour établir les formules correspondantes en dimension 3, ainsi que l'espérance du nombre de sommets (de faces en dimension 3), du périmètre et de l'aire de l'enveloppe convexe d'un nuage de points indépendants tirés d'une distribution normale en dimension 2 ou 3, ou d'une distribution uniforme à l'intérieur du disque ou de la sphère (résultats dont Efron montrent qu'ils peuvent facilement être adaptés à toute ellipse dans le plan et tout ellipsoïde dans l'espace, ou à toute distribution normale « à symétrie elliptique ») :

1. Par exemple, pour $n > 3$ points du plan tirés indépendamment selon une distribution gaussienne, en notant $\phi(x) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)$ et $\Phi(x) = \int_{-\infty}^x \phi(y) dy$:

$$\mathbb{E}(V_n) = 4\sqrt{\pi} \binom{n}{2} \int_{-\infty}^{\infty} \Phi^{n-2}(p) \phi^2(p) dp, \quad (5.15)$$

$$\mathbb{E}(L_n) = 4\pi \binom{n}{2} \int_{-\infty}^{\infty} \Phi^{n-2}(p) \phi^2(p) dp, \quad (5.16)$$

$$\mathbb{E}(A_n) = 3\pi \binom{n}{3} \int_{-\infty}^{\infty} \Phi^{n-3}(p) \phi^3(p) dp, \quad (5.17)$$

où V_n , L_n et A_n désignent respectivement le nombre de sommets, le périmètre et l'aire de l'enveloppe convexe du nuage de points.

2. En dimension 3, on a :

$$\mathbb{E}(F_n) = 4\sqrt{3\pi} \binom{n}{3} \int_{-\infty}^{\infty} \Phi^{n-3}(p) \phi^3(p) dp, \quad (5.18)$$

$$\mathbb{E}(H_n) = \frac{3}{2} \mathbb{E}(F_n), \quad (5.19)$$

$$\mathbb{E}(V_n) = \frac{1}{2} \mathbb{E}(F_n) + 2, \quad (5.20)$$

$$\mathbb{E}(L_n) = 24\sqrt{3\pi} \binom{n}{3} \int_{-\infty}^{\infty} \Phi^{n-3}(p) \phi^3(p) dp, \quad (5.21)$$

$$\mathbb{E}(A_n) = 12\pi \binom{n}{3} \int_{-\infty}^{\infty} \Phi^{n-3}(p) \phi^3(p) dp, \quad (5.22)$$

F_n et H_n désignant respectivement le nombre de faces et le nombre d'arêtes de l'enveloppe convexe (L_n est ici la somme de la longueur des arêtes et A_n la somme des aires des faces ; V_n désigne toujours le nombre de sommets).

Efron calcule également le volume moyen de l'enveloppe convexe de n vecteurs tirés indépendamment selon une même distribution normale (moyenne nulle, variance unité) dans un espace de dimension $d < n$:

$$\mathbb{E}(\text{Vol}_n) = \frac{2\pi^{\frac{1}{2}d}}{\Gamma(\frac{1}{2}d)} \left(\frac{d+1}{d} \right) \binom{n}{d+1} \int_{-\infty}^{\infty} \Phi^{n-d-1}(p) \phi^{d+1}(p) dp \quad (5.23)$$

(expression à multiplier par 2 pour $n = d + 1$).

- En 1965 toujours, H. Raynaud présente, dans les Comptes Rendus de l'Académie des Sciences [91], une généralisation à \mathbb{R}^d des formules de Rényi-Sulanke et d'Efron relatives à l'espérance du nombre de sommets de l'enveloppe convexe d'un nuage de n points indépendants, choisis soit selon une même loi normale dans tout \mathbb{R}^d , soit selon une loi uniforme à l'intérieur d'un domaine convexe. En dimension p , pour une loi normale de variance $\frac{a}{2}$ et de moyenne nulle, Raynaud calcule la densité de probabilité de l'enveloppe convexe d'un nuage de n points indépendants. Il montre que, dans la limite où n devient grand, cette distribution converge vers une distribution de Poisson uniforme sur la sphère de centre l'origine et de rayon $\sqrt{a \ln n}$.

- En 1970, H. Carnal [92] revient sur le cas d'un nuage de points dans le plan, pour des distributions dont la seule restriction est qu'elles soient à symétrie sphérique. Il étudie les comportements asymptotiques du nombre moyen de côtés de l'enveloppe convexe, de son périmètre moyen et de son aire moyenne. En particulier, il montre que le nombre moyen de côtés de l'enveloppe convexe, pour certaines distributions, tend vers une constante lorsque le nombre de points du nuage devient très grand.
- Raynaud consacre en 1970 un deuxième article [93] à l'enveloppe convexe de nuages de n points indépendants (distribués suivant une loi uniforme à l'intérieur d'une sphère ou suivant une loi normale gaussienne dans tout l'espace) dans \mathbb{R}^d . Il donne les détails des résultats annoncés plus tôt [91] ainsi que des expressions asymptotiques du nombre de faces $F_n^{(d)}$ (ou de côtés si $d = 2$) de l'enveloppe convexe lorsque n devient grand. Il montre en particulier que pour une gaussienne normalisée² :

$$\mathbb{E}(F_n^{(d)}) \underset{n \rightarrow \infty}{\sim} \frac{2^d}{\sqrt{d}} (\pi \ln n)^{\frac{1}{2}(d-1)} \quad (5.24)$$

- Dix ans plus tard, en 1980, W. Eddy [94] introduit le concept de fonction support dans l'étude des enveloppes convexes aléatoires. En prenant dans le plan n points $P_i = (x_i, y_i)$ de distribution gaussienne, il associe à chacun d'eux le processus aléatoire :

$$B_i(\theta) = x_i \cos \theta + y_i \sin \theta,$$

θ variant de 0 à 2π ($B_i(\theta)$ est la projection du point P_i sur la droite de direction θ). Il définit ensuite le processus

$$M(\theta) = \sup_i \{B_i(\theta)\}$$

dont il démontre que la loi est celle du maximum point à point des processus stochastiques indépendants et identiquement distribués $B_i(\theta)$, en s'appuyant sur les travaux de Brown et Resnick sur ce sujet [95]. Il

²Pour $d = 2$, on retrouve la formule obtenue par Rényi et Sulanke (Eq. 5.10) ; pour $d = 3$, on a $\mathbb{E}(F_n^{(d)}) \sim \frac{8}{\sqrt{3}} \pi \ln n$.

montre que la distribution ponctuelle du processus $M(\theta)$ est une loi de Gumbel et indique, mais sans aller plus loin, que certaines fonctionnelles de ce processus peuvent donner accès à des propriétés géométriques de l'enveloppe convexe du nuage de points :

$$\mathbb{E}(L_n) = \int_0^{2\pi} M(\theta) d\theta \quad (5.25)$$

pour le périmètre ; et :

$$\mathbb{E}(A_n) = \frac{1}{2} \int_0^{2\pi} [M^2(\theta) - (M'(\theta))^2] d\theta \quad (5.26)$$

pour l'aire, avec $M'(\theta) = \frac{dM}{d\theta}$. Ces fonctionnelles correspondent à ce que l'on appellera ici les formules de Cauchy.

- C'est précisément une de ces formules de Cauchy que L. Takács [96] suggère d'utiliser pour calculer le périmètre moyen de l'enveloppe convexe du mouvement brownien plan, en réponse au problème posé par G. Letac dans l'*American Mathematical Monthly* en 1978 [97]. Si l'on note L_T le périmètre de l'enveloppe convexe d'un mouvement brownien plan de durée T , on a :

$$\boxed{\mathbb{E}(L_T) = \sqrt{8\pi T}}. \quad (5.27)$$

- En 1981, W. Eddy et J. Gale [98] poursuivent le travail entrepris par W. Eddy [94] à partir de l'isomorphisme entre l'ensemble des convexes du plan et l'ensemble de leurs fonctions supports (qui est un sous-ensemble des fonctions continues). Leur démarche consiste à décrire la distribution de l'enveloppe convexe d'un échantillon aléatoire à partir d'une distribution sur les fonctions continues, autrement dit à partir d'un processus stochastique, comme l'a fait Eddy dans le cas d'un échantillon gaussien [94]. Eddy et Gale soulignent le lien entre les statistiques de valeurs extrêmes appliquées à n variables aléatoires unidimensionnelles et la distribution de l'enveloppe convexe de n variables aléatoires multidimensionnelles, puisque la distribution de l'enveloppe convexe de n points de \mathbb{R}^d correspond au processus

donné par le maximum point à point des n processus stochastiques que sont les fonctions support $B_i(\theta)$ définies dans le premier article d'Eddy [94]. En considérant des distributions de l'échantillon à symétrie sphérique classées en trois catégories selon que leurs queues sont à décroissance exponentielle, algébrique ou tronquée (distribution sur un disque par exemple), Eddy et Gale calculent la distribution asymptotique du processus stochastique associé (c'est-à-dire la fonction support de l'échantillon $M(\theta)$) quand le nombre n de points de l'échantillon devient grand). Les trois cas correspondent à trois types de distribution du processus limite, données par les lois de Gumbel, Fréchet et Weibull. Eddy et Gale notent également que l'espérance du nombre de sommets de l'enveloppe aléatoire dans le cas « Fréchet », c'est-à-dire pour des distributions de l'échantillon initial ayant des queues en loi de puissance, tend vers une constante (comme l'avait montré Carnal [92]).

- Suivant une voie différente, N. Jewell et J. Romano montrent l'année suivante, en 1982, la correspondance entre le problème de l'enveloppe convexe d'un nuage de points aléatoires et le recouvrement du cercle unité par un nombre fini d'arcs dont la position et la longueur suivent une loi bivariée. Ainsi, pour des arcs de longueur π :

$$\text{Proba (cercle recouvert)} = \text{Proba (env. conv. du nuage contient l'origine)}$$

et plus généralement, pour des arcs de longueur différente de π :

$$\text{Proba (cercle recouvert)} = \text{Proba (env. conv. du nuage contient un cercle donné)}$$

- En 1983, M. El Bachir, dans sa thèse [99], étudie l'enveloppe convexe $C(T)$ du mouvement brownien plan $\mathcal{B}(t)$ considéré jusqu'au temps T . Il démontre en particulier la conjecture de Lévy [84] selon laquelle $C(T)$ est presque sûrement sans point anguleux. Il montre également que $C(T)$ est un processus de Markov sur l'ensemble des convexes compacts d'intérieurs non vides contenant l'origine O . Si l'on note $\partial C(T)$ la frontière de $C(T)$, il établit que :

1. $\text{Prob}(\mathcal{B}(T) \in \partial C(T)) = \text{Prob}(O \in \partial C(T)) = 0$
2. $\{t : \mathcal{B}(t) \in \partial C(t)\}$ est de mesure de Lebesgue nulle presque

sûrement.

El Bachir calcule ensuite explicitement, à partir des formules de Cauchy, l'espérance du périmètre et de l'aire de l'enveloppe convexe du mouvement brownien plan. Pour le périmètre, il obtient une formule générale pour des mouvements avec une dérive μ , formule dont un cas particulier est $\mu = 0$ pour lequel on retrouve le $\sqrt{8\pi T}$ de Takács. Pour l'aire, il obtient :

$$\boxed{\mathbb{E}(A_t) = \frac{\pi T}{2}.} \quad (5.28)$$

- Durant la décennie suivante, l'étude de l'enveloppe convexe d'un nuage de points indépendants identiquement distribué suscite beaucoup d'intérêt. C. Buchta [100] établit une formule exacte pour l'aire moyenne de l'enveloppe convexe de n points choisis uniformément à l'intérieur d'un domaine convexe K (les formules existant jusque là étant essentiellement asymptotiques). Quelques années plus tard, F. Affentranger [101] étend le résultat de Buchta à des dimensions supérieures, *via* une relation de récurrence. On trouvera de nombreux détails et références bibliographique dans les synthèses et revues de Buchta [102], R. Schneider [103], W. Weil et J. Wieacker [104].
- Autre voie de recherche active, celle ouverte par Eddy [94] et Gale [98], dont H. Brozius et J. de Haan [105] prolongent les travaux pour des distributions non nécessairement à symétrie sphérique. Brozius [106] s'intéressera ensuite à la convergence en loi vers des processus de Poisson ponctuels des distributions de grandeurs telles que le nombre de sommets de l'enveloppe convexe de n points aléatoires indépendants et identiquement distribués. Les travaux de Davis [107] et Aldous *et al.* [108] développent également ce type d'approche.
- Cranston *et al.* [109] reviennent, après El Bachir [99], sur l'enveloppe convexe $C(t)$ du mouvement brownien et en particulier sur la continuité de sa frontière $\partial C(t)$. Ils démontrent que $\partial C(t)$ est presque sûrement C^1 et font mention de plusieurs travaux importants liés à ces questions, dont ceux de Shimura [110, 111] et Burdzy [112] montrant que pour

tout $\alpha \in (\frac{\pi}{2}, \pi)$, il existe des temps τ aléatoires tels que $C(\tau)$ a des coins d'ouverture α ; et également ceux de Le Gall [113] montrant que la dimension de Hausdorff de l'ensemble des temps auxquels le mouvement brownien visite un coin de $C(t)$ d'angle α est presque sûrement égale à $1 - \frac{\pi}{2\alpha}$. Ils mentionnent également l'article de Lévy [114] et la thèse de S. N. Evans [115] sur le taux de croissance de l'enveloppe convexe du mouvement brownien, ainsi que l'article de Burdzy et J. San Martin [116] sur la courbure de $C(t)$ près du point le plus bas de la trajectoire du mouvement brownien.

- En 1992, D. Khoshnevisan [117] prolonge les travaux de Cranston *et al.* et démontre une inégalité qui permet, en quelque sorte, de transposer à l'enveloppe convexe certaine des propriétés d'échelle du mouvement brownien lui-même.
- En 1993, deux articles s'intéressent à l'enveloppe convexe de points corrélés, en l'espèce les sommets d'une marche aléatoire dans le plan. G. Letac d'abord [118] indique que l'utilisation de la formule de Cauchy permet d'écrire l'espérance du périmètre L_n de l'enveloppe convexe de n importe quelle marche aléatoire de n pas directement en termes de la fonction support $M_n(\theta) = \max_{0 \leq i \leq n} \{x_i \cos \theta + y_i \sin \theta\}$ de la trajectoire :

$$\mathbb{E}(L_n) = \int_0^{2\pi} \mathbb{E}(M_n(\theta)) d\theta, \quad (5.29)$$

ce qui peut fournir une méthode alternative à celles de Spitzer-Widom ou Baxter pour calculer le périmètre de l'enveloppe convexe d'une marche aléatoire.

C'est justement sur la formule de Spitzer-Widom-Baxter (Eq. 5.2) que reviennent T. Snyder et J. Steele [119] en utilisant, à l'instar de Baxter, des méthodes purement combinatoires qui leur permettent de généraliser les résultats antérieurs comme suit :

Soit H_n le nombre de côtés de l'enveloppe convexe d'une marche aléatoire plane de n pas et soit e_i la longueur du i -ième côté. Si f est une fonction à valeurs réelles et si l'on

note $G_n = \sum_{i=1}^{H_n} f(e_i)$, alors :

$$\mathbb{E}(G_n) = 2 \sum_{k=1}^n \frac{\mathbb{E}(f(|S_k|))}{k}, \quad (5.30)$$

S_k désignant comme précédemment la position de la marche après k pas.

- En prenant $f(x) = 1$, on a $G_n = H_n$ (le nombre de côtés de l'enveloppe convexe) et on retrouve le résultat de Baxter :

$$\mathbb{E}(H_n) = 2 \sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow \infty}{\sim} 2 \ln n.$$

- En prenant $f(x) = x$, on a $G_n = L_n$ et on retrouve, comme Baxter, le résultat de Spitzer et Widom sans utiliser la formule de Cauchy :

$$\mathbb{E}(L_n) = 2 \sum_{k=1}^n \frac{\mathbb{E}(|S_k|)}{k}.$$

- En prenant $f(x) = x^2$, G_n est la somme des carrés des côtés, que l'on peut noter $L_n^{(2)}$, et on obtient alors :

$$\mathbb{E}(L_n^{(2)}) = 2n(\sigma_X^2 + \sigma_Y^2),$$

σ_X^2 et σ_Y^2 étant les variances des marginales sur X et Y de la distribution des incréments de la marche.

Snyder et Steele établissent également deux autres résultats d'importance :

1. Un majorant pour la variance $\mathbb{E}(L_n^2)$ (à ne pas confondre avec $\mathbb{E}(L_n^{(2)})$) du périmètre de l'enveloppe convexe de n'importe quelle marche aléatoire plane à n pas :

$$\mathbb{E}(L_n^2) \leq \frac{\pi^2}{2} n(\sigma_X^2 + \sigma_Y^2) \quad (5.31)$$

2. Un résultat sur la probabilité de grandes déviations du périmètre de l'enveloppe convexe d'une marche aléatoire plane de n pas :

$$\text{Prob}(|L_n - \mathbb{E}(L_n)| \geq t) \leq 2e^{-\frac{t^2}{8\pi^2 n}} \quad (5.32)$$

- Dans un article publié en 1996 [120], A. Goldman apporte un nouveau point de vue sur l'enveloppe convexe du mouvement brownien plan. Il établit en effet un lien entre les mosaïques poissoniennes du plan (qui partitionnent \mathbb{R}^2 en domaines convexes aléatoires) et le périmètre de l'enveloppe convexe du pont brownien. Plus précisément :

Soit $D(R)$ le disque ouvert de rayon R et D_i ($i = 1..N_R$) les domaines convexes de la mosaïque poissonienne contenus dans $D(R)$. Soit

$$\phi_i(t) = \sum_{n=1}^{\infty} \exp(-t\lambda_{n,i})$$

la fonction spectrale du domaine D_i (les $\lambda_{n,i}$ sont les valeurs propres du laplacien pour le problème de Dirichlet sur D_i). Enfin, posons :

$$\phi_R(t) = \frac{1}{N_R} \sum_{i=1}^{N_R} \phi_i(t).$$

Alors :

$\phi_R(t)$ a presque sûrement une limite finie $\Phi(t)$ (appelée fonction spectrale infinie) quand R tend vers l'infini, et :

$$\Phi(t) = \frac{1}{4\pi^2 t} \mathbb{E} \left(e^{-\sqrt{2t}L} \right) \quad (5.33)$$

où L désigne le périmètre de l'enveloppe convexe d'un pont brownien plan de durée unité.

Goldman calcule également le premiers moment de L à partir de la formule de Cauchy :

$$\mathbb{E}(L) = \sqrt{\frac{\pi^3}{2}}. \quad (5.34)$$

Pour le calcul du second moment :

$$\mathbb{E}(L^2) = \frac{\pi^2}{3} \left(\pi \int_0^\pi \frac{\sin u}{u} du - 1 \right), \quad (5.35)$$

Goldman se ramène au calcul de $\mathbb{E}(M(\theta)M(0))$, la fonction de corrélation à deux points du processus aléatoire $M(\theta)$ (définie comme précédemment comme étant la fonction support de la trajectoire du pont brownien) :

$$\mathbb{E}(M(\theta)M(0)) = \frac{\sin \theta}{2} \left[\frac{\theta(2\pi - \theta)}{6(\pi - \theta)} + \cotan \theta \right] \quad (5.36)$$

Goldman obtient cette dernière expression à partir de la probabilité qu'une trajectoire brownienne plane $\mathcal{B}_{0,1}$ fermée, de durée unité et issue de l'origine, demeure toute entière à l'intérieur d'un secteur ξ d'ouverture angulaire β :

$$\text{Prob}(\mathcal{B}_{0,1} \in \xi) = \frac{4\pi e^{-r^2}}{\beta} \sum_{k=1}^{\infty} \sin^2 \left(\frac{k\pi\alpha}{\beta} \right) I_\nu(r^2), \quad (5.37)$$

où $\nu = \frac{k\pi}{\beta}$, et, en supposant que l'origine O soit à l'intérieur du secteur ξ , r est la distance entre O et la pointe S du secteur, et α l'angle entre la droite (OS) et le bord le plus proche du secteur. (I_ν est la fonction de Bessel modifiée du premier type d'ordre ν .)

- Dans un article ultérieur [121], Goldman exploite plus avant le lien entre mosaïques poissonniennes du plan et enveloppe convexe du pont brownien plan. Il commence par montrer que le lien établi dans son premier article demeure valide en remplaçant l'enveloppe convexe du pont brownien par celle du mouvement brownien libre. Il rappelle ensuite la conjecture de Kendall sur ce que l'on appelle la cellule de Crofton du plan (il s'agit, dans une mosaïque poissonnienne, du domaine D_0 contenant l'origine) : lorsque l'aire V_0 de cette cellule est grande sa « forme » est « proche » de celle d'un disque. Goldman démontre un résultat allant dans ce sens (en termes de valeurs propres du laplacien pour le problème de Dirichlet) et, grâce au lien qu'il a mis au jour avec l'enveloppe convexe du mouvement brownien en déduit que celle-ci, lorsqu'elle est « petite », a une forme presque « circulaire ». Plus

précisément, si C désigne l'enveloppe convexe du mouvement brownien plan \mathcal{W} de durée unité, $D(r)$ le disque de centre l'origine et de rayon $r \in (0, \infty)$ et si $M = \sup\{\|\mathcal{W}(s)\|, 0 \leq s \leq 1\}$, alors pour tout $\epsilon \in (0, 1)$:

$$\limsup_{a \rightarrow 0} \text{Prob}[D((1 - \epsilon)a) \subset C \subset D(a) | M = a] = 1 \quad (5.38)$$

- Parallèlement aux travaux consacrés à l'enveloppe convexe du mouvement brownien plan, les études sur les distributions asymptotiques (et non pas seulement leurs valeurs moyennes) de grandeurs telles que le nombre de sommets, le périmètre ou l'aire de l'enveloppe convexe d'un nuage de points indépendants tirés uniformément à l'intérieur d'un convexe K du plan se poursuivent, avec notamment les travaux de P. Groeneboom [122] (complété par l'article de S. Finch et I. Huetter [123]), Hsing [124] pour l'aire lorsque K est un disque, Cabo et Groeneboom [125] pour l'aire également mais lorsque K est polygonal, Bräker et Hsing [105] pour la loi jointe de l'aire et du périmètre dans un cadre plus général de distribution des points initiaux, et les travaux plus récents de Vu [126], de Calka et Schreiber [127], de Reitzner [128–130] et de Bárány [131–133].
- En 2009, P. Biane et G. Letac [134] reviennent sur l'enveloppe convexe du mouvement brownien plan, en s'intéressant à l'enveloppe convexe globale de plusieurs copies d'une même trajectoire, obtenues par rotation, dont ils calculent le périmètre.

Nous voyons que les problèmes d'enveloppes convexes aléatoires ont généré beaucoup de travaux au cours des cinquante dernières années. Les principaux résultats pertinents pour l'étude que nous allons présenter ici sont ceux donnant des expressions explicites (exactes ou asymptotiques) pour le périmètre et l'aire moyens de l'enveloppe convexe d'un échantillon aléatoire dans le plan. Nous tentons d'en rassembler les références dans le tableau suivant :

RÉSULTATS CONNUS		Périmètre (moyen)	Aire (moyenne)
Points Indépendants		<i>Rényi et Sulanke [90]</i> (éq. (5.11) et (5.13)); <i>Efron [135]</i> (éq. (5.16) et (5.17)); <i>Carnal [92]</i> ; <i>Buchta [100]</i> ; <i>Affentranger [101]</i>	
Marche aléatoire	ouverte	<i>Spitzer et Widom [87]</i> (éq. (5.2)); <i>Baxter [88]</i> ; <i>Snyder et Steele [119]</i> ; <i>Letac [118]</i>	?
	(un seul marcheur) fermée	?	?
Mouvt brownien	ouvert	<i>Takács [96]</i> (éq. (5.27))	<i>El Bachir [99]</i> (éq. (5.28))
	(un seul marcheur) fermé	<i>Goldmann [120]</i> (éq. (5.34))	?

Nous présenterons plus loin une méthode permettant, à partir des résultats exposés dans la première partie de ce mémoire, non seulement de compléter une partie des cases vides du tableau ci-dessus mais aussi de traiter d'une généralisation particulièrement intéressante sur le plan physique, à savoir les propriétés géométriques de l'enveloppe convexe globale de $n > 1$ chemins browniens plans. Il s'agit à notre connaissance d'un problème n'ayant jusqu'ici pas été abordé. Tous les travaux mentionnés plus haut, pour ceux

qui concernent des ensembles de points corrélés, s'intéressent à une seule marche aléatoire ou un seul mouvement brownien, à l'exception de [134] qui s'intéresse à plusieurs copies d'un même mouvement.

L'enveloppe convexe de plusieurs chemins aléatoires apparaît pourtant très naturellement, à la fois sur le plan théorique et sur le plan physique, comme nous le verrons plus loin (§ 9.1). En outre, le cas le plus simple, celui de n mouvements indépendants (que nous traiterons ici), révèle déjà une structure à nos yeux particulièrement intéressante, en ceci que les propriétés géométriques de l'enveloppe convexe globale dépendent d'une façon *a priori* non triviale de la combinaison des mouvements. Certes, chacun des marcheurs browniens se meut comme s'il était seul ; mais quel impact peut avoir la multiplicité du nombre de marcheurs sur la forme et la taille de l'enveloppe convexe globale ? Celle-ci sera-t-elle semblable à l'enveloppe d'un seul mouvement, ou bien à une enveloppe n fois plus « grande » (en un sens) ? On voit que ces questions sont liées à celle de savoir si les chemins vont avoir tendance à beaucoup se recouvrir ou au contraire à « s'égailler » dans le plan. Dans un contexte distinct mais proche, cette question a été posée par Larralde *et al.* [136] et Yuste [137] : il s'agissait pour eux de déterminer l'ensemble des sites visités par n marcheurs aléatoires indépendants sur un réseau bidimensionnel. Nous verrons plus loin le lien entre leurs résultats et ceux que nous allons obtenir ici en développant une méthode générale inspirée de l'approche de Takacs [96] et d'El Bachir [99].

Chapitre 6

Du maximum unidimensionnel à la fonction support

Avant d'aborder pleinement la question des enveloppes convexes aléatoires, nous examinons un problème qui nous paraît apporter un éclairage physique intéressant sur la méthode que nous introduirons ensuite. (C'est d'ailleurs, chronologiquement, par ce problème que nous sommes arrivés aux fonctions supports et à l'utilisation des formules de Cauchy dans l'étude de l'enveloppe convexe du mouvement brownien plan.)

6.1 Présentation du problème

Considérons un chemin brownien plan fermé de durée T contraint à rester au-dessus d'une droite passant par l'origine et formant un angle θ avec une direction de référence, par exemple l'axe des abscisses (*cf.* fig. 6.1). Physiquement, ce chemin pourrait par exemple correspondre à un polymère, dans le modèle de Rouse [138], dont les extrémités seraient fixées en un même point à la surface d'un substrat sous-jacent incliné¹. Dans un tel contexte,

¹Dans le modèle de Rouse, les monomères composant un polymère sont vus comme des billes reliées entre elles par des ressorts, à l'équilibre thermique avec un réservoir de chaleur. En étiquetant par un indice i les monomères, ceci revient à donner à une configuration $\mathcal{C}(\{\vec{r}_i\})$ le poids statistique suivant :

$$P[\mathcal{C}(\{\vec{r}_i\})] = e^{-\beta[\kappa \sum_i |\vec{r}_i - \vec{r}_{i-1}|^2]},$$

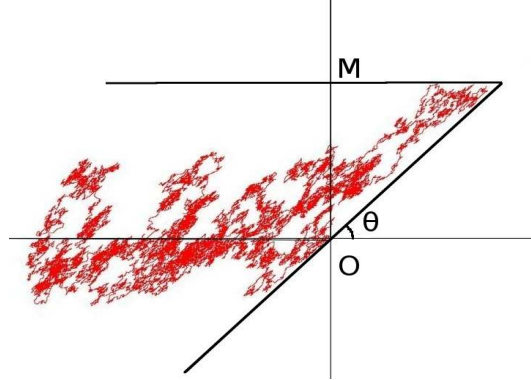


FIG. 6.1 – Maximum suivant l’axe des ordonnées d’un chemin brownien fermé contraint à rester au-dessus d’une droite formant un angle θ avec l’axe des abscisses.

sachant que l’action chimique et/ou biologique d’une molécule dépend souvent de sa configuration spatiale, il peut être intéressant de connaître son extension maximale dans une direction donnée.

Nous cherchons donc ici à calculer la distribution de l’extension maximale, dans la direction de l’axe des ordonnées, du chemin brownien ; autrement dit, nous souhaitons connaître la distribution du maximum de la coordonnée $y(t)$, pour $0 \leq t \leq T$, le mouvement brownien plan considéré étant donné par

$$\mathcal{B}(t) = (x(t), y(t)).$$

Pour $\theta \in (0, \frac{\pi}{2})$, $x(t)$ et $y(t)$ sont des ponts browniens linéaires couplés par la contrainte voulant que, à chaque instant, $y(t) > x(t) \tan \theta$.

$E[\mathcal{C}(\{\vec{r}_i\})] = \kappa \sum_i |\vec{r}_i - \vec{r}_{i-1}|^2$ correspondant à l’énergie de la configuration. Or, dans la représentation par intégrale de chemin que nous utilisons depuis le début de ce mémoire (éq. (2.24)), le poids d’un chemin brownien plan $\vec{x}(t)$ sur l’intervalle $[0, T]$ s’écrit :

$$P[\{\vec{x}(t')\}] \propto \exp \left[-\frac{1}{2} \int_0^T \left(\frac{d\vec{x}}{dt} \right)^2 dt' \right]. \quad (6.1)$$

L’analogie entre les deux modèles se fait donc en envisageant la limite continue dans le modèle de Rouse et en considérant, dans le modèle brownien, t non plus comme le temps mais comme l’abscisse curviligne le long du polymère.

Notons :

$$F_\theta(M) = \text{Prob} \left(\max_{0 \leq t \leq T} \{y(t)\} \leq M \right). \quad (6.2)$$

De manière équivalente, la valeur de $F_\theta(M)$ peut être vue comme la probabilité qu'un chemin brownien fermé issu de l'origine demeure à l'intérieur d'un secteur d'ouverture θ centré sur le point de coordonnées $(M \cotan \theta, M)$ et dont l'un des bords passe par l'origine, tel qu'illustré sur la figure 6.2 (après rotation).

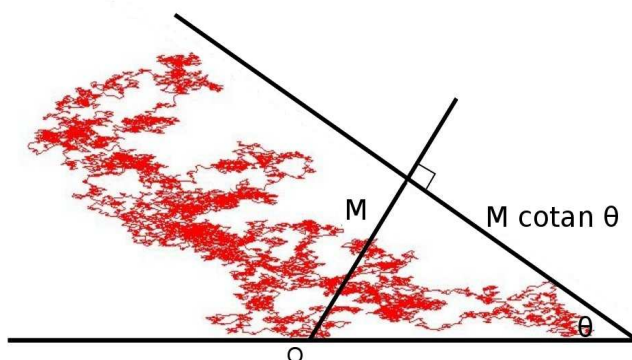


FIG. 6.2 – Chemin brownien fermé contraint à demeurer à l'intérieur d'un secteur d'ouverture θ . M correspond alors à la distance entre le point initial (et final) du mouvement et le bord opposé du secteur.

Bien sûr dans les cas limites $\theta = 0$ et $\theta = \frac{\pi}{2}$, nous connaissons déjà la fonction de répartition F_θ (pour l'avoir calculé dans la première partie, éq. (2.15) (remplacée ici par une forme équivalente [139]) et (2.21)), puisque $y(t)$ décrit dans le premier cas simplement une excursion brownienne standard et dans le second cas un pont brownien standard :

$$F_0(M) = \sqrt{2\pi} \frac{T^{\frac{3}{2}}}{M^3} \sum_{n=1}^{\infty} n^2 e^{-\frac{n^2 \pi^2 T}{2M^2}} \quad (6.3)$$

$$F_{\frac{\pi}{2}}(M) = 1 - e^{-\frac{2M^2}{T}}. \quad (6.4)$$

Pour des valeurs générales de θ , notre raisonnement s'inspire de celui mené pour des mouvements linéaires, en le transposant à un contexte bidimension-

nel. Nous faisons donc à nouveau appel à la formule de Feynman-Kac :

$$F_\theta(M) \propto \langle \vec{\varepsilon}' | e^{-\hat{H}T} | \vec{\varepsilon}' \rangle, \quad (6.5)$$

avec :

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} - \frac{1}{2} \frac{\partial^2}{\partial y^2} + V(x, y), \quad (6.6)$$

V étant un potentiel visant à assurer le confinement du mouvement à l'intérieur du secteur d'ouverture θ : il est donc nul à l'intérieur de ce secteur et infini à l'extérieur.

$\vec{\varepsilon}'$ est l'analogie ici du paramètre de coupure déjà rencontré dans l'étude de mouvements linéaires. Il assure la même fonction, à savoir pallier l'impossibilité d'imposer simultanément au mouvement de commencer sur un bord du secteur et de ne pas le franchir. $\vec{\varepsilon}'$ est donc un « petit » vecteur orienté de l'origine vers l'intérieur du secteur et pris comme point initial et final du mouvement. Nous ferons ensuite, comme dans le cas unidimensionnel, tendre $\vec{\varepsilon}'$ vers $\vec{0}$.

6.2 Cas particulier des angles $\theta = \frac{\pi}{n}$

Avant d'examiner un cas tout à fait général, étudions le cas particulier des angles θ qui sont une « fraction entière » de π , c'est-à-dire qui ont la forme $\theta = \frac{\pi}{n}$.

Remarquons tout d'abord que, dans un cas encore plus particulier, celui où $n = 4$ et donc $\theta = \frac{\pi}{4}$, $x(t)$ et $y(t)$ sont des ponts browniens unidimensionnels avec la contrainte de voir y demeurer toujours « au-dessus » de x ; autrement dit, x et y sont deux marcheurs browniens « malveillants » (cf. § 4.3). Les fonctions propres du hamiltonien \hat{H} (éq. (6.6)) s'écrivent dans ce cas particulier assez aisément, par (anti)symétrie :

$$\psi_{k_1, k_2; M} = \sin k_1(M-x) \sin k_2(M-y) - \sin k_1(M-y) \sin k_2(M-x), \quad (6.7)$$

k_1 et k_2 étant des nombres réels. La valeur propre associée est donnée par l'action du hamiltonien :

$$\hat{H}\psi_{k_1, k_2; M} = \frac{k_1^2 + k_2^2}{2}\psi_{k_1, k_2; M} \quad (6.8)$$

Évaluons $\psi_{k_1, k_2; M}$ au point $(x, y) = (\varepsilon, 0)$, avec $\varepsilon \ll 1$:

$$\psi_{k_1, k_2; M}(\varepsilon, 0) = \sin k_1(M - \varepsilon) \sin k_2 M - \sin k_1 M \sin k_2(M - \varepsilon) \quad (6.9)$$

$$\underset{\varepsilon \ll 1}{\sim} \varepsilon [\sin(k_1 M) k_2 \cos(k_2 M) - \sin(k_2 M) k_1 \cos(k_1 M)] \quad (6.10)$$

Nous pouvons écrire le propagateur (6.5) :

$$F_{\frac{\pi}{4}}(M) = \iint_{-\infty}^{\infty} [\psi_{k_1, k_2; M}(\varepsilon, 0)]^2 e^{-\frac{(k_1^2 + k_2^2)T}{2}} dk_1 dk_2. \quad (6.11)$$

Nous obtenons donc :

$$F_{\frac{\theta}{4}}(M) \propto \varepsilon^2 \iint_{-\infty}^{\infty} [(\sin(k_1 M) k_2 \cos(k_2 M) - \sin(k_2 M) k_1 \cos(k_1 M))]^2 e^{-\frac{(k_1^2 + k_2^2)T}{2}} dk_1 dk_2. \quad (6.12)$$

Le calcul peut s'effectuer par exemple en séparant les variables k_1 et k_2 , et la constante de proportionnalité est fixée, comme précédemment, par la condition de normalisation :

$$\lim_{M \rightarrow \infty} F_{\theta}(M) = 1. \quad (6.13)$$

Ceci nous conduit au résultat suivant :

$$\boxed{F_{\frac{\theta}{4}}(M) = 1 - e^{-\frac{4M^2}{T}} - \frac{4M^2}{T} e^{-\frac{2M^2}{T}}.} \quad (6.14)$$

Pour aborder le cas plus général d'angles $\theta = \frac{\pi}{n}$ avec n un entier supérieur ou égal à 3 quelconque, nous effectuons un changement de repère consistant à prendre la pointe du secteur d'ouverture θ comme origine et le bord du secteur ne contenant pas le point d'attache du mouvement comme axe des abscisses. Les nouvelles coordonnées x' et y' peuvent s'écrire en fonction des anciennes, x et y , comme suit :

$$\begin{cases} x' = M \cotan \theta - x \\ y' = M - y \end{cases} \quad (6.15)$$

Nous passons ensuite en coordonnées polaires :

$$\begin{cases} r^2 = x'^2 + y'^2 \\ x' = r \cos \phi \\ y' = r \sin \phi \end{cases} \quad (6.16)$$

et choisissons comme suit le paramètre de coupure ε :

$$\vec{\varepsilon} = \left(\frac{M}{\sin \theta}, \theta - \varepsilon \right), \quad (6.17)$$

ce qui revient à fixer le point d'ancrage du mouvement légèrement à l'intérieur du secteur, tout près de l'ancienne origine du repère.

Dans ce contexte, pour des angles $\theta = \frac{\pi}{n}$, le propagateur d'un mouvement brownien contraint à demeurer dans un secteur d'ouverture θ peut être calculé [140] à l'aide d'une élégante généralisation de la méthode des images évoquée dans la première partie (§ 2.1.1). Ainsi, la probabilité qu'un mouvement brownien plan restreint à un secteur d'ouverture $\theta = \frac{\pi}{n}$ aille d'un point \vec{r}_0 à un point \vec{r} en un temps T est donné par :

$$\begin{aligned} \langle \vec{r} | e^{-\hat{H}T} | \vec{r}_0 \rangle = \\ \frac{1}{2\pi T} e^{-\frac{r^2+r_0^2}{2T}} \sum_{m=0}^{n-1} \left\{ e^{\frac{r \cdot r_0}{T} \cos(\phi - \phi_0 - \frac{2m\pi}{n})} - e^{\frac{r \cdot r_0}{T} \cos(\phi + \phi_0 - \frac{2(m+1)\pi}{n})} \right\}. \end{aligned} \quad (6.18)$$

Dans le cas qui nous intéresse ici, les points \vec{r}_0 et \vec{r} sont identiques :

$$\begin{cases} r = r_0 = \frac{M}{\sin \frac{\pi}{n}} \\ \phi = \phi_0 = \frac{\pi}{n} - \varepsilon. \end{cases} \quad (6.19)$$

En utilisant ces relations dans la formule (6.18) donnant le propagateur, il vient :

$$\begin{aligned} \langle \vec{\varepsilon} | e^{-\hat{H}T} | \vec{\varepsilon} \rangle = \\ \frac{1}{2\pi T} e^{-\frac{r^2}{T}} \sum_{m=0}^{n-1} \left\{ e^{\frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right)} - e^{\frac{r^2}{T} \cos\left(\frac{2m\pi}{n} + 2\varepsilon\right)} \right\}. \end{aligned} \quad (6.20)$$

Après simplification et en considérant $\varepsilon \ll 1$, nous obtenons :

$$\langle \vec{\varepsilon} | e^{-\hat{H}T} | \vec{\varepsilon} \rangle \sim \frac{1}{2\pi T} e^{-\frac{r^2}{T}} \sum_{m=0}^{n-1} e^{\frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right)}. \\ \left\{ 2\varepsilon^2 \frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right) + 2\varepsilon \frac{r^2}{T} \sin\left(\frac{2m\pi}{n}\right) - 2\varepsilon^2 \frac{r^4}{T^2} \sin^2\left(\frac{2m\pi}{n}\right) \right\}. \quad (6.21)$$

Les propriétés de parité des fonctions sin et cos font que les termes $2\varepsilon \frac{r^2}{T} \sin\left(\frac{2m\pi}{n}\right) \exp\left[\frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right)\right]$ s'annulent lorsque l'on en fait la somme pour m variant de 0 à $n-1$; ceci nous conduit donc à :

$$F_{\frac{\pi}{n}}(M = r \sin \frac{\pi}{n}) \propto \\ \frac{\varepsilon^2}{\pi T} e^{-\frac{r^2}{T}} \sum_{m=0}^{n-1} e^{\frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right)} \left\{ \frac{r^2}{T} \cos\left(\frac{2m\pi}{n}\right) - \frac{r^4}{T^2} \sin^2\left(\frac{2m\pi}{n}\right) \right\}. \quad (6.22)$$

Enfin, la condition de normalisation de la fonction de répartition nous permet à nouveau de déterminer la constante de proportionnalité (égale en l'occurrence à $[\frac{\varepsilon^2 r^2}{\pi T^2}]^{-1}$). Nous en tirons notre expression finale pour la fonction de répartition de l'ordonnée maximale M dans le cas d'une inclinaison $\theta = \frac{\pi}{n}$:

$$F_{\frac{\pi}{n}}(M) = \\ e^{-\frac{M^2}{T \sin^2 \frac{\pi}{n}}} \sum_{m=0}^{n-1} e^{\frac{M^2}{T \sin^2 \frac{\pi}{n}} \cos\left(\frac{2m\pi}{n}\right)} \left\{ \cos\left(\frac{2m\pi}{n}\right) - \frac{M^2}{T \sin^2 \frac{\pi}{n}} \sin^2\left(\frac{2m\pi}{n}\right) \right\}. \quad (6.23)$$

Nous pouvons aisément vérifier que, pour $n = 4$, cette formule redonne bien le résultat obtenu précédemment par une voie différente (éq. (6.14)).

6.3 Cas général

Pour des valeurs quelconques de l'inclinaison θ de la droite au-dessus de laquelle le polymère brownien que nous étudions est confiné, le calcul de la fonction de répartition de l'ordonnée maximale se fait en suivant les mêmes lignes que dans les cas particuliers traités dans la section précédente. La différence réside en ce qu'il n'est plus possible d'utiliser la méthode des

images. Il faut donc résoudre directement l'équation aux fonctions propres correspondant au hamiltonien \hat{H} donné par (6.6). Celui-ci est simplement le hamiltonien d'une particule libre à l'intérieur du secteur d'angle θ , avec une condition de nullité pour les fonctions propres sur les bords du secteur. Les fonctions propres, agissant sur les variables polaires (r, ϕ) , peuvent donc s'écrire comme un produit de fonctions sinus et de fonctions de Bessel de première espèce J_ν , avec un paramètre k réel et un paramètre n entier :

$$\Psi_{k,n}(r, \phi) = \sin\left(\frac{n\pi}{\theta}\phi\right) J_{\frac{n\pi}{\theta}}(kr). \quad (6.24)$$

Ces fonctions satisfont en effet :

$$\begin{cases} -\frac{1}{2}\Delta\Psi_{k,n}(r, \phi) = k^2\Psi_{k,n}(r, \phi) & \text{pour } 0 < \phi < \theta \\ \Psi_{k,n}(r, \phi) = 0 & \text{pour } \phi = 0, \phi = \theta, \end{cases} \quad (6.25)$$

Δ désignant l'opérateur laplacien.

L'équivalent du propagateur (6.18) est donc ici, avec la bonne normalisation [141, 142] :

$$\begin{aligned} \langle \vec{r} | e^{-\hat{H}T} | \vec{r}_0 \rangle = \\ \frac{2}{\theta T} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi\phi}{\theta}\right) \sin\left(\frac{n\pi\phi_0}{\theta}\right) \int_0^{\infty} J_{\frac{n\pi}{\theta}}(kr) J_{\frac{n\pi}{\theta}}(kr_0) e^{-\frac{k^2 T}{2}} k dk, \end{aligned} \quad (6.26)$$

Une identité relative aux fonctions de Bessel [56] nous permet ensuite de réaliser l'intégration et d'obtenir :

$$\langle \vec{r} | e^{-\hat{H}T} | \vec{r}_0 \rangle = \frac{2}{\theta T} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi\phi}{\theta}\right) \sin\left(\frac{n\pi\phi_0}{\theta}\right) e^{-\frac{r^2+r_0^2}{2T}} I_{\frac{n\pi}{\theta}}\left(\frac{rr_0}{T}\right), \quad (6.27)$$

I_ν étant la fonction de Bessel modifiée de première espèce d'ordre ν .

À l'instar de ce que nous avons fait dans la section précédente, nous utilisons maintenant ce propagateur entre le point de départ et le point d'arrivée du chemin brownien considéré, points qui sont bien sûr identiques et donnés par :

$$\begin{cases} r = r_0 = \frac{M}{\sin\frac{\pi}{n}} \\ \phi = \phi_0 = \frac{\pi}{n} - \varepsilon. \end{cases} \quad (6.28)$$

Il vient alors :

$$F_\theta(M) \propto \frac{2}{\theta T} \sum_{k=1}^{\infty} \sin^2 \left(\frac{k\pi(\theta - \varepsilon)}{\theta} \right) e^{-\frac{M^2}{T \sin^2 \theta} I_{\frac{k\pi}{\theta}}} \left(\frac{M^2}{T \sin^2 \theta} \right) \quad (6.29)$$

$$\underset{\varepsilon \ll 1}{\propto} \frac{2\pi^2 \varepsilon^2}{\theta^3 T} e^{-\frac{M^2}{T \sin^2 \theta}} \sum_{k=1}^{\infty} k^2 I_{\frac{k\pi}{\theta}} \left(\frac{M^2}{T \sin^2 \theta} \right). \quad (6.30)$$

Il nous faut bien sûr nous assurer de la normalisation de cette fonction de répartition et, par ailleurs, nous devons pouvoir retrouver les résultats des cas particuliers précédents à partir de cette formule générale. Empruntons cette voie en fixant $\theta = \frac{\pi}{n}$ avec n entier. La forme de (6.30) requiert que nous soyons capable de calculer des sommes du type :

$$\sum_{k=1}^{\infty} k^2 I_{kn}(u), \quad (6.31)$$

I_ν désignant toujours une fonction de Bessel modifiée de première espèce, et u désignant une variable réelle. Par un calcul que nous présentons en annexe (B), nous obtenons le résultat suivant :

$$\sum_{k=1}^{\infty} k^2 I_{kn}(u) = \frac{u^2}{2n^3} \sum_{k=0}^{n-1} \left[\frac{1}{u} \cos \frac{2\pi k}{n} - \sin^2 \frac{2\pi k}{n} \right] e^{u \cos \frac{2\pi k}{n}}. \quad (6.32)$$

Il nous reste à insérer ce résultat dans la formule générale (6.30). En notant $u = \frac{M^2}{T \sin^2 \theta}$, nous aboutissons à :

$$F_{\frac{\pi}{n}} \propto \frac{\varepsilon^2}{\pi T} u e^{-u} \sum_{k=0}^{n-1} e^{u \cos \frac{2\pi k}{n}} \left[\cos \frac{2\pi k}{n} - u \sin^2 \frac{2\pi k}{n} \right] \quad (6.33)$$

Il nous suffit alors de comparer cette équation avec (6.23) pour fixer la constante de proportionnalité qui, comme précédemment, fait « disparaître » le paramètre de coupure, nous conduisant au résultat final, valable **pour tout angle θ** :

$$\boxed{F_\theta(M) = \frac{2\pi^3 T \sin^2 \theta}{M^2 \theta^3} e^{-\frac{M^2}{T \sin^2 \theta}} \sum_{k=1}^{\infty} k^2 I_{\frac{k\pi}{\theta}} \left(\frac{M^2}{T \sin^2 \theta} \right)}. \quad (6.34)$$

L'intérêt de ce résultat, à nos yeux, est double. D'une part, dans le contexte du problème énoncé initialement, à savoir celui de la hauteur maximale, selon l'axe vertical, d'un chemin brownien contraint à demeurer au-dessus d'une droite (problème pouvant correspondre dans le modèle de Rouse à un polymère demeurant au-dessus d'un substrat sous-jacent incliné), notre résultat apporte une réponse complète en donnant la distribution exacte de cette hauteur maximale. En second lieu, si l'on examine le problème tel que représenté sur la figure 6.2, alors la possibilité d'un autre point de vue se dégage : ce chemin fermé peut être une excursion brownienne dans le demi-plan supérieur, dont l'extension maximale dans la direction $\phi = \frac{\pi}{2} - \theta$ est égale à M . Par extension maximale, nous entendons la chose suivante : si nous projetons tous les points du chemin sur la droite de direction ϕ et mesurons les distances (algébriques) entre l'origine O et les projetés, alors la plus grande de ces distances (possiblement négatives) sera M . Cette grandeur, M , varie non seulement en fonction de ϕ mais aussi stochastiquement à chaque réalisation du chemin brownien — or $F_{\frac{\pi}{2}-\phi}$ telle que calculée ci-dessus donne justement la fonction de répartition de la variable aléatoire $M(\phi)$.

Nous verrons au chapitre suivant que $M(\phi)$ est ce l'on appelle la fonction support de l'ensemble considéré (en l'occurrence un chemin brownien fermé) et qu'elle peut se révéler fort utile dans l'étude de l'enveloppe convexe de l'ensemble en question. Mais avant de passer à ce point, intéressons-nous à la valeur moyenne de M , pour un angle donné, sur l'ensemble des chemins browniens fermés satisfaisant les contraintes choisies. Les moments de la distribution de M joueront en effet un rôle de premier plan dans notre étude de l'enveloppe convexe du mouvement brownien. En outre, le calcul de $\langle M(\theta) \rangle$ nous paraît assez intéressant pour figurer ici dans le corps de texte.

6.4 Valeur moyenne de M

Nous partons de la fonction de répartition $F_\theta(M)$ — donnée par l'éq. (6.34) — afin de calculer (pour une valeur de θ fixée) la valeur moyenne de la variable aléatoire M , considérée sur l'ensemble des réalisations possibles du chemin brownien représenté sur la figure 6.1.

$$\langle M \rangle = \int_0^\infty M F'_\theta(M) dM \quad (6.35)$$

$$= \int_0^\infty (1 - F_\theta(M)) dM, \quad (6.36)$$

où nous avons noté F'_θ la dérivée de F_θ et réalisé une intégration par parties. En insérant l'expression (6.34) et en posant $x = \frac{M}{\sqrt{T} \sin \theta}$ puis $y = x^2$, il vient :

$$\langle M \rangle = \int_0^\infty \sqrt{T} \sin \theta \left(1 - \frac{2\pi^3}{\theta^3 x^2} e^{-x^2} \sum_{k=1}^\infty k^2 I_{\frac{k\pi}{\theta}}(x^2) \right) dx \quad (6.37)$$

$$= \frac{\sqrt{T} \sin \theta}{2} \int_0^\infty \left(1 - \frac{2\pi^3}{\theta^3 y} e^{-y} \sum_{k=1}^\infty k^2 I_{\frac{k\pi}{\theta}}(y) \right) \frac{dy}{\sqrt{y}}. \quad (6.38)$$

Il va nous être utile, pour mener le calcul à son terme, d'introduire une fonction de coupure avec un paramètre c que nous ferons ensuite tendre vers 0 :

$$\langle M \rangle = \frac{\sqrt{T} \sin \theta}{2} \lim_{c \rightarrow 0} \int_0^\infty \left(1 - \frac{2\pi^3}{\theta^3 y} e^{-y} \sum_{k=1}^\infty k^2 I_{\frac{k\pi}{\theta}}(y) \right) e^{-c y} \frac{dy}{\sqrt{y}}. \quad (6.39)$$

Cette ré-écriture nous permet en effet de séparer les termes de l'intégrale.

En posant $1 + c = \text{ch } \alpha$, avec $0 < \alpha \ll 1$, nous avons d'une part :

$$\int_0^\infty e^{-c y} \frac{dy}{\sqrt{y}} = \sqrt{\frac{\pi}{c}} \simeq \frac{\sqrt{2\pi}}{\alpha}; \quad (6.40)$$

et d'autre part

$$\int_0^\infty y^{-\frac{3}{2}} e^{-y(1+c)} I_{\frac{k\pi}{\theta}}(y) dy = \int_0^\infty y^{-\frac{3}{2}} e^{-y \text{ch } \alpha} I_{\frac{k\pi}{\theta}}(y) dy \quad (6.41)$$

$$= -\sqrt{\frac{2}{\pi}} \text{sh } \alpha Q_{\frac{k\pi}{\theta} - \frac{1}{2}}^{-1}(\text{ch } \alpha) \quad (6.42)$$

$$= \frac{2}{\sqrt{\pi}} \int_\alpha^\infty \sqrt{\text{ch } u - \text{ch } \alpha} e^{-\frac{k\pi}{\theta} u} du \quad (6.43)$$

où Q_ν^μ est une fonction de Legendre associée de seconde espèce et où nous avons utilisé deux identités (6.622.(3) et 8.715.(2)) tirées de [56].

Nous pouvons désormais calculer la somme qui apparaît dans (6.39) :

$$\begin{aligned} \int_0^\infty y^{-\frac{3}{2}} e^{-y(1+c)} \sum_{k=1}^\infty k^2 I_{\frac{k\pi}{\theta}}(y) dy \\ = \frac{2}{\sqrt{\pi}} \int_\alpha^\infty \sqrt{\operatorname{ch} u - \operatorname{ch} \alpha} \sum_{k=1}^\infty k^2 \left(e^{-\frac{\pi}{\theta}u}\right)^k du \end{aligned} \quad (6.44)$$

$$= \frac{1}{2\sqrt{\pi}} \int_\alpha^\infty \sqrt{\operatorname{ch} u - \operatorname{ch} \alpha} \frac{\operatorname{ch} \left(\frac{u\pi}{2\theta}\right)}{\operatorname{sh}^3 \left(\frac{u\pi}{2\theta}\right)} du. \quad (6.45)$$

Nous allons maintenant étudier le comportement de (6.45) lorsque α tend vers 0. L'idée est que le terme asymptotiquement dominant de (6.45) doit venir annuler celui de (6.40) (qui est divergent) lorsque les deux sont combinés dans (6.39).

Commençons donc par le terme dominant, en posant $a = \frac{\pi}{2\theta}$ et $u = \alpha v$:

$$\int_\alpha^\infty \sqrt{\operatorname{ch} u - \operatorname{ch} \alpha} \frac{\operatorname{ch} \left(\frac{u\pi}{2\theta}\right)}{\operatorname{sh}^3 \left(\frac{u\pi}{2\theta}\right)} du = \alpha \int_1^\infty \sqrt{\operatorname{ch} \alpha v - \operatorname{ch} \alpha} \frac{\operatorname{ch} (a\alpha v)}{\operatorname{sh}^3 (a\alpha v)} dv \quad (6.46)$$

$$\simeq \alpha \int_1^\infty \frac{\sqrt{\frac{\alpha^2}{2}(v^2 - 1)}}{a^3 \alpha^3 v^3} dv \quad (6.47)$$

$$\simeq \frac{1}{\alpha a^3 \sqrt{2}} \int_1^\infty \frac{\sqrt{v^2 - 1}}{v^3} dv \quad (6.48)$$

$$\simeq \frac{\pi}{4\alpha a^3 \sqrt{2}} \quad (6.49)$$

Nous voyons d'ores et déjà que, multiplié par l'ensemble de ses préfacteurs, ce terme viendra bien annuler (6.40).

La démarche consiste maintenant à retrancher ce terme dominant, (6.49), à (6.46) pour obtenir un terme constant qui nous donnera la valeur de (6.39).

En conservant les mêmes notations :

$$\int_{\alpha}^{\infty} \left[\frac{\sqrt{\operatorname{ch} u - \operatorname{ch} \alpha}}{\operatorname{sh}^3 (au)} \operatorname{ch} (au) - \frac{\sqrt{u^2 - \alpha^2}}{\sqrt{2}a^3u^3} \right] du$$

$$\underset{\alpha \rightarrow 0, 0 < \delta \ll 1}{\simeq} \int_{\delta}^{\infty} \left[\frac{\sqrt{\operatorname{ch} u - 1}}{\operatorname{sh}^3 (au)} \operatorname{ch} (au) - \frac{1}{\sqrt{2}a^3u^2} \right] du \quad (6.50)$$

$$\simeq \int_{\delta}^{\infty} \frac{\sqrt{\operatorname{ch} u - 1}}{2a} \frac{d}{du} \left[-\frac{1}{\operatorname{sh}^2 (au)} \right] du - \frac{1}{\sqrt{2}a^3\delta} \quad (6.51)$$

$$\simeq \frac{1}{2\sqrt{2}a^3\delta} + \frac{1}{2a} \int_{\delta}^{\infty} \frac{\operatorname{sh} u}{2\sqrt{\operatorname{ch} u - 1}} \frac{du}{\operatorname{sh}^2 (au)} - \frac{1}{\sqrt{2}a^3\delta} \quad (6.52)$$

$$\simeq -\frac{1}{2\sqrt{2}a^3\delta} + \frac{1}{2\sqrt{2}a} \int_{\delta}^{\infty} \frac{\operatorname{ch} \frac{u}{2}}{\operatorname{sh}^2 (au)} du \quad (6.53)$$

$$\simeq -\frac{1}{2\sqrt{2}a^3\delta} + \frac{1}{2\sqrt{2}a} \int_{\delta}^{\infty} \frac{\operatorname{ch} \frac{u}{2} - 1 + 1}{\operatorname{sh}^2 (au)} du \quad (6.54)$$

Nous séparons ensuite les deux termes de l'intégrale :

$$\simeq -\frac{1}{2\sqrt{2}a^3\delta} + \frac{1}{2\sqrt{2}a} \left\{ \left[-\frac{\operatorname{coth} (au)}{a} \right]_{\delta}^{\infty} + \int_{\delta}^{\infty} \frac{\operatorname{ch} \frac{u}{2} - 1}{\operatorname{sh}^2 au} du \right\} \quad (6.55)$$

$$\simeq -\frac{1}{2\sqrt{2}a^3\delta} + \frac{1}{2\sqrt{2}a} \left\{ -\frac{1}{a} \left(1 - \frac{1}{a\delta} \right) + \int_0^{\infty} \frac{2\operatorname{sh}^2 \frac{u}{4}}{\operatorname{sh}^2 au} du \right\} \quad (6.56)$$

Nous avons changé la borne inférieure de la dernière intégrale en 0 puisque celle-ci converge désormais dans la limite $\delta \rightarrow 0$. Les termes en $\frac{1}{\delta}$ s'annulent

et l'intégrale est donnée dans [56] (3.511.9). Nous obtenons donc :

$$\int_{\alpha}^{\infty} \left[\frac{\sqrt{\operatorname{ch} u - \operatorname{ch} \alpha}}{\operatorname{sh}^3 (au)} \operatorname{ch} (au) - \frac{\sqrt{u^2 - \alpha^2}}{\sqrt{2}a^3u^3} \right] du$$

$$\simeq -\frac{1}{2\sqrt{2}a^2} + \frac{1}{2\sqrt{2}a^2} \left(1 - \frac{\pi}{4a} \operatorname{cotan} \left(\frac{\pi}{4a} \right) \right) \quad (6.57)$$

$$\simeq -\frac{\pi}{8\sqrt{2}a^3} \frac{1 + \cos \frac{\pi}{2a}}{\sin \left(\frac{\pi}{2a} \right)}. \quad (6.58)$$

Il ne nous reste plus qu'à combiner (6.58) avec (6.49) puis à insérer le résultat obtenu, ainsi que (6.40), dans (6.39) en redonnant aux notations leur sens en termes des variables initiales pour aboutir au résultat final :

$$\boxed{\langle M \rangle = \frac{1 + \cos \theta}{2} \sqrt{\frac{\pi T}{2}}} \quad (6.59)$$

T étant, rappelons-le, la durée du mouvement brownien considéré et θ l'inclinaison de la droite déterminant le demi-plan dans lequel le mouvement est confiné.

Sans détailler le calcul, qui suit exactement les mêmes lignes, nous donnons également le second moment de M :

$$\boxed{\langle M^2 \rangle = \frac{\pi^2 \sin^2 \theta}{6\theta^2} T} \quad (6.60)$$

Notons les formes compactes et simples de ces valeurs moyennes, que ne laissait peut-être pas espérer la fonction de répartition (6.34). Notons aussi, en traduisant l'équation (6.59) dans le langage correspondant au second point de vue évoqué à la fin de la section précédente, que, pour une excursion dans le demi-plan supérieur, la valeur moyenne de la projection maximale du chemin brownien sur la droite de direction ϕ est donnée par :

$$\langle M(\phi) \rangle = \frac{1 + \sin \phi}{2} \sqrt{\frac{\pi T}{2}}. \quad (6.61)$$

Nous verrons plus loin tout l'intérêt de ce résultat, notamment son interprétation en termes d'enveloppe convexe de la trajectoire. Ceci nous sera permis par les propriétés de M vue comme une fonction associée à la trajectoire : ce sont quelques propriétés de cette fonction dite « support » que nous allons préciser dans le chapitre suivant.

Chapitre 7

Fonction support et formules de Cauchy : une approche globale des enveloppes convexes aléatoires

Les méthodes employées dans les différents travaux mentionnés dans notre aperçu des résultats connus sont diverses et habituellement spécifiques à un contexte donné : points indépendants d'une part, points corrélés d'autre part, marche aléatoire (en temps discret) d'une part, mouvement brownien (temps continu) d'autre part. Nous développons ici une méthode générale présentant l'avantage de fonctionner dans l'ensemble de ces contextes.

Le point de départ de notre approche s'inspire des points de vues et remarques de Takács [96], Eddy [94] et El Bachir [99], reprenant en particulier l'idée d'utiliser les formules de Cauchy. Leur application, en vue d'obtenir des résultats explicites nouveaux, nécessite une systématisation et le recours à des techniques nouvelles que nous détaillons dans ce qui suit. Nous verrons en particulier apparaître des liens explicites entre enveloppes convexes aléatoires et statistiques des valeurs extrêmes dans le cas de points indépendants, et entre enveloppes convexes aléatoires et temps d'atteinte du maximum dans le cas de chemins browniens.

La méthode présentée ici se fonde donc sur l'idée que la connaissance

de l'extension spatiale d'un objet dans toutes les directions θ devrait permettre d'accéder aux propriétés géométriques de l'enveloppe convexe de l'objet. Mathématiquement, cette idée se traduit par les formules dites de Cauchy, Cauchy-Crofton ou encore Cauchy-Barbier [143–148]. Ces formules, comme nous le verrons, permettent de calculer le périmètre d'une courbe plane et l'aire enclose par celle-ci, au moyen d'intégrales portant sur les projections de la courbe selon toutes les directions du plan. Ces projections sont décrites par la fonction support de l'objet.

7.1 Notion de fonction support

Dans \mathbb{R}^2 , la fonction support $M(\theta)$ d'un ensemble $C = \{(x_i, y_i), i \in I\}$ est définie par :

$$M(\theta) = \max_{i \in I} \{x_i \cos \theta + y_i \sin \theta\}. \quad (7.1)$$

Physiquement, on peut la visualiser comme suit : imaginons que l'on sonde

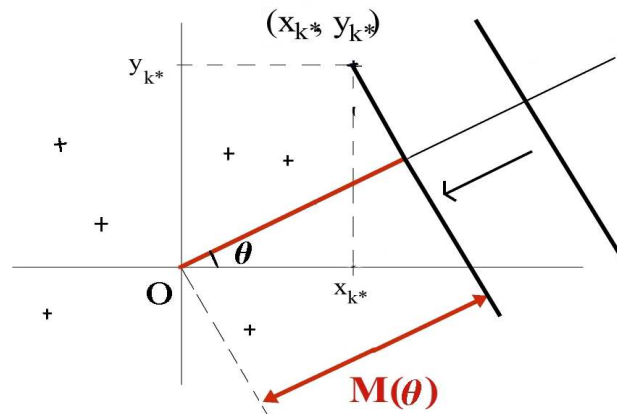


FIG. 7.1 – Fonction support d'un nuage de points

l'extension spatiale d'un objet bidimensionnel dans toutes les directions du plan ; pour ce faire, étant donné un angle $\theta \in [0, 2\pi]$, on approche une droite (qui pourrait être un faisceau laser ou une baguette) perpendiculairement à la direction θ , depuis l'infini jusqu'au contact avec l'objet. La distance

(algébrique — elle peut être négative) entre la ligne droite et l'origine O est, par définition, $M(\theta)$. La figure 7.1 illustre ce point, dans le cas où l'objet « sondé » est un nuage de points.

Notons que, par définition de $M(\theta)$, pour tout θ fixé il existera un point (x_{k^*}, y_{k^*}) du nuage tel que :

$$M(\theta) = x_{k^*} \cos \theta + y_{k^*} \sin \theta \quad (7.2)$$

Remarquons qu'à partir de (7.2), nous pouvons exprimer simplement la dérivée de M :

$$M'(\theta) = -x_{k^*} \sin \theta + y_{k^*} \cos \theta \quad (7.3)$$

En d'autres termes, $M'(\theta)$ est la distance (algébrique) entre le point de l'ensemble donnant la projection maximale $M(\theta)$ et la droite de direction θ , comme on peut le voir sur la figure 7.2.

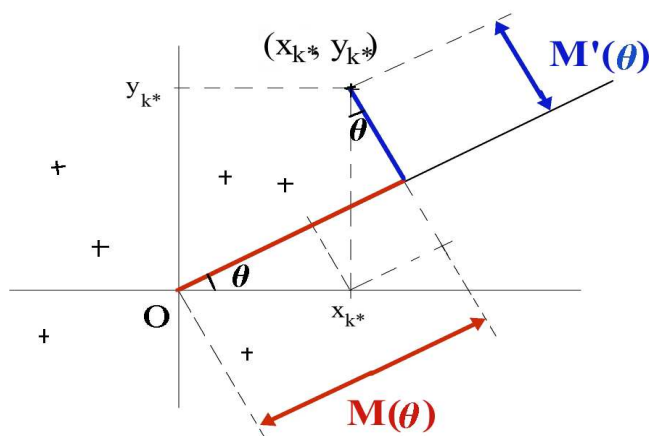


FIG. 7.2 – Dérivée de la fonction support d'un ensemble

Pour nous familiariser avec le concept de fonction support, examinons un exemple simple : celui d'un triangle (Fig. 7.3) dont nous représentons graphiquement la fonction support $M(\theta)$ et sa dérivée $M'(\theta)$ sur les figures 7.4(a) et 7.4(b).

$M(\theta)$ est, bien sûr, 2π -périodique. Surtout, l'on remarque la présence de points anguleux, correspondant à des discontinuités de la dérivée de $M(\theta)$. Ainsi, $M(\theta)$ apparaît C^1 (et même C^∞) par morceaux, et sa dérivée présente

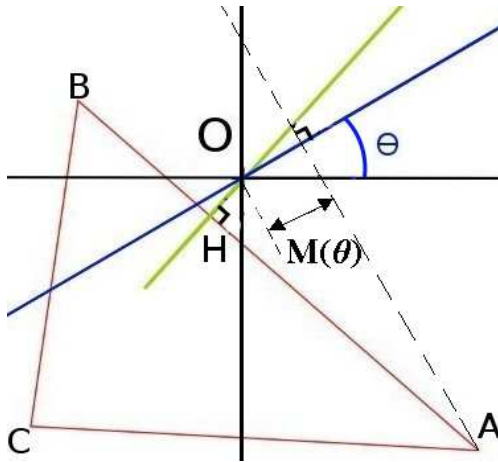


FIG. 7.3 – Triangle ABC (en rouge), avec la droite de direction θ (en bleu) et la perpendiculaire à AB passant par O (en vert).

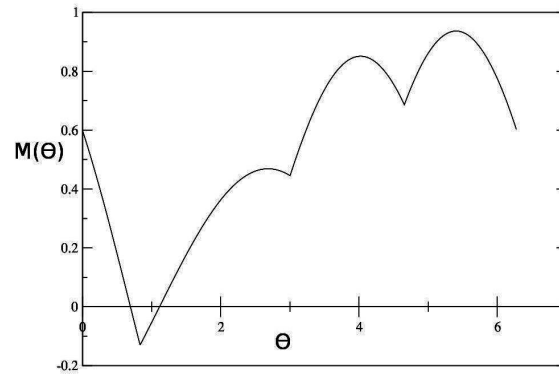
un nombre fini de discontinuités. Ce nombre est, dans ce cas particulier, égal à 3, le nombre de points de l'ensemble convexe dont $M(\theta)$ est la fonction support. Ceci n'est pas un hasard et la coïncidence de ces deux nombres se comprend en revenant aux équations (7.2) et (7.3) : sur un certain intervalle de valeurs de θ , l'un des sommets du triangle ABC donnera la projection maximale sur la direction θ et déterminera donc la valeur de $M(\theta)$, par exemple :

$$M(\theta) = x_A \cos \theta + y_A \sin \theta, \quad (7.4)$$

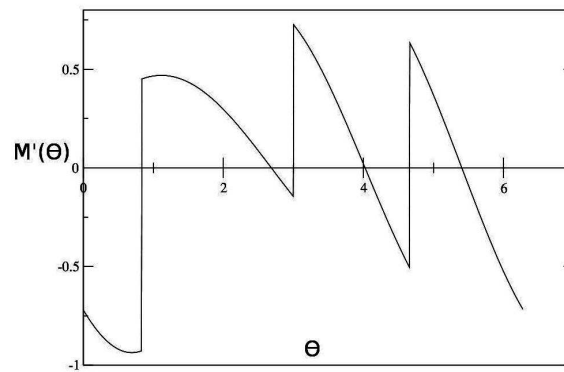
Dès lors, sur le même intervalle de valeurs θ :

$$M'(\theta) = -x_A \sin \theta + y_A \cos \theta. \quad (7.5)$$

Nous pouvons spécifier l'intervalle de valeurs de θ sur lequel ces deux équations (7.4 et 7.5) sont valides. Remarquons en effet que par définition de $M(\theta)$, lorsque θ correspond à la droite passant par l'origine O perpendiculaire au côté AB , A et B ont la même projection sur la direction θ (Fig. 7.3). Dans la limite θ^- , c'est-à-dire pour des angles infinitésimalement plus petit que θ , la valeur de M sera donnée par la projection de A , et la valeur de $|M'(\theta)|$ par la longueur du segment $[AH]$ (H étant le pied de la perpendiculaire à AB passant par O). Dans la limite θ^+ , pour des angles très légèrement supérieurs



(a)



(b)

FIG. 7.4 – (a) Graphe de la fonction support $M(\theta)$ du triangle ABC et (b) de sa dérivée $M'(\theta)$

à θ , la valeur de $M(\theta)$ sera toujours donnée par la projection commune de A et B mais $|M'(\theta)|$ sera cette fois donnée par la distance BH . Ainsi, lorsque θ correspond à la direction de la perpendiculaire à AB passant par O , la fonction support M est continue, mais sa dérivée présente une discontinuité.

Nous comprenons donc maintenant complètement le graphe 7.4. En effet, si l'on se reporte à la figure 7.3, et que l'on commence en $\theta = 0$, on voit que le point A donne alors la projection maximale sur la direction θ , et que la longueur (algébrique) de cette projection diminue quand θ croît, jusqu'à ce que la direction donnée par θ coïncide avec la droite (OH) , c'est-à-dire la perpendiculaire au côté AB passant par l'origine. À ce stade, comme nous

venons de le voir, $M(\theta)$ présente un point de « rebroussement »¹ : B va alors donner la projection maximale, dont la valeur va augmenter jusqu'à ce que θ corresponde à la droite (OB) où $M(\theta)$ atteindra un maximum local avant de diminuer jusqu'à ce que θ coïncide avec la perpendiculaire au côté BC passant par O , où $M(\theta)$ présente un second point de rebroussement ; et ainsi de suite².

La fonction support M peut être définie de manière analogue pour des ensembles non discrets de points, et elle apparaît dans la version des formules de Cauchy que nous utiliserons à la fois pour des points indépendants et pour des mouvements browniens. Nous donnerons le détail de ces formules au paragraphe suivant ; avant de clore celui-ci, introduisons un léger changement de perspective quant à la définition de la fonction support d'un ensemble de points indexés par i . Pour tout i et tout θ , définissons :

$$z_i(\theta) = x_i \cos \theta + y_i \sin \theta \quad (7.6)$$

$$h_i(\theta) = -x_i \sin \theta + y_i \cos \theta. \quad (7.7)$$

z_i est simplement la projection du i -ième point de l'échantillon sur la direction θ et h_i sa projection sur la direction perpendiculaire à θ . Par définition (éq. 7.1) :

$$M(\theta) = \max_i \{z_i(\theta)\} \equiv z_{k^*}(\theta) \quad (7.8)$$

¹Nous utilisons ici par abus de langage l'expression « point de rebroussement » pour désigner un cas particulier de point anguleux où la dérivée présente une discontinuité et change de signe.

²Dans l'exemple choisi ici, tous les sommets du triangle sont « visibles » à travers un maximum local de la fonction $M(\theta)$. Cependant, ce n'est pas toujours le cas. Il suffit pour s'en convaincre de considérer une configuration dans laquelle le point H , tout en étant par définition sur la droite (AB) , n'appartient pas au segment $[AB]$. Si H est à l'extérieur du côté de A , A passera « inaperçu ». En revanche, la coïncidence de la direction θ avec la droite (OH) donnera, elle, toujours une discontinuité dans la dérivée $M'(\theta)$ (mais sans changement de signe), ce qui se traduira par un point anguleux (mais non de rebroussement) pour $M(\theta)$. Ainsi, les points anguleux de $M(\theta)$ comptent-ils bien le nombre de côtés (et donc, en dimension 2, de sommets) de l'enveloppe convexe du nuage de points considérés. Les maximums locaux de $M(\theta)$ comptent pour leur part uniquement les sommets E tels que la projection maximale sur la droite (OE) est donnée par E — on pourrait qualifier ces sommets d'« extrémaux » ou d'« auto-extrêmes »

pour un certain indice k^* .

On a alors :

$$M'(\theta) = h_{k^*}(\theta) \quad (7.9)$$

Nous verrons plus loin l'intérêt de ce point de vue.

7.2 Les formules de Cauchy

Il existe différentes versions des formules de Cauchy. Nous utilisons ici celles qui, moyennant connaissance de la fonction support $M(\theta)$ d'une courbe convexe fermée et de sa dérivée, permettent d'accéder au périmètre L de la courbe et à l'aire A délimitée par celle-ci :

$$L = \int_0^{2\pi} M(\theta) d\theta \quad (7.10)$$

$$A = \frac{1}{2} \int_0^{2\pi} \left(M^2(\theta) - (M'(\theta))^2 \right) d\theta \quad (7.11)$$

Ces formules sont assez simples à établir pour une courbe polygonale convexe, et le passage à la limite continue donne le résultat désiré pour les courbes lisses (*cf.* § C et [146–148]).

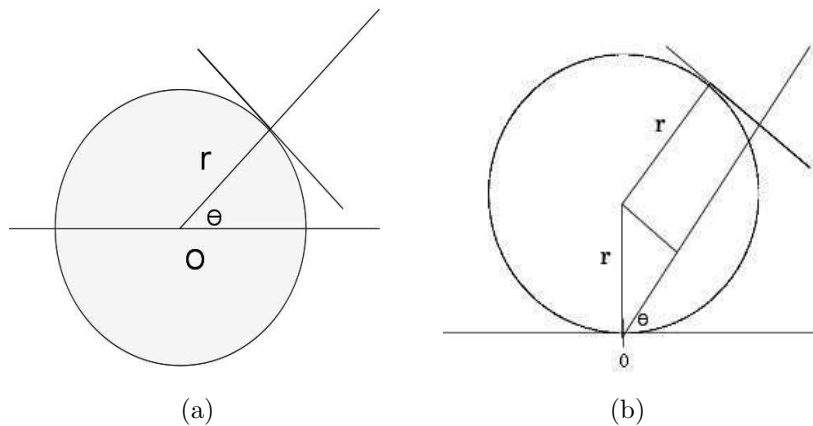


FIG. 7.5 – Deux exemples simples pour les formules de Cauchy : cercle centré sur l'origine, et cercle « posé » sur l'origine.

Dans le cas élémentaire d'un cercle centré sur l'origine de rayon r (Fig. 7.5(a)), $M(\theta)$ est constante et égale à r pour tout θ . Sa dérivée est nulle et les formules de Cauchy donnent bien les résultats habituels. Le second exemple est légèrement moins trivial (Fig. 7.5(b)), puisque $M(\theta)$ n'y est plus constante mais égale³ à $r(1 + \sin \theta)$. On retrouve bien sûr à nouveau les résultats habituels :

$$L = \int_0^{2\pi} d\theta r(1 + \sin \theta) = 2\pi r$$

$$A = \frac{1}{2} \int_0^{2\pi} d\theta r^2 [(1 + \sin \theta)^2 - \cos^2 \theta] = \pi r^2.$$

Que se passe-t-il si l'on applique les formules de Cauchy à une courbe non convexe, voire à un nuage de points ? À notre connaissance, E. Barbier [145] fut le premier à envisager que les formules de Cauchy, sous la forme que nous utilisons ici (forme d'ailleurs parfois dite de Cauchy-Barbier), donnent alors les grandeurs relatives à l'enveloppe convexe de la courbe ou du nuage. Il est possible de démontrer ce point, en observant que l'enveloppe convexe d'un objet présente la même fonction support que l'objet lui-même.

7.3 Les formules de Cauchy appliquées à un échantillon aléatoire

Examinons donc maintenant la manière dont les formules de Cauchy s'appliquent à des ensembles aléatoires. Considérons un ensemble de n points choisis au hasard, par exemple n points distribués uniformément à l'intérieur du disque unité. Pour chaque réalisation, on peut déterminer la fonction support de l'enveloppe convexe du nuage de points obtenu et utiliser les formules de Cauchy afin de calculer le périmètre et l'aire de cette enveloppe convexe. Mais, si l'on envisage l'ensemble des réalisations possibles, $M(\theta)$ et $M'(\theta)$ sont, pour θ fixé, des variables aléatoires, dont les valeurs diffèrent d'une

³Remarquons que $M(\theta)$ a ici la même forme que $\langle M(\theta) \rangle$ dans le cas d'une excursion brownienne dans le demi-plan supérieur (éq. 6.61).

réalisation à une autre. Il est dès lors possible de passer à la moyenne sur l'ensemble des réalisations dans les formules de Cauchy (7.10, 7.11) :

$$\langle L \rangle = \int_0^{2\pi} d\theta \langle M(\theta) \rangle \quad (7.12)$$

$$\langle A \rangle = \frac{1}{2} \int_0^{2\pi} d\theta \left(\langle M^2(\theta) \rangle - \langle (M'(\theta))^2 \rangle \right) \quad (7.13)$$

[$\langle \cdot \rangle$ désigne la moyenne sur l'ensemble des réalisations, et nous supposons que cette opération commute avec l'intégration sur θ]

Si nous supposons que les points de l'échantillon sont indexés par i , si nous définissons les processus aléatoires $z_i(\theta)$ et $h_i(\theta)$ comme ci-dessus (eqs. 7.6, 7.7) et si nous notons :

- μ_θ la densité de probabilité du maximum des $z_i(\theta)$,
- ρ_θ la densité de probabilité de l'indice k^* tel que $M(\theta) = \max_i \{z_i(\theta)\} \equiv z_{k^*}(\theta)$,
- et $\sigma_{i,\theta}$ la densité de probabilité de $h_i(\theta)$,

alors :

$$\langle M(\theta) \rangle = \int_{-\infty}^{\infty} z \mu_\theta(z) dz \quad (7.14)$$

$$\langle M^2(\theta) \rangle = \int_{-\infty}^{\infty} z^2 \mu_\theta(z) dz \quad (7.15)$$

$$\begin{aligned} \langle [M'(\theta)]^2 \rangle &= \int_I \int_{-\infty}^{\infty} h^2 \rho_\theta(k) \sigma_{k,\theta}(h) dh dk \\ &= \int_I \rho_\theta(k) \langle h_k^2(\theta) \rangle dk \end{aligned} \quad (7.16)$$

Dans cette formulation, il apparaît explicitement que les enveloppes convexes aléatoires sont directement liées à la statistique des valeurs extrêmes, c'est-à-dire l'étude des extréma d'échantillons aléatoires. En effet, lorsque I est fini et que les n points indexés par $i \in I$ sont choisis indépendamment selon une distribution de probabilité identique, μ_θ correspond à la distribution du maximum de n variables aléatoires réelles (en l'occurrence les $z_i(\theta)$) — un exemple classique des statistiques de valeurs extrêmes [13, 14, 149, 150].

Dans le cas où les points de l'échantillon sont les sommets d'une marche aléatoire bidimensionnelle à N pas, les $z_i(\theta)$ peuvent être vus, à θ fixé, comme les sommets d'une marche aléatoire unidimensionnelle à N pas, et μ_θ est alors la distribution du maximum d'une telle marche. Notons que dans ce cas, ρ_θ est la distribution du « temps » (du pas) auquel la marche aléatoire unidimensionnelle $z_i(\theta)$ atteint son maximum [151–153].

Il est également possible de considérer des cas dans lesquels I n'est pas un ensemble fini ni même discret : par exemple l'ensemble de points aléatoires en question peut être les positions $\mathcal{B}(\tau) = (x(\tau), y(\tau))$ d'un mouvement brownien plan aux temps $\tau \in I = [0, T]$. Dans ce cas, $z_\tau(\theta)$ et $h_\tau(\theta)$ sont tous deux des mouvements browniens unidimensionnels, et μ_θ est la distribution du maximum d'un mouvement brownien linéaire sur $[0, T]$, ρ_θ est la distribution du temps auquel un mouvement brownien linéaire atteint son maximum sur $[0, T]$ (donnée par la loi de l'arcsinus de Lévy [40], comme nous l'avons vu au chapitre 2), et $\sigma_{\tau, \theta}$ le propagateur du mouvement brownien linéaire entre 0 et τ (c'est-à-dire la distribution de la position du brownien à l'instant τ).

Nous montrerons d'abord comment l'approche basée sur la fonction support permet d'obtenir relativement aisément les résultats relatifs à l'enveloppe convexe de nuages de points indépendants dans \mathbb{R}^2 , puis nous la mettrons en œuvre dans le contexte de mouvements browniens multiples, où nous présenterons des résultats nouveaux.

Chapitre 8

Points indépendants

8.1 Cas général

Considérons ici un échantillon de N points choisis indépendamment selon une distribution bivariable identique pour tous :

$$\text{Prob}(x_i \in [x, x + dx], y_i \in [y, y + dy]) = p(x, y) dx dy,$$

Suivant le raisonnement indiqué à la section précédente (éqs. (7.6), (7.7)), nous posons :

$$z_i(\theta) = x_i \cos \theta + y_i \sin \theta,$$

and :

$$h_i(\theta) = -x_i \sin \theta + y_i \cos \theta,$$

Si nous nous souvenons des équations (7.14) et (7.12), il nous faut d'abord déterminer la distribution μ_θ du maximum des $z_i(\theta)$. Une hypothèse simplificatrice de ce point de vue est celle dans laquelle la distribution des points est isotrope, ce qui entraîne

$$\forall \theta, \quad \mu_\theta = \mu_0 \equiv \mu.$$

Nous nous plaçons dans ce cas pour ce qui suit.

8.2 Cas isotropes

Soient $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ N points du plans choisis de manière indépendante selon une distribution bivariee à symétrie sphérique $p(x, y)$. Dans un cas isotrope comme celui-ci, la distribution de la fonction support $M(\theta)$ ne dépend pas de θ et $M_n \equiv M(0)$ contient toute l'information dont nous avons besoin, comme nous venons de le voir. Les variables aléatoires $z_i(0)$ et $h_i(0)$ sont, respectivement, les abscisses x_i et les ordonnées y_i des points de l'échantillon. En combinant (7.12) et (7.14), nous pouvons écrire la valeur moyenne du périmètre de l'enveloppe convexe comme suit :

$$\langle L_N \rangle = 2\pi \langle \max_i \{x_i\} \rangle \equiv 2\pi \langle M_N \rangle \quad (8.1)$$

En posant :

$$F_N(M) = \text{Prob}[M_N \leq M] \quad (8.2)$$

nous avons :

$$F_N(M) = \left[\int_{-\infty}^M p_X(x) dx \right]^N, \quad (8.3)$$

où $p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy$ est la marginale sur sa première variable de la distribution de l'échantillon. Ainsi, pour un cas isotrope général :

$$\begin{aligned} \langle M_N \rangle &= \int_{-\infty}^{\infty} M F'_N(M) dM \\ \therefore \langle L_N \rangle &= 2\pi N \int_{-\infty}^{\infty} M p_X(M) \left[\int_{-\infty}^M p_X(x) dx \right]^{N-1} dM \\ \therefore \langle L_N \rangle &= 2\pi N \int_{-\infty}^{\infty} M p_X(M) F_{N-1}(M) dM \end{aligned} \quad (8.4)$$

Quant à l'aire moyenne de l'enveloppe convexe dans les cas isotropes, nous pouvons l'exprimer sous la forme suivante (éqs. (7.13), (7.15) et (7.16)) :

$$\langle A_N \rangle = \pi \langle M_N^2 \rangle - \pi \langle y_{k^*}^2 \rangle, \quad (8.5)$$

où y_{k^*} est l'ordonnée du point (x_{k^*}, y_{k^*}) ayant la plus grande abscisse, c'est-à-dire satisfaisant :

$$x_{k^*} = \max_i \{x_i\} = M_N.$$

Nous pouvons aisément exprimer le second moment de M_n qui apparaît dans (8.5) :

$$\langle M_N^2 \rangle = \int_{-\infty}^{\infty} M^2 F'_N(M) dM \quad (8.6)$$

$$= n \int_{-\infty}^{\infty} M^2 p_X(M) F_{N-1}(M) dM. \quad (8.7)$$

Pour calculer le second terme qui apparaît dans (8.5), à savoir, le second moment de l'ordonnée du point ayant la plus grande abscisse, nous calculons d'abord la densité de probabilité \hat{p} de ce point, définie par :

$$\text{Prob} \{(x_{k^*}, y_{k^*}) \in [(x, y), (x + dx, y + dy)]\} = \hat{p}(x, y) dx dy \quad (8.8)$$

(Notons que $F_N(M) = \int \hat{p}(M, y) dy$.)

Il n'est pas difficile de voir que $\hat{p}(x_{k^*}, y_{k^*})$ peut s'exprimer comme la densité de probabilité qu'un des N points ait pour coordonnées (x_{k^*}, y_{k^*}) et que les $N - 1$ points restants aient des abscisses inférieures à x^* :

$$\hat{p}(x_{k^*}, y_{k^*}) = N p(x_{k^*}, y_{k^*}) \left[\int_{-\infty}^{x_{k^*}} p_X(x) dx \right]^{N-1} \quad (8.9)$$

D'où :

$$\langle y_{k^*}^2 \rangle = N \iint_{-\infty}^{\infty} y_{k^*}^2 p(x_{k^*}, y_{k^*}) F_{N-1}(x_{k^*}) dx_{k^*} dy_{k^*} \quad (8.10)$$

Il suffit ensuite d'insérer (8.7) et (8.10) dans (8.5) pour obtenir une expression générale de l'aire moyenne de l'enveloppe convexe de n points choisis indépendamment selon une distribution bivariable isotrope p dont la marginale sur la première variable est donnée par p_X :

$$\begin{aligned} \langle A_N \rangle = N \pi \int_{-\infty}^{\infty} u^2 p_X(u) F_{N-1}(u) du \\ - N \pi \iint_{-\infty}^{\infty} v^2 p(u, v) F_{N-1}(u) du dv \quad (8.11) \end{aligned}$$

Nous examinons tout de suite un exemple.

Exemple d'une distribution gaussienne

Dans le cas particulier où p est une distribution gaussienne :

$$p(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}, \quad (8.12)$$

nous avons :

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \equiv \phi(x) \quad (8.13)$$

et :

$$\int_{-\infty}^x p_X(x') dx' = \int_{-\infty}^x \phi(x') dx' \equiv \Phi(x) \quad (8.14)$$

En insérant ces expressions dans (8.4) et (8.11), et en intégrant par parties, nous obtenons :

$$\langle L_N \rangle = 4 \pi \binom{N}{2} \int_{-\infty}^{\infty} \Phi^{N-2}(x) \phi^2(x) dx \quad (8.15)$$

$$\langle A_N \rangle = 3 \pi \binom{N}{3} \int_{-\infty}^{\infty} \Phi^{N-3}(x) \phi^3(x) dx \quad (8.16)$$

Ces résultats sont identiques à ceux obtenus par Efron [135] à l'aide d'une méthode différente.

8.3 Comportement asymptotique du périmètre et de l'aire moyens

Du fait des équations (7.14-7.16), le comportement asymptotique du périmètre et de l'aire moyens de l'enveloppe convexe de points indépendants distribués de manière identique et isotrope, sera lié, *via* les formules de Cauchy, à celui du maximum de leurs abscisses, qui ne sont rien d'autre que n variables aléatoires réelles, indépendantes et identiquement distribuées.

L'étude du comportement asymptotique du maximum de telles variables est aujourd'hui connue sous le nom de statistique des valeurs extrêmes, comme nous le mentionnions plus haut. Ses grandes lignes, qui ont été établies dès la première moitié et le milieu du vingtième siècle [11–14, 149, 154] et qui continuent de trouver nombre d'applications en physique [15, 16], peuvent être résumées ainsi :

Éléments de statistique des valeurs extrêmes

Soient z_1, z_2, \dots, z_N des variables aléatoires réelles, indépendantes et de densité de probabilité commune $p(z)$, et soit $M_N = \max_{\kappa=1..N} \{z_\kappa\}$ leur maximum. Alors :

$$F_N(M) = \text{Prob}(M_N \leq M) = \left[\int_{-\infty}^M p(z) dz \right]^N.$$

Dans la limite où N devient très grand, la fonction de répartition $F_N(M)$ du maximum M_N présente un des trois comportements suivants, selon la forme des « queues » de la distribution initiale p :

1. « Des queues exponentielles conduisent à une loi de type Gumbel » :

$$p(z) \underset{z \rightarrow \infty}{\sim} A e^{-z^\alpha} \rightarrow F_N(M) \underset{N \rightarrow \infty}{\sim} e^{-e^{-(M^\alpha - \log N)}}$$

2. « Des queues en loi de puissance conduisent à une loi de type Fréchet » :

$$p(z) \underset{z \rightarrow \infty}{\sim} A z^{-(\alpha+1)} \rightarrow F_N(M) \underset{N \rightarrow \infty}{\sim} e^{-\frac{A}{\alpha} N M^{-\alpha}}$$

3. « Des queues tronquées (c-à-d une loi à support fini, par exemple de rayon a) conduisent à une loi de type Weibull (de paramètre a) » :

$$p(z) \underset{z \rightarrow a}{\sim} A (a - z)^{\alpha-1} \rightarrow F_N(M) \underset{N \rightarrow \infty}{\sim} e^{-\frac{A}{\alpha} N (a - M)^\alpha}$$

Un point intéressant que peut révéler l'intuition et que confirment ces formules est que, dans les trois cas, lorsque N devient grand, la valeur du maximum croît¹ : la multitude des points tend à pousser leur maximum vers la droite sur l'axe réel.

¹En l'occurrence comme une puissance de $\ln N$ dans le premier cas, comme une puissance de N dans le second, et en s'approchant en puissance inverse de N du rayon a de l'intervalle dans le troisième.

Comme l'on peut s'y attendre, *via* les formules de Cauchy, les classes d'universalité qui apparaissent dans l'étude classique des valeurs extrêmes se traduisent dans les comportements asymptotiques du périmètre et de l'aire moyens de l'enveloppe convexe de n points du plan, indépendants et identiquement distribués. Notre méthode met donc justement ce lien en exergue, c'est pourquoi nous avons souhaité, avant d'aborder le mouvement brownien plan, présenter les résultats relatifs à des nuages de points indépendants. Toutefois, pour ne pas surcharger le texte, nous ne détaillons qu'une partie des calculs, par quelques exemples, avant de donner un résumé des résultats que notre méthode permet d'obtenir relativement aisément.

Exemple 1 : périmètre moyen dans le cas « Weibull »

Voyons donc ce qu'il advient dans le cas de N points choisis indépendamment à l'intérieur d'un cercle de rayon a suivant une distribution ayant des queues de type Weibull :

$$p(x, y) \underset{\sqrt{x^2+y^2} \rightarrow a}{\sim} A (a - \sqrt{x^2 + y^2})^{\gamma-1}, \quad (8.17)$$

A assurant la normalisation de la distribution :

$$\iint p(x, y) dx dy = 1.$$

En notant comme précédemment F_N la fonction de répartition du maximum M_n des abscisses, une intégration par parties donne :

$$\begin{aligned} \langle M_N \rangle &= \int_{-a}^a x F'_N(x) dx \\ &= a - \int_{-a}^a F_N(x) dx \end{aligned} \quad (8.18)$$

Nous nous concentrons sur le second terme de (8.18) et écrivons :

$$I_N = \int_{-a}^a F_N(x) dx.$$

Alors :

$$I_N = \int_{-a}^a \left[1 - \int_x^a p_X(x') dx' \right]^N dx \quad (8.19)$$

$$= \int_{-a}^a \exp \left[N \ln \left(1 - \int_x^a p_X(x') dx' \right) \right] dx, \quad (8.20)$$

où comme précédemment nous notons $p_X(x) = \int p(x, y) dy$.

Nous choisissons maintenant $0 < \delta \ll 1$ tel que

$$\text{pour } (a - \delta) < x < a,$$

$$p(x, y) \simeq A (a - \sqrt{x^2 + y^2})^{\gamma-1}.$$

L'idée étant que lorsque N devient grand, des points de l'échantillon s'approcheront toujours plus du bord du support (le cercle de rayon a) et, par conséquent, l'on peut se concentrer sur les « queues » de la distribution. Nous écrivons donc :

$$I_N = I_N^{(1)} + I_N^{(2)} \quad (8.21)$$

avec :

$$I_N^{(1)} = \int_{-a}^{a-\delta} \exp \left[N \ln \left(1 - \int_x^a p_X(x') dx' \right) \right] dx \quad (8.22)$$

et

$$I_N^{(2)} = \int_{a-\delta}^a \exp \left[N \ln \left(1 - \int_x^a p_X(x') dx' \right) \right] dx \quad (8.23)$$

Il est possible de montrer que $I_N^{(1)}$ décroît exponentiellement avec N et que I_N est, comme on peut s'y attendre heuristiquement, dominée par $I_N^{(2)}$ qui, comme nous allons le voir, décroît en puissance inverse de N .

La distribution de l'échantillon, $p(x, y)$, est à symétrie circulaire et donc :

$$p_X(x) = \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} p(x, y) dy \quad (8.24)$$

$$= 2 \int_0^{\sqrt{a^2-x^2}} p(x, y) dy \quad (8.25)$$

Or, pour $x \lesssim a$, nous avons : $p(x, y) \sim A (a - \sqrt{x^2 + y^2})^{\gamma-1}$. Par conséquent, en posant $y = xu$ et en considérant $(a - \delta) < x < a$:

$$p_X(x) \sim 2 \int_0^{\sqrt{a^2-x^2}} A (a - \sqrt{x^2 + y^2})^{\gamma-1} dy \quad (8.26)$$

$$\sim 2Ax \int_0^{\sqrt{\frac{a^2}{x^2}-1}} (a - x \sqrt{1+u^2})^{\gamma-1} du \quad (8.27)$$

$$\sim 2Aa \int_0^{\sqrt{\frac{2(a-x)}{a}}} (a-x)^{\gamma-1} du \quad (8.28)$$

$$\sim 2 A \sqrt{2a} (a-x)^{\gamma-\frac{1}{2}} \quad (8.29)$$

Nous repartons maintenant de l'équation (8.23) :

$$\begin{aligned} I_N^{(2)} &= \int_{a-\delta}^a \exp \left[N \ln \left(1 - \int_x^a p_X(x') dx' \right) \right] dx \\ &\sim \int_{a-\delta}^a \exp \left[N \ln \left(1 - \int_x^a 2 A \sqrt{2a} (a-x')^{\gamma-\frac{1}{2}} dx' \right) \right] dx \\ &\sim \int_{a-\delta}^a \exp \left[N \ln \left(1 - \frac{4 A \sqrt{2a}}{2 \gamma + 1} (a-x)^{\gamma+\frac{1}{2}} \right) \right] dx \\ &\sim \int_{a-\delta}^a \exp \left[-\frac{4 A N \sqrt{2a}}{2 \gamma + 1} (a-x)^{\gamma+\frac{1}{2}} \right] dx \end{aligned} \quad (8.30)$$

(8.31)

Pour aller plus loin, nous effectuons le changement de variable suivant :

$$u = \frac{4 A N \sqrt{2a}}{2 \gamma + 1} (a-x)^{\gamma+\frac{1}{2}} \quad (8.32)$$

Dans la limite N grand où nous nous sommes placés, ce change-

ment de variable conduit à :

$$I_N^{(2)} \sim \frac{2}{[4AN\sqrt{2a}]^{\frac{2}{1+2\gamma}}} \int_0^\infty e^{-u} [(2\gamma + 1) u]^{\frac{1-2\gamma}{1+2\gamma}} du$$

$$\sim \frac{2(2\gamma + 1)^{\frac{1-2\gamma}{1+2\gamma}} \Gamma\left(\frac{2}{2\gamma+1}\right)}{[4AN\sqrt{2a}]^{\frac{2}{1+2\gamma}}} \quad (8.33)$$

De la combinaison de (8.33), (8.18) et (8.1) vient le résultat final :

$$\langle L_N \rangle_{N \rightarrow \infty} \sim 2\pi a - \frac{4\pi(2\gamma + 1)^{\frac{1-2\gamma}{1+2\gamma}} \Gamma\left(\frac{2}{2\gamma+1}\right)}{[4AN\sqrt{2a}]^{\frac{2}{1+2\gamma}}} \quad (8.34)$$

Pour prendre un exemple plus concret, considérons N points choisis indépendamment et uniformément dans le disque unité :

$$p(x, y) = \frac{1}{\pi} \Theta(1 - x^2 - y^2), \quad (8.35)$$

où Θ est la fonction de Heaviside.

Dans la notation utilisée au cours des calculs ci-dessus, cela signifie :

$$a = 1, \quad (8.36)$$

$$A = \frac{1}{\pi}, \quad (8.37)$$

$$\gamma = 1. \quad (8.38)$$

Nous trouvons donc :

$$\boxed{\langle L_N \rangle_{N \rightarrow \infty} \sim 2\pi \left(1 - \frac{\Gamma\left(\frac{2}{3}\right) \pi^{\frac{2}{3}}}{12^{\frac{1}{3}} N^{\frac{2}{3}}}\right)}, \quad (8.39)$$

en parfait accord avec le résultat de Rényi et Sulanke (éq. 5.11) [90].

Nous retrouvons le fait que, lorsque N devient grand, la probabilité de trouver des points très près du bord du disque unité, tout

autour de celui-ci, s'accroît², si bien qu'en moyenne le périmètre de l'enveloppe convexe du nuage tendra vers celui du disque, le rapprochement se faisant en loi de puissance.

Bien sûr, au-delà de cette expression asymptotique, nous pouvons obtenir, en procédant de la même manière que dans l'exemple des distributions gaussiennes (§8.2), des expressions exactes pour le périmètre et l'aire moyens de l'enveloppe convexe dans le cas de points tirés uniformément dans le disque unité :

Périmètre :

$$\langle L_N \rangle = 2\pi \left[1 - \int_{-1}^1 F_N(M) dM \right] \quad (8.40)$$

Aire :

$$\langle A_N \rangle = \pi \left[1 - \frac{8}{3} \int_{-1}^1 M F_N(M) dM \right] \quad (8.41)$$

avec :

$$F_N(M) = \frac{1}{\pi} \left[\arcsin(M) + M \sqrt{1 - M^2} \right]^N \quad (8.42)$$

Nous allons examiner dans un autre exemple le calcul du comportement asymptotique de l'aire moyenne de l'enveloppe convexe d'un nuage de points aléatoires, mais donnons ici, sans en détailler le calcul, le comportement asymptotique de (8.41) :

$$\boxed{\langle A_N \rangle_{N \rightarrow \infty} \sim \pi \left(1 - 2 \frac{\Gamma\left(\frac{2}{3}\right) 2^{\frac{7}{3}} \pi^{\frac{2}{3}}}{3^{\frac{4}{3}} N^{\frac{2}{3}}} \right)}, \quad (8.43)$$

Notons à nouveau l'accord avec le résultat de Rényi et Sulanke (éq. 5.12) [90] et le fait que l'aire de l'enveloppe convexe tend vers celle du disque dans lequel les points sont jetés. Remarquons

²Ceci est à rapprocher de la tendance à se décaler vers les grandes valeurs dont fait preuve le maximum de N variables aléatoires simples quand N devient grand, dans la statistique standard des valeurs extrêmes que nous rappelions plus haut.

aussi que l'exposant de N , contrôlant la vitesse de convergence de l'aire moyenne vers celle du disque, est le même que pour le périmètre moyen — seul le préfacteur de la puissance de N change³.

Rappelons aussi les résultats de Hilhorst *et al.* [155] pour le problème de Sylvester⁴ : lorsque N devient grand, l'enveloppe convexe du nuage de N point (sachant que les N points en sont les sommets) est comprise dans un anneau d'épaisseur $\sim N^{-\frac{4}{5}}$ — plus proche, donc, en termes de périmètre et d'aire, que les $N^{-\frac{2}{3}}$ de notre cas : ce que l'on peut comprendre physiquement en voyant qu'exiger d'avoir les N points sur l'enveloppe convexe tend à augmenter la taille de celle-ci et donc à la rapprocher du bord du disque.

Exemple 2 : aire moyenne dans le cas « Fréchet »

Considérons N points choisis indépendamment selon une distribution isotrope ayant des queues de type Fréchet :

$$p(x, y) \underset{\sqrt{x^2+y^2} \rightarrow \infty}{\sim} \frac{A}{(x^2 + y^2)^{\frac{\gamma+2}{2}}}, \quad (8.44)$$

A assurant la normalisation de la distribution :

$$\iint p(x, y) dx dy = 1.$$

Reprenant l'équation (8.11), nous commençons par son premier terme. En notant comme précédemment F_N la fonction de répartition du maximum M_N des abscisses, nous avons :

$$\langle M_N^2 \rangle = \int_{-\infty}^{\infty} x^2 F_N'(x) dx \equiv I_N. \quad (8.45)$$

³Rényi et Sulanke [90] ont montré que ceci est en fait vrai pour tout support à bord lisse, avec, en outre, toujours le même exposant : $N^{-\frac{2}{3}}$.

⁴Si N points sont choisis selon une distribution uniforme sur le disque unité, quelle est la probabilité p_N qu'ils soient les sommets d'un polygone convexe — autrement dit, qu'ils soient les sommets de leur propre enveloppe convexe ?

Où nous avons conservé la notation :

$$F_N(x) = \left[1 - \int_x^\infty p_X(x') dx' \right]^N \quad (8.46)$$

avec, toujours, $p_X(x) = \int p(x, y) dy$.

Nous choisissons $K \gg 1$ tel que pour

$$x \geq K,$$

$$p(x, y) \simeq \frac{A'}{(x^2 + y^2)^{\frac{\gamma+2}{2}}}.$$

L'idée étant que lorsque N devient grand, les points de l'échantillon auront tendance à se disséminer toujours plus loin à travers le plan, et par conséquent nous pouvons nous concentrer sur les queues de la distribution. Nous écrivons donc :

$$I_N = I_N^{(1)} + I_N^{(2)} \quad (8.47)$$

avec :

$$I_N^{(1)} = \int_{-\infty}^K x^2 F'_N(x) dx \quad (8.48)$$

et

$$I_N^{(2)} = \int_K^\infty x^2 F'_N(x) dx \quad (8.49)$$

Comme précédemment, il est possible de montrer que I_N est, comme on peut s'y attendre heuristiquement, dominée par $I_N^{(2)}$.

La distribution de l'échantillon $p(x, y)$ est à symétrie circulaire et donc :

$$p_X(x) = \int_{-\infty}^\infty p(x, y) dy \quad (8.50)$$

$$= 2 \int_0^\infty p(x, y) dy \quad (8.51)$$

Or, pour $x \gg 1$, nous avons : $p(x, y) \sim \frac{A}{(x^2+y^2)^{\frac{\gamma+2}{2}}}$. Par conséquent, en posant $y = ux$ et en considérant les cas où $x \gg 1$:

$$p_X(x) \sim 2 \int_0^\infty \frac{A}{(x^2 + y^2)^{\frac{\gamma+2}{2}}} dy \quad (8.52)$$

$$\sim 2Ax \int_0^\infty \frac{1}{x^{\gamma+2}(1+u^2)^{\frac{\gamma+2}{2}}} du \quad (8.53)$$

$$\sim \frac{A\sqrt{\pi}\Gamma\left(\frac{\gamma+1}{2}\right)}{x^{\gamma+1}\Gamma\left(\frac{\gamma}{2} + 1\right)} \quad (8.54)$$

$$\sim \frac{C}{x^{\gamma+1}} \quad (8.55)$$

où nous avons noté $C = A\sqrt{\pi}\frac{\Gamma\left(\frac{\gamma+1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}+1\right)}$.

Nous pouvons maintenant donner une expression de $F_N(x)$ pour x grand et N grand :

$$\begin{aligned} F_N(x) &\sim \left[1 - \int_x^\infty \frac{C}{x'^{\gamma+1}} dx'\right]^N \\ &\sim \left[1 - \frac{C}{\gamma x^\gamma}\right]^N \\ &\sim e^{-\frac{NC}{\gamma x^\gamma}} \end{aligned} \quad (8.56)$$

D'où, toujours pour x et N grands :

$$F'_N(x) \sim \frac{NC}{x^{\gamma+1}} e^{-\frac{NC}{\gamma x^\gamma}} \quad (8.57)$$

Nous insérons ensuite (8.57) dans (8.49), et posons $u = \frac{NC}{\gamma x^\gamma}$:

$$\begin{aligned} I_N^{(2)} &\sim \int_0^\infty \left(\frac{NC}{\gamma}\right)^{\frac{2}{\gamma}} u^{-\frac{2}{\gamma}} e^{-u} du \\ &\sim \left(\frac{NC}{\gamma}\right)^{\frac{2}{\gamma}} \Gamma\left(1 - \frac{2}{\gamma}\right) \end{aligned} \quad (8.58)$$

Examinons le second terme de l'équation (8.11). Comme précédemment, nous nous concentrons sur la partie à grand x de

l'intégrale, qui sera dominante. Nous la ré-écrivons, de manière à nous ramener au calcul effectué au paragraphe précédent :

$$\begin{aligned}
& N \int_K \int_{-\infty}^{\infty} y^2 p(x, y) F_{N-1}(x) dx dy \\
&= \int_K F'_N(x) \frac{\int_{-\infty}^{\infty} y^2 p(x, y) dy}{p_X(x)} dx \\
&\sim \int_K F'_N(x) x^2 \frac{A}{C} \int_{-\infty}^{\infty} \frac{u^2}{(1+u^2)^{\frac{\gamma+2}{2}}} du dx \\
&\sim \frac{1}{\gamma-1} \int_K F'_N(x) x^2 dx \tag{8.59}
\end{aligned}$$

Cette dernière intégrale est, au facteur $\frac{1}{\gamma-1}$ près, la même que (8.49) ; par conséquent, il vient :

$$\langle A_N \rangle_{N \rightarrow \infty} \sim \left(1 - \frac{1}{\gamma-1}\right) \int_K x^2 F'_N(x) dx,$$

qui, associée à (8.58) et simplifiée, donne :

$$\langle A_N \rangle_{N \rightarrow \infty} \sim \frac{\gamma}{\gamma-1} \left(\frac{C}{\gamma}\right)^{\frac{2}{\gamma}} \Gamma\left(2\left(1 - \frac{1}{\gamma}\right)\right) N^{\frac{2}{\gamma}}. \tag{8.60}$$

On peut montrer sans trop de difficultés qu'il s'agit bien du même résultat que celui de Carnal [92]. Nous retrouvons ici une croissance de la valeur moyenne de l'aire du même type que celle du maximum de N variables aléatoires distribuées suivant une loi appartenant à la classe d'universalité de la distribution de Fréchet, c'est-à-dire en puissance de N . Nous revenons sur ce point et ces implications géométriques dans la discussion d'ensemble qui suit le résumé des résultats pour des nuages de points indépendants.

Les calculs dans les autres cas de points indépendants (tirés d'une distribution à symétrie circulaire) étant complètement analogues, que ce soit pour le périmètre moyen ou pour l'aire moyenne, nous n'en reproduisons pas les détails mais donnons un résumé des résultats que l'on obtient relativement aisément à l'aide de la méthode de la fonction support.

**Périmètre moyen de l'enveloppe convexe de N points
indépendants
quand N est grand**

- *Queues exponentielles* ($p(x, y) \sim Ae^{-(x^2+y^2)^{\alpha/2}}$ quand $(x^2 + y^2) \rightarrow \infty$)

$$\boxed{\langle L_N \rangle \sim 2\pi \ln^{1/\alpha} N}$$

- *Queues en loi de puissance* ($p(x, y) \sim \frac{A}{(x^2+y^2)^{(\frac{\gamma+1}{2})}}$) :

$$\boxed{\langle L_N \rangle \sim 2\pi \left(\frac{A B(\frac{1}{2}, \frac{\gamma+1}{2})}{\gamma} \right)^{\frac{1}{\gamma}} \Gamma(1 - \frac{1}{\gamma}) N^{\frac{1}{\gamma}}$$

où B est la fonction beta.

- *Queues tronquées* ($p(x, y) \sim A(a - \sqrt{x^2 + y^2})^{\gamma-1}$) :

$$\boxed{\langle L_N \rangle \sim 2\pi a \left(1 - \frac{f(a, \gamma)}{N^{\frac{2}{2\gamma+1}}} \right)}$$

avec

$$f(a, \gamma) = \frac{(\gamma + \frac{1}{2})^{\frac{1-2\gamma}{1+2\gamma}} \Gamma(\frac{2}{1+2\gamma})}{a(2A\sqrt{2a})^{\frac{2}{1+2\gamma}}}$$

Nous retrouvons bien comme attendu la distinction entre les trois classes d'universalités des statistiques de valeurs extrêmes. Les nuages de points tirés de distributions présentant des « queues exponentielles » ont, en moyenne lorsque le nombre N de points devient grand, une enveloppe convexe dont le périmètre croît plus lentement (en puissance de $\ln N$) que les nuages tirés de distributions à queues en loi de puissance (pour lesquelles la croissance du périmètre de l'enveloppe convexe est en puissance de N), traduisant ainsi la moindre probabilité d'avoir des points très éloignés de l'origine dans le premier cas que dans le second. Ces deux cas sont à rapprocher de celui de points distribués à l'intérieur d'un support fini, que nous avons déjà discuté plus haut, et dans lequel la probabilité de trouver des points proches du bord du support croît avec le nombre de points du nuage, conduisant ainsi l'enveloppe convexe à tendre vers le bord du support, en l'occurrence le cercle de rayon a .

Aire moyenne enclose par l'enveloppe convexe de N points indépendants, quand N est grand

- *Queues exponentielles* :

$$\langle A_N \rangle \sim \pi \ln^\alpha N$$

- *Queues en loi de puissance* :

$$\langle A_N \rangle \sim \pi \left(\frac{\gamma}{\gamma-1} \right) \left(\frac{A B(\frac{1}{2}, \frac{\gamma+1}{2})}{\gamma} \right)^{\frac{2}{\gamma}} \Gamma \left(2 - \frac{2}{\gamma} \right) N^{\frac{2}{\gamma}}$$

où B est la fonction beta.

- *Queues tronquées* :

$$\langle A_N \rangle \sim \pi a^2 \left(1 - \frac{8 f(a, \gamma)}{3 N^{\frac{2}{2\gamma+1}}} \right)$$

avec

$$f(a, \gamma) = \frac{(\gamma + \frac{1}{2})^{\frac{1-2\gamma}{1+2\gamma}} \Gamma(\frac{2}{1+2\gamma})}{a(2A\sqrt{2a})^{\frac{2}{1+2\gamma}}}$$

Nous retrouvons pour l'aire les mêmes caractéristiques que pour le périmètre quant aux croissances relatives des enveloppes convexes selon la forme de la distribution initiale des points. Il est particulièrement intéressant de noter que dans le cas de distributions à queues exponentielles, les comportements asymptotiques de l'aire et du périmètre moyens de l'enveloppe convexe correspondent aux grandeurs géométriques d'un cercle de centre l'origine et de rayon $\ln^{\frac{1}{\alpha}} N$ (α étant l'exposant caractéristique des queues exponentielles de la distribution initiale)⁵.

8.4 Un cas non isotrope : points choisis uniformément dans un carré

Examinons maintenant un cas non isotrope : le calcul du périmètre moyen de l'enveloppe convexe de N points indépendants distribués uniformément

⁵Ceci est à rapprocher du résultat de Geffroy [85], p. 62.

dans un carré de côté a . La densité bivariée de l'échantillon peut s'écrire :

$$p(x, y) = \frac{1}{a^2} \Theta \left(\frac{a^2}{4} - x^2 \right) \Theta \left(\frac{a^2}{4} - y^2 \right), \quad (8.61)$$

où Θ est la fonction de Heaviside.

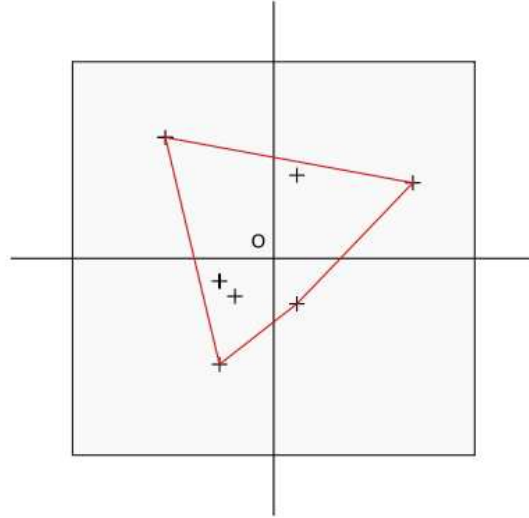


FIG. 8.1 – Enveloppe convexe d'un ensemble aléatoire de points choisis selon une distribution uniforme dans un carré

Nous considérerons, comme décrit dans la partie introductive de cette section, la projection de l'échantillon sur la droite passant par l'origine et faisant un angle θ avec l'axe des x . Nous notons la variable aléatoire correspondant à la projection d'un point de l'échantillon $z \equiv x \cos \theta + y \sin \theta$. Sa densité est donnée par :

$$p(z) = \frac{1}{a^2} \iint_{-\frac{a}{2}}^{\frac{a}{2}} \delta(z - x \cos \theta - y \sin \theta) \Theta \left(\frac{a^2}{4} - x^2 \right) \Theta \left(\frac{a^2}{4} - y^2 \right) dx dy, \quad (8.62)$$

où δ est la fonction delta de Dirac.

De par la symétrie du carré, nous pouvons nous concentrer sur $0 \leq \theta \leq \frac{\pi}{4}$ et écrire

$$x = \frac{z - y \sin \theta}{\cos \theta}.$$

Ceci nous permet de simplifier (8.62) :

$$p(z) = \frac{1}{a^2 \cos \theta} \int_{-\frac{a}{2}}^{\frac{a}{2}} \Theta \left(\frac{a^2}{4} - \left(\frac{z - y \sin \theta}{\cos \theta} \right)^2 \right) \Theta \left(\frac{a^2}{4} - y^2 \right) dy. \quad (8.63)$$

En appliquant la condition qui veut que chacun des points de l'échantillon appartienne à l'intérieur du carré, ce qui correspond à faire en sorte que les fonctions de Heaviside soient non nulles, nous avons :

$$(i) \quad -\frac{a}{2} \leq y \leq \frac{a}{2} \quad (8.64)$$

$$(ii) \quad -\frac{a}{2} \leq \frac{z - y \sin \theta}{\cos \theta} \leq \frac{a}{2} \quad (8.65)$$

$$\therefore \max \left(-\frac{a}{2}, \frac{z - \frac{a}{2} \cos \theta}{\sin \theta} \right) \leq y \leq \min \left(\frac{a}{2}, \frac{z + \frac{a}{2} \cos \theta}{\sin \theta} \right) \quad (8.66)$$

et :

$$-\frac{a}{2}(\cos \theta + \sin \theta) \leq z \leq \frac{a}{2}(\cos \theta + \sin \theta) \quad (8.67)$$

Il y aura trois cas :

1. Pour

$$-\frac{a}{2}(\cos \theta + \sin \theta) \leq z \leq \frac{a}{2}(\sin \theta - \cos \theta),$$

la coordonnée y variera entre $-\frac{a}{2}$ et $\frac{z + \frac{a}{2} \cos \theta}{\sin \theta}$, d'où :

$$p(z) = \frac{z + \frac{a}{2}(\cos \theta + \sin \theta)}{a^2 \cos \theta \sin \theta}. \quad (8.68)$$

2. Pour

$$\frac{a}{2}(\sin \theta - \cos \theta) \leq z \leq \frac{a}{2}(\cos \theta - \sin \theta),$$

la coordonnée y variera entre $-\frac{a}{2}$ et $\frac{a}{2}$, d'où :

$$p(z) = \frac{1}{a \cos \theta}. \quad (8.69)$$

3. Pour

$$\frac{a}{2}(\cos \theta - \sin \theta) \leq z \leq \frac{a}{2}(\cos \theta + \sin \theta),$$

la coordonnée y variera entre $\frac{z - \frac{a}{2} \cos \theta}{\sin \theta}$ et $\frac{a}{2}$, d'où :

$$p(z) = \frac{\frac{a}{2}(\cos \theta + \sin \theta) - z}{a^2 \cos \theta \sin \theta}. \quad (8.70)$$

Pour alléger les notations, nous écrirons désormais :

$$a_\theta = \frac{a}{2}(\cos \theta + \sin \theta) \quad (8.71)$$

$$b_\theta = a^2 \cos \theta \sin \theta \quad (8.72)$$

En désignant comme précédemment par $M_N(\theta)$ la valeur de la fonction support de l'échantillon pour un angle θ , c'est-à-dire, la valeur de la projection maximale sur la direction θ , nous avons :

$$\langle M_N(\theta) \rangle = \int_{-a_\theta}^{a_\theta} z F'_{\theta,N}(z) dz \quad (8.73)$$

$$= [z F'_{\theta,N}(z)]_{-a_\theta}^{a_\theta} - \int_{-a_\theta}^{a_\theta} F_{\theta,N}(z) dz \quad (8.74)$$

$$= a_\theta - I_\theta, \quad (8.75)$$

où :

$$F_{\theta,N}(z) = \left[\int_{-a_\theta}^z p(z') dz' \right]^N \quad (8.76)$$

$$I_\theta = \int_{-a_\theta}^{a_\theta} F_{\theta,N}(z) dz \quad (8.77)$$

Pour calculer I_θ nous utilisons notre connaissance de $p(z)$ (éqs. (8.68), (8.69), (8.70)) et nous obtenons :

$$\begin{aligned} \langle M_N(\theta) \rangle = a_\theta - \frac{\sin \theta \tan^N \theta}{2^{N-1}(2N+1)} - \frac{\cos \theta}{2^N(N+1)} [(2 - \tan \theta)^{N+1} - \tan^{N+1} \theta] \\ - \sqrt{b_\theta \tan \theta} {}_2F_1 \left(\frac{1}{2}, -N; \frac{3}{2}; \frac{\tan \theta}{2} \right), \end{aligned} \quad (8.78)$$

${}_2F_1$ étant la fonction hypergéométrique de Gauss.

Les propriétés asymptotiques des fonctions hypergéométriques [51], permettent d'analyser le comportement de $\langle M_N(\theta) \rangle$ lorsque N devient grand :

$$\langle M_N(\theta) \rangle \sim a_\theta - \sqrt{\frac{\pi b_\theta}{2N}} + o\left(\frac{1}{\sqrt{N}}\right) \quad (8.79)$$

Nous obtenons alors le résultat désiré :

$$\begin{aligned} \langle L_N \rangle &= 8 \int_0^{\frac{\pi}{4}} \langle M_N(\theta) \rangle \\ \langle L_N \rangle &\sim 4 a \left(1 - \pi \frac{\Gamma(\frac{3}{4})}{\Gamma(\frac{1}{4}) \sqrt{N}} \right). \end{aligned} \quad (8.80)$$

Ce résultat est identique à celui que Rényi et Sulanke obtinrent par une autre méthode [90]. Remarquons que, comme dans le cas de points distribués uniformément à l'intérieur d'un disque, le périmètre moyen de l'enveloppe convexe du nuage tend vers celui du support contenant le nuage — soit ici $4a$, le périmètre du carré — lorsque le nombre N de points devient grand. Cependant, la convergence est dans le cas présent moins rapide que pour le disque : en $N^{-\frac{1}{2}}$ contre $N^{-\frac{2}{3}}$. On peut penser que, physiquement et statistiquement, les coins du carré sont en quelque sorte plus « difficiles » à atteindre pour les points du nuage, et que la convergence du périmètre moyen de l'enveloppe convexe vers la valeur du périmètre du carré s'en trouve ralentie.

Chapitre 9

Points corrélés

Comme nous le mentionnions au début de cette partie, une des principales caractéristiques de l'approche présentée ici est sa généralité : elle peut être appliquée à des échantillons de points corrélés aussi bien qu'à des points indépendants. Nous choisissons ici d'étudier au moyen de cette méthode l'enveloppe convexe de n mouvements browniens plans, qui n'a fait l'objet que dans le cas $n = 1$ d'un certain nombre de travaux chez les mathématiciens [84, 96, 99, 120, 121].

Outre son intérêt théorique, l'étude de l'enveloppe convexe de n chemins browniens plans peut être motivée par une question particulièrement pertinente pour la préservation des espèces animales dans leur habitat, comme nous allons le voir dans le prochain paragraphe, avant de donner le détail de nos calculs et résultats.

9.1 Chemins browniens plans *et domaine vital*

Une question récurrente à laquelle doivent répondre les écologistes, en particulier lors de la création d'une réserve naturelle visant à protéger une population animale donnée [156], est celle du domaine vital de cette population, à savoir : quel espace un animal ou un groupe d'animaux ont-ils besoin de pouvoir explorer au cours d'un certain laps de temps pour survivre ? On peut par exemple, pour des espèces ayant un lieu d'attache permanent auquel

elles retournent chaque jour (un lieu de séjour nocturne par exemple), prendre comme laps de temps la durée d'une journée. Le domaine vital se définit alors simplement comme l'espace exploré par le groupe lors de sa recherche quotidienne de nourriture. Les spécialistes mettent en œuvre diverses méthodes pour évaluer ce territoire à partir du relevé des positions des animaux [82]. L'une de ces méthodes consiste simplement à considérer l'enveloppe convexe des positions relevées — ce qui peut paraître grossier, puisqu'il n'est alors pas tenu compte de certains facteurs topographiques pouvant interdire l'accès à telle ou telle partie du territoire ainsi dessiné ; cependant, dans certaines circonstances, notamment de faiblesse des données statistiques, cette méthode demeure la meilleure [157].

Les positions relevées en suivant un animal (par exemple à l'aide d'une puce GPS) apparaîtront comme les sommets d'un chemin dont les caractéristiques varieront en fonction du type de mouvement réalisé : lors de phases de déplacements correspondant à une recherche de nourriture, les positions décrites peuvent être vues comme les sommets d'une marche aléatoire dans le plan [158–160]. Ainsi, pour des animaux dont les déplacements quotidiens se composent essentiellement de telles phases, les grandeurs géométriques du territoire (telles que son périmètre et son aire) pourront être estimées par l'intermédiaire des grandeurs moyennes de l'enveloppe convexe de la marche aléatoire correspondante (Fig. 9.1). Si le relevé des positions

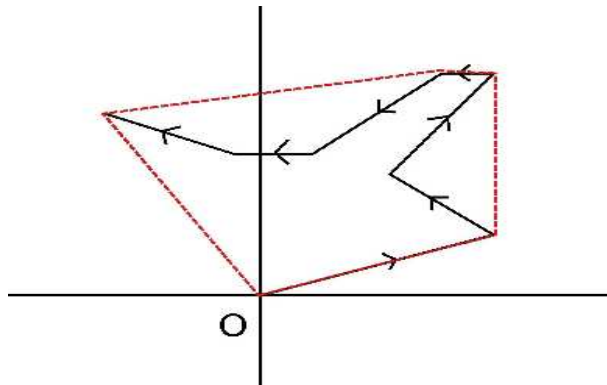


FIG. 9.1 – Enveloppe convexe d'une marche aléatoire (7 pas)

est très fin dans l'espace et/ou très long dans le temps, le nombre de pas de

la marche aléatoire plane sera très grand et son parcours pourra, selon une bonne approximation, être considéré comme la trajectoire d'un mouvement brownien plan de durée T égale à la durée d'observation (fig. 9.2). (Notons que cette approximation vaut pour une marche aléatoire dont la distribution de la taille des pas est à variance finie.)

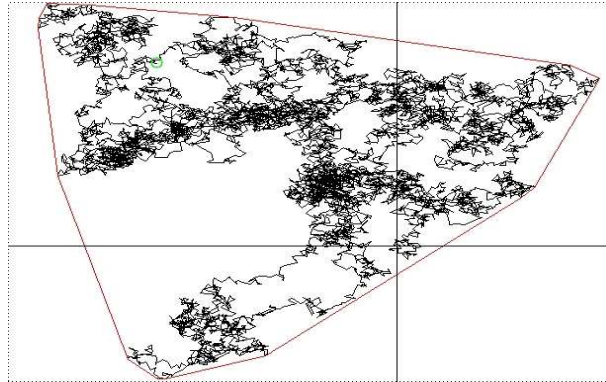


FIG. 9.2 – Enveloppe convexe d'un mouvement brownien plan

Le territoire d'un animal peut donc être caractérisé par le périmètre moyen et l'aire moyenne de l'enveloppe convexe d'un mouvement brownien plan, issu de l'origine et de trajectoire ouverte ou fermée, au choix, ce dernier cas correspondant bien, par exemple, à un animal regagnant son nid ou sa tanière chaque soir après avoir passé la journée à explorer son territoire en quête de nourriture. Comme nous l'avons vu lors de notre présentation des travaux antérieurs sur ce sujet, le périmètre moyen et l'aire moyenne de l'enveloppe convexe d'un chemin brownien ouvert sont connus [96, 99], de même que le périmètre moyen dans le cas d'un chemin fermé [120]. S'il est sur le plan théorique naturel de vouloir prolonger ces résultats au cas de plusieurs mouvements brownien plans (fig. 9.3), il l'est tout autant sur le plan écologique, beaucoup d'animaux vivant en groupe. Or, aucun résultat ne semble avoir été donné dans ce cas jusqu'à présent. C'est cette lacune que nous nous proposons ici de combler, en appliquant la méthode de la fonction support exposée plus haut, combinée aux résultats obtenus dans la première partie de ce mémoire. Nous mettons d'abord ceci en oeuvre pour $n = 1$ mouvement brownien dans le plan, puis pour $n > 1$.

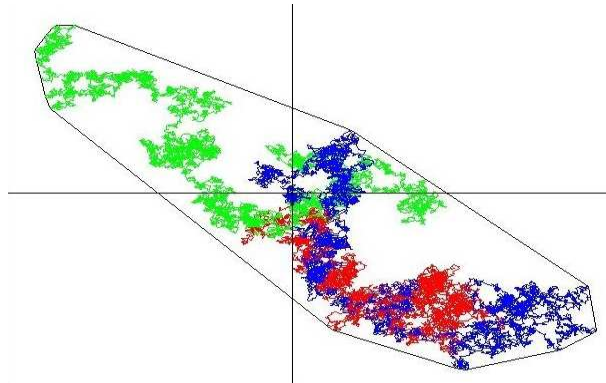


FIG. 9.3 – Enveloppe convexe de 3 mouvements browniens plans indépendants, fermés, issus de l'origine O .

9.2 Enveloppe convexe du mouvement brownien plan

Nous considérons ici un mouvement brownien plan de durée T , issu de l'origine O :

$$\mathcal{B}(\tau) = (x(\tau), y(\tau))$$

avec

$$0 \leq \tau \leq T,$$

$x(\tau)$ et $y(\tau)$ étant des mouvements browniens linéaires standard de durée T , tels que définis au premier chapitre de ce mémoire, c'est-à-dire obéissant aux équations de Langevin suivantes :

$$\dot{x}(\tau) = \eta_x(\tau)$$

et

$$\dot{y}(\tau) = \eta_y(\tau)$$

où $\eta_x(\tau)$ et $\eta_y(\tau)$ sont des bruits blancs gaussiens, de moyenne nulle et de corrélation *delta* :

$$\langle \eta_x(\tau) \eta_x(\tau') \rangle = \delta(\tau - \tau').$$

Notons en passant un point qui nous sera utile plus tard et qui découle de cette définition :

$$\langle x^2(\tau) \rangle = \tau$$

et

$$\langle y^2(\tau) \rangle = \tau. \quad (9.1)$$

[$\langle \rangle$ désigne la moyenne sur toutes les réalisations possibles d'un mouvement brownien libre de durée T .]

Fixons une direction θ . Nous allons comme précédemment (éqs. (7.6) et (7.7)) utiliser les projections sur la direction θ et sur la direction qui lui est perpendiculaire :

$$z_\theta(\tau) = x(\tau) \cos \theta + y(\tau) \sin \theta$$

et

$$h_\theta(\tau) = -x(\tau) \sin \theta + y(\tau) \cos \theta.$$

Ici, z_θ et h_θ sont deux mouvements browniens linéaires (chacun de durée T), paramétrés par θ . Il s'ensuit que $M(\theta)$ est simplement le maximum du brownien $z_\theta(\tau)$ sur l'intervalle $\tau \in [0, T]$, autrement dit :

$$M(\theta) = \max_{\tau \in [0, T]} [z_\theta(\tau)].$$

En outre, si nous notons τ^* le temps auquel ce maximum est atteint, alors :

$$M(\theta) = z_\theta(\tau^*) = x(\tau^*) \cos \theta + y(\tau^*) \sin \theta.$$

En dérivant par rapport à θ , il vient :

$$M'(\theta) = -x(\tau^*) \sin \theta + y(\tau^*) \cos \theta = h_\theta(\tau^*).$$

Pour le dire en mots : si $M(\theta)$ est le maximum du premier mouvement brownien $z_\theta(\tau)$, $M'(\theta)$ correspond à la position du second mouvement (indépendant du premier¹) $h_\theta(\tau)$ à l'instant $\tau = \tau^*$ auquel le premier atteint son maximum. (cf. Fig. 9.4(a) et 9.4(b)).

¹En effet :

$$\begin{aligned} \langle z_\theta(\tau) h_\theta(\tau') \rangle &= \langle x(\tau) y(\tau') \rangle \cos^2 \theta - \langle x(\tau') y(\tau) \rangle \sin^2 \theta \\ &\quad - \langle x(\tau) x(\tau') \rangle \cos \theta \sin \theta + \langle y(\tau) y(\tau') \rangle \cos \theta \sin \theta = 0, \end{aligned}$$

en utilisant l'indépendance de x et y et le fait que $\langle x(\tau) x(\tau') \rangle = \langle y(\tau) y(\tau') \rangle = \delta(\tau - \tau')$.

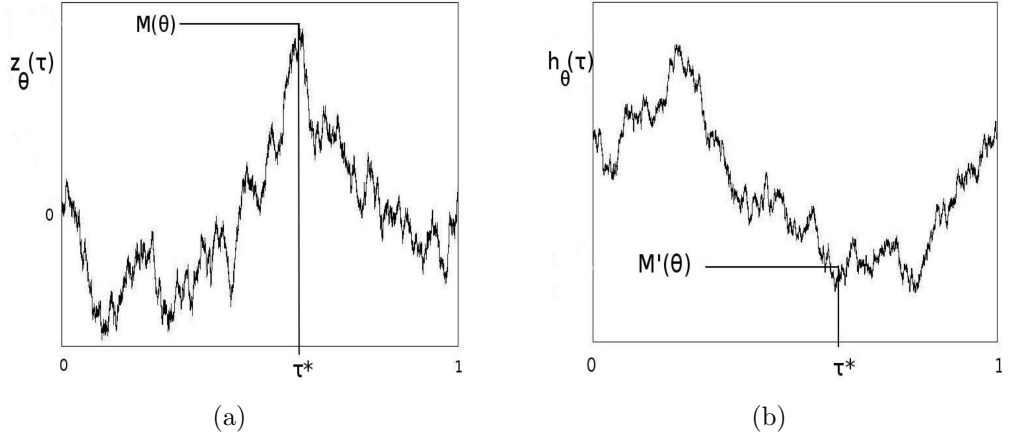


FIG. 9.4 – (a) Temps τ^* d'atteinte du maximum $M(\theta)$ de $z_\theta(\tau)$ et (b) valeur correspondante de $M'(\theta) = h_\theta(\tau^*)$

En particulier, lorsque $\theta = 0$, $z_0(\tau) = x(\tau)$ et $h_0(\tau) = y(\tau)$, et $M(0)$ est alors le maximum de $x(\tau)$ sur l'intervalle $\tau \in [0, T]$ tandis que $M'(0) = y(\tau^*)$ est la valeur de y à l'instant τ^* où x atteint son maximum.

Rappelons que dans les cas isotropes, les formules de Cauchy (éqs. 7.12 et 7.13) deviennent :

$$\langle L \rangle = 2\pi \langle M(0) \rangle \quad (9.2)$$

$$\langle A \rangle = \pi (\langle [M(0)]^2 \rangle - \langle [M'(0)]^2 \rangle). \quad (9.3)$$

Le mouvement plan que nous considérons ici est, de par sa définition, isotrope, et nous pourrions donc utiliser cette version des formules de Cauchy.

Comme nous l'avons vu au chapitre 2, la distribution du maximum d'un mouvement brownien linéaire $x(\tau)$ sur $[0, T]$ est connue, et donnée par la fonction de répartition suivante :

$$F(M) = \text{Prob}[M(0) \leq M] = \text{erf}\left(\frac{M}{\sqrt{2T}}\right), \quad (9.4)$$

où

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du.$$

Les deux premiers moments de cette distribution se calculent aisément :

$$\langle M(0) \rangle = \sqrt{\frac{2T}{\pi}}$$

et

$$\langle [M(0)]^2 \rangle = T.$$

L'équation (9.2) conduit alors au périmètre moyen de l'enveloppe convexe d'un mouvement brownien plan :

$$\boxed{\langle L \rangle = \sqrt{8\pi T}.} \quad (9.5)$$

Il est légèrement plus compliqué de calculer la valeur moyenne de l'aire enclose par l'enveloppe convexe car il faut alors connaître $\langle [M'(0)]^2 \rangle$. Rappelons d'abord (éq. 9.1) que pour un τ^* donné,

$$\langle y^2(\tau^*) \rangle = \tau^*$$

(puisque y est un mouvement brownien standard), l'espérance portant sur toute les valeurs possibles de y à τ^* fixé. Or τ^* est lui-même une variable aléatoire : il s'agit du temps auquel le mouvement brownien x atteint son maximum . Il nous faut donc également faire porter la moyenne sur la distribution de τ^* , que nous avons étudié au chapitre 2 de ce mémoire et qui est donnée par la loi de l'arcsinus de Lévy : $P(\tau^*, T) = 1/[\pi\sqrt{\tau^*(T - \tau^*)}]$; il vient alors :

$$\langle [M'(0)]^2 \rangle = \langle \tau^* \rangle = T/2;$$

d'où nous obtenons, *via* l'équation (9.3), l'expression exacte de l'aire enclose par l'enveloppe convexe du mouvement brownien plan :

$$\boxed{\langle A \rangle = \frac{\pi T}{2}.} \quad (9.6)$$

Si nous considérons maintenant une trajectoire brownienne plane fermée, c'est-à-dire contrainte à revenir à l'origine après un temps T , le raisonnement

est totalement analogue, à ceci près que $x(\tau)$ et $y(\tau)$ sont désormais des ponts browniens de durée T plutôt que des browniens libres :

$$x(0) = x(T) = 0$$

$$y(0) = y(T) = 0.$$

Comme nous l'avons vu également au chapitre 2, la distribution du maximum d'un pont brownien est également connue et ses deux premiers moments sont donnés par :

$$\langle M(0) \rangle = \sqrt{\frac{\pi T}{8}}$$

et

$$\langle [M(0)]^2 \rangle = \frac{T}{2}.$$

L'équation (9.2) conduit ensuite au périmètre moyen de l'enveloppe convexe d'un pont brownien plan :

$$\boxed{\langle L \rangle = \sqrt{\frac{\pi^3 T}{2}}.} \quad (9.7)$$

Pour calculer l'aire moyenne de l'enveloppe convexe, notons d'abord que, pour un brownien, à un instant donné τ^* :

$$\langle y^2(\tau^*) \rangle = \frac{\tau^*(T - \tau^*)}{T}.$$

Rappelons ensuite un résultat indiqué au chapitre 2 : la distribution de τ^* est uniforme. Dès lors, en prenant la moyenne sur τ^* , de distribution $P(\tau^*, T) = 1/T$, il vient :

$$\langle M'(0)^2 \rangle = \langle y^2(\tau^*) \rangle = \frac{T}{6}.$$

Enfin, comme précédemment, l'équation (9.3) nous conduit à une expression exacte de **l'aire enclose par l'enveloppe convexe d'un chemin brownien plan fermé** :

$$\boxed{\langle A \rangle = \frac{\pi T}{3}.} \quad (9.8)$$

Les résultats (9.5) et (9.6) ont été donnés par M. El Bachir [99], avec la même approche que celle utilisée ici, suggérée par L. Takács [96]. L'équation

(9.7) a été donnée par A. Goldman, à partir d'une analogie avec les mosaïques poissonniennes du plan. Le dernier résultat, (9.8), est, pour autant que nous sachions, nouveau.

9.3 Cas (particulier) d'une excursion brownienne

Par analogie avec les mouvements linéaires, nous appelons excursion brownienne plane un chemin brownien fermé (issu de l'origine et y retournant) confiné au demi-plan supérieur. Ce contexte est loin de nous être inconnu, puisqu'il s'agit, à une rotation près, de celui du polymère brownien envisagé au chapitre 6. Comme nous l'avons alors souligné, la fonction de répartition calculée était identique à celle de la valeur de la fonction support pour une direction donnée.

Ainsi, le calcul que nous avons mené pour obtenir le premier moment de la distribution correspondante (éq. (6.61)), nous conduit directement, *via* la formule de Cauchy (9.2), au **périmètre moyen de l'enveloppe convexe d'une excursion brownienne dans le demi-plan supérieur** :

$$\boxed{\langle L \rangle = \sqrt{\frac{\pi^3 T}{2}}} \quad (9.9)$$

À notre connaissance, ce résultat est nouveau. Remarquons qu'il est identique à celui pour un pont brownien fermé (éq. (9.7)), autrement dit le fait que nous nous restreignons ici, parmi tous les mouvements browniens fermés (issus de l'origine) de durée T , à ceux qui demeurent dans le demi-plan supérieur ne modifie pas la valeur moyenne du périmètre de l'enveloppe convexe du mouvement. Pour le dire encore autrement : la moyenne du périmètre de l'enveloppe convexe sur le sous-ensemble des chemins fermés (issus de l'origine) confinés au demi-plan supérieur est égale à la moyenne prise sur l'ensemble des chemins fermés issus de l'origine.

On peut comprendre cette identité en remarquant qu'il est possible de construire, entre l'ensemble des chemins fermés et le sous-ensembles des excursions, une application surjective qui conserve le périmètre de l'enveloppe

convexe. Pour ce faire, considérons d'abord des marches aléatoires fermées constituées de k pas : une telle marche peut être vue comme un ensemble de k points dont l'ordre est donné par les numéros 1 à k (le k^{e} pas ramenant la marche à son point de départ). Imaginons donc k points dans un plan orienté (avec une verticale et une horizontale) mais pour l'instant dépourvu d'origine. À un même ensemble de k points correspondent $k!$ marches possibles, si la numérotation est entièrement libre. Rien n'empêche ensuite de dire que le point numéro 1 coïncide avec l'origine O du plan et de fixer ainsi celle-ci. Nous pouvons ainsi construire toutes les marches de k pas fermées issues de l'origine.

Autre manière de procéder : on peut attribuer systématiquement le numéro 1 au point le plus « bas » des k puis numéroter les autres librement, de $(k - 1)!$ façons. En faisant toujours coïncider l'origine du plan avec le point numéro 1, les marches ainsi obtenues sont des excursions de k pas dans le demi-plan supérieur (et toutes les marches de ce type peuvent être construites ainsi).

Nous voyons alors qu'un même ensemble de k points pourra correspondre à la fois à $k!$ marches fermées non contraintes ou $(k - 1)!$ marches fermées contraintes à demeurer dans un demi-plan. Toutes ces marches sont équiprobables (dans leurs ensembles respectifs) et auront bien sûr la même enveloppe convexe. Intuitivement, cela suffit pour comprendre que les deux types de marche auront des enveloppes convexes aux grandeurs moyennes égales².

En supposant que notre raisonnement supportent le passage à la limite continue, nous pouvons donc obtenir un autre résultat nouveau, donnant **l'aire moyenne enclose par l'enveloppe convexe d'une excursion brownienne (issue de l'origine) de durée T dans le demi-plan supérieur**, identique à celle d'un simple chemin brownien fermé de même

²Formellement, on peut finaliser le raisonnement en créant des classes d'équivalence de marches passant par les « mêmes » points (l'identité des points concerne ici leurs positions relatives, puisque nous forçons l'origine du plan à correspondre au premier point de la marche). On pourra alors définir une bijection préservant les grandeurs géométriques de l'enveloppe convexe entre l'ensemble quotient des marches non contraintes et celui des marches contraintes.

durée :

$$\boxed{\langle A \rangle = \frac{\pi T}{3}}. \quad (9.10)$$

9.4 Enveloppe convexe de n chemins browniens plans

Comme mentionné plus haut, la méthode utilisée ici peut se généraliser à n chemins browniens plans indépendants, ouverts ou fermés. Dans un tel cas, nous avons alors deux jeux de n mouvements browniens linéaires indépendants :

$$\mathcal{B}_j(\tau) = (x_j(\tau), y_j(\tau))$$

($j = 1, 2, \dots, n$). Par isotropie, nous pouvons à nouveau utiliser les éqs. (9.2) et (9.3), à ceci près que désormais $M(0)$ désigne le maximum global de n mouvements browniens linéaires indépendants (n ponts dans le cas de chemins fermés) $x_j(\tau)$ ($j = 1, 2, \dots, n$), chacun de durée T :

$$M(0) = \max_{1 \leq j \leq n} \max_{0 \leq \tau \leq T} [x_1(\tau), x_2(\tau), \dots, x_n(\tau)]. \quad (9.11)$$

Soient j_* et τ^* l'indice du chemin qui atteint le maximum global et le temps auquel le maximum global est atteint. Pour les mêmes raisons que dans le cas $n = 1$, nous avons $M'(0) = y_{j_*}(\tau^*)$, autrement dit, $M'(0)$ est égal à la position du j_* -ième chemin y à l'instant où le maximum global des chemins x est atteint.

Nous avons calculé au chapitre 4 de ce mémoire la distribution jointe de $M(0)$, le maximum global de n mouvements browniens linéaires, et de son temps d'atteinte, à la fois pour des mouvements libres et pour des ponts ; nous allons donc pouvoir utiliser ces résultats pour accéder aux grandeurs géométriques de l'enveloppe convexe de n chemins browniens dans le plan. Nous donnons dans cette section les calculs permettant d'accéder à ces grandeurs, avant d'en discuter les résultats à la section suivante (§ 9.5).

9.4.1 Chemins ouverts

Considérons d'abord n chemins ouverts. À partir de la fonction de répartition $F_n(M(0))$ que nous avons donnée au chapitre 4 (éq. 4.3), les deux premiers moments $\langle M(0) \rangle$ et $\langle [M(0)]^2 \rangle$ peuvent être calculés pour tout n . Nous obtenons alors grâce à l'éq. (9.2) le périmètre moyen de l'enveloppe convexe, $\langle L_n \rangle = \alpha_n \sqrt{T}$ avec

$$\boxed{\alpha_n = 4n\sqrt{2\pi} \int_0^\infty du u e^{-u^2} [\operatorname{erf}(u)]^{n-1}.} \quad (9.12)$$

Les premières valeurs sont :

$$\begin{aligned} \alpha_1 &= \sqrt{8\pi} = 5.013\dots \\ \alpha_2 &= 4\sqrt{\pi} = 7.089\dots \\ \alpha_3 &= 24 \frac{\tan^{-1}(1/\sqrt{2})}{\sqrt{\pi}} = 8.333\dots \end{aligned}$$

(*cf.* Fig. 9.5 pour une représentation graphique de α_n en fonction de n). Pour les grandes valeurs de n , il est possible d'analyser l'intégrale apparaissant dans l'éq. (9.12) par la méthode de Laplace :

$$\boxed{\alpha_n \sim 2\pi\sqrt{2 \ln n}.} \quad (9.13)$$

(Nous donnons les détails de cette analyse en annexe, p. 165)

Notons la dépendance logarithmique en n , sur l'importance de laquelle nous reviendrons à la fin de ce chapitre.

Pour calculer l'aire moyenne de l'enveloppe convexe, il nous faut calculer le terme $\langle [M'(0)]^2 \rangle$ de l'éq. (9.3). Nous procédons comme dans le cas $n = 1$. Pour un indice j et un temps τ donnés :

$$\langle y_j^2(\tau) \rangle = \tau,$$

puisque $y_j(\tau)$ n'est rien autre qu'un mouvement brownien linéaire. Ainsi :

$$\langle y_{j^*}^2(\tau^*) \rangle = \tau^*.$$

Il nous faut ensuite prendre la moyenne sur τ^* qui est le temps auquel le maximum global qui apparaît dans l'éq. (9.11) survient. Nous avons calculé cette distribution au chapitre 4 et obtenu (eq. (4.7)) :

$$\sigma_n(\tau^*) = \frac{1}{T} f_n(\tau^*/T),$$

avec

$$f_n(z) = \frac{2}{\pi \sqrt{z(1-z)}} \int_0^\infty dx x e^{-x^2} [\operatorname{erf}(x\sqrt{z})]^{n-1}. \quad (9.14)$$

En prenant la moyenne sur τ^* suivant cette distribution, nous pouvons calculer

$$\langle [M'(0)]^2 \rangle = \int_0^T \tau^* \sigma_n(\tau^*) d\tau^*.$$

Nous insérons ensuite ce résultat dans l'éq. (9.3) et obtenons l'aire exacte de l'enveloppe convexe de n mouvements browniens plans, pour tout n :

$$\langle A_n \rangle = \beta_n T$$

avec

$$\boxed{\beta_n = 4n \sqrt{\pi} \int_0^\infty du u [\operatorname{erf}(u)]^{n-1} (ue^{-u^2} - g(u))}, \quad (9.15)$$

où

$$g(u) = \frac{1}{2\sqrt{\pi}} \int_0^1 \frac{e^{-u^2/t} dt}{\sqrt{t(1-t)}}.$$

Les quelques premières valeurs sont :

$$\beta_1 = \pi/2 = 1.570\dots \quad (9.16)$$

$$\beta_2 = \pi = 3.141\dots \quad (9.17)$$

$$\beta_3 = \pi + 3 - \sqrt{3} = 4.409\dots \quad (9.18)$$

(La fig. 9.5 reproduit un graphique de β_n en fonction de n).

L'analyse pour n grand (cf. D.2) donne :

$$\boxed{\beta_n \sim 2\pi \ln n.} \quad (9.19)$$

9.4.2 Chemins fermés

Pour n ponts browniens plans, le raisonnement est complètement analogue. Les différences sont :

- le propagateur du mouvement linéaire obtenu par projection sur l'axe des x est celui d'un pont brownien linéaire et donc :

$$\langle y_{j_*}^2(\tau^*) \rangle = \frac{\tau^*(T - \tau^*)}{T}$$

- la fonction de répartition du maximum $\langle M_n \rangle$ de n ponts browniens linéaires (de durée T) est donnée par (cf. éq. (4.3) et (2.21)) :

$$F_n(M) = \left[1 - e^{-\frac{2M^2}{T}} \right]^n.$$

- la distribution du temps auquel le maximum $\langle M_n \rangle$ de n ponts browniens linéaires (de durée T) est atteint est donnée par (cf. éq. (4.11)) : $\sigma_n(\tau) = \frac{1}{T} g_n\left(\frac{\tau}{T}\right)$ avec

$$g_n(z) = n \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{(-1)^k}{[1 + 4kz(1-z)]^{\frac{3}{2}}}.$$

Nous suivons ensuite la même voie qu'au paragraphe précédent. Le périmètre moyen de l'enveloppe convexe découle directement de la moyenne de M_n :

$$\begin{aligned} \langle M_n \rangle &= \int_0^\infty M F_n'(M) dM \\ &= \int_0^\infty [1 - F_n(M)] dM \\ &= \int_0^\infty [1 - [1 - e^{-\frac{2M^2}{T}}]^n] dM \\ &= \frac{1}{2} \sqrt{\frac{\pi T}{2}} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^{k+1}}{\sqrt{k}} \end{aligned} \tag{9.20}$$

Nous obtenons donc, *via* la formule de Cauchy (9.2), l'expression suivante pour le périmètre moyen de l'enveloppe convexe de n trajectoires browniennes planes fermées :

$$\langle L_n \rangle = \alpha_n(c) \sqrt{T}$$

avec, pour tout n :

$$\boxed{\alpha_n(c) = \frac{\pi^{3/2}}{\sqrt{2}} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^{k+1}}{\sqrt{k}}} \quad (9.21)$$

Pour l'aire moyenne, nous procédons également comme au paragraphe précédent, avec bien sûr les différences précisées ci-dessus.

Pour le premier terme de la formule de Cauchy (9.3), nous avons :

$$\begin{aligned} \langle [M_n(0)]^2 \rangle &= \int_0^\infty M^2 F_n'(M) dM \\ &= 2 \int_0^\infty M[1 - F_n(M)] dM \\ &= 2 \int_0^\infty M[1 - [1 - e^{-\frac{2M^2}{T}}]^n] dM \\ &= \frac{\pi T}{2} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^{k+1}}{k} \end{aligned} \quad (9.22)$$

$$= \frac{\pi T}{2} \sum_{k=1}^n \frac{1}{k}. \quad (9.23)$$

Pour le second terme, nous avons :

$$\begin{aligned} \langle [M_n'(0)]^2 \rangle &= \int_0^T \frac{\tau^*(T - \tau^*)}{T} \frac{1}{T} g_n \left(\frac{\tau^*}{T} \right) d\tau^* \\ &= T \int_0^1 z(1 - z) g_n(z) dz \end{aligned}$$

avec

$$g_n(z) = n \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{(-1)^k}{[1 + 4kz(1-z)]^{\frac{3}{2}}}.$$

L'intégrale peut se calculer par exemple à l'aide d'un logiciel de calcul formel.

En combinant les résultats pour $\langle [M_n(0)]^2 \rangle$ et $\langle [M_n'(0)]^2 \rangle$ et en posant

$$\langle A_n \rangle = \beta_n(c)T,$$

nous obtenons finalement :

$$\boxed{\beta_n(c) = \frac{\pi}{2} \left[\sum_{k=1}^n \frac{1}{k} - \frac{n}{3} + \frac{1}{2} \sum_{k=2}^n (-1)^k f(k) \right]} \quad (9.24)$$

avec

$$f(k) = \binom{n}{k} (k-1)^{-3/2} \left(k \tan^{-1}(\sqrt{k-1}) - \sqrt{k-1} \right).$$

Les quelques premières valeurs des coefficients des valeurs moyennes du périmètre et de l'aire de n ponts browniens plans sont :

$$\begin{aligned} \alpha_1(c) &= \sqrt{\pi^3/2} = 3.937\dots \\ \alpha_2(c) &= \sqrt{\pi^3}(\sqrt{2} - 1/2) = 5.090\dots \\ \alpha_3(c) &= \sqrt{\pi^3}(3/\sqrt{2} - 3/2 + 1/\sqrt{6}) = 5.732\dots \end{aligned}$$

et

$$\beta_1(c) = \pi/3 = 1.047\dots \quad (9.25)$$

$$\beta_2(c) = \pi(4 + 3\pi)/24 = 1.757\dots \quad (9.26)$$

$$\beta_3(c) = 2.250\dots \quad (9.27)$$

(*cf.* Fig. 9.5 pour un graphique de $\alpha_n(c)$ et $\beta_n(c)$ en fonction de n).

L'analyse pour n grand (*cf.* annexes D.3 et D.4) montre que :

$$\boxed{\alpha_n(c) \sim \pi\sqrt{2 \ln n}} \quad (9.28)$$

et

$$\boxed{\beta_n(c) \sim \frac{\pi}{2} \ln n,} \quad (9.29)$$

plus petit respectivement d'un facteur $1/2$ et $1/4$ que les résultats correspondants pour des chemins ouverts — conformément à l'intuition que l'on peut avoir, le chemin fermé « ayant un retour » vers l'origine à effectuer que le chemin ouvert n'a pas.

9.5 Simulations numériques et discussion

Pour illustrer les résultats analytiques relatifs à l'enveloppe convexe du mouvement brownien plan, nous avons réalisé des simulations numériques.

Celles-ci nécessitent d'une part de générer des chemins browniens dans le plan, et d'autre part de calculer numériquement l'enveloppe convexe de ces chemins. Nous expliquons en annexe (§ E) la manière dont nous procédons, à partir d'un algorithme proposé par R. L. Graham [70]. Cet algorithme n'est pas le plus rapide, et ne serait de ce fait pas le plus adapté à une simulation à grande échelle. Les simulations réalisées ici visent avant tout à illustrer notre propos, aussi n'avons nous pas poussé plus avant la recherche et l'implémentation d'un algorithme plus efficace³.

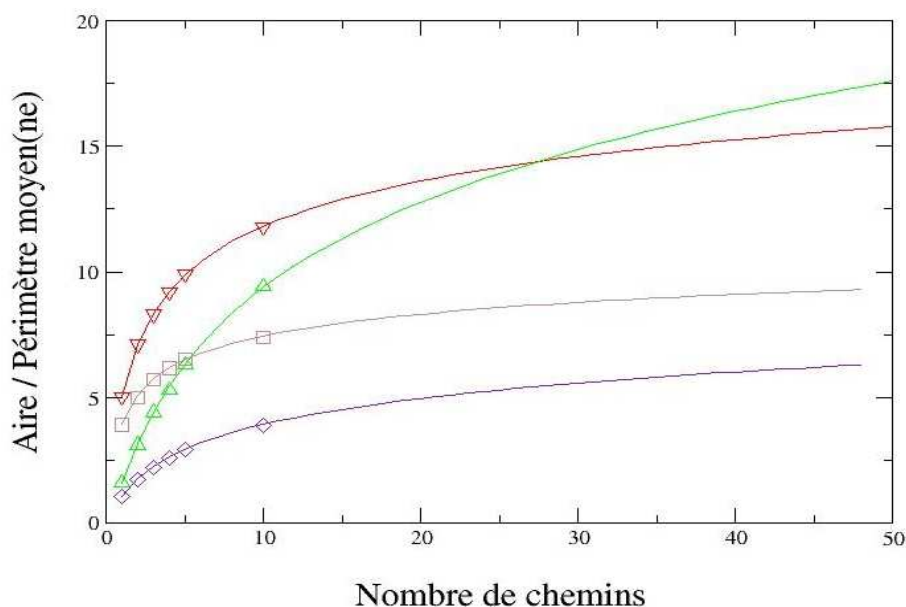


FIG. 9.5 – Avec $T = 1$, résultats analytiques pour le périmètre moyen α_n [équ. (9.12), en rouge] et l'aire moyenne β_n [équ. (9.15), en vert] de n chemins browniens ouverts, et de même pour le périmètre moyen $\alpha_n(c)$ [équ. (9.21), en brun] et l'aire moyenne $\beta_n(c)$ [équ. (9.24), en violet] de n chemins browniens fermés, en fonction de n . Les symboles correspondent aux résultats de simulations numériques (jusqu'à $n = 10$, avec 10^3 réalisations pour chaque point.).

Rappelons les comportements asymptotiques des formules exactes

³Il s'agit d'un problème classique en algorithmique, bien documenté dans la littérature [74, 78].

représentées sur la figure 9.5, qui sont donnés par les équations (9.13), (9.19), (9.28) et (9.29).

Pour n chemins browniens plans ouverts :

$$\langle L_n \rangle \sim 2\pi\sqrt{2T \ln n} \quad (9.30)$$

$$\langle A_n \rangle \sim 2\pi T \ln n; \quad (9.31)$$

et pour n chemins browniens plans fermés :

$$\langle L_n^{(c)} \rangle \sim \pi\sqrt{2T \ln n} \quad (9.32)$$

$$\langle A_n^{(c)} \rangle \sim \frac{\pi}{2} T \ln n. \quad (9.33)$$

Dans les deux cas le rapport entre la valeur moyenne de l'aire de l'enveloppe convexe et le carré de celle du périmètre présente asymptotiquement la même limite que dans le cas d'un cercle :

$$\frac{\langle A_n \rangle}{\langle L_n \rangle^2} \underset{n \rightarrow \infty}{\sim} \frac{1}{4\pi}. \quad (9.34)$$

Autrement dit, en moyenne et lorsque n devient grand, l'enveloppe convexe de n chemins browniens plan sature l'inégalité isopérimétrique⁴ :

$$L^2 - 4\pi A \geq 0. \quad (9.35)$$

Cette enveloppe convexe moyenne doit donc être un cercle, centré sur l'origine, et dont le rayon s'obtient en divisant $\langle L_n \rangle$ par 2π :

$$\boxed{R_n = \sqrt{2T \ln n}} \quad (9.36)$$

pour n chemins ouverts; et :

$$\boxed{R_n^{(c)} = \sqrt{\frac{T \ln n}{2}}} \quad (9.37)$$

pour n chemins fermés.

⁴Sur ce sujet, le lecteur intéressé se reportera avec profit à l'article de B. Teissier dans les *Leçons de mathématiques d'aujourd'hui* [161]. Il y trouvera également une démonstration des formules de Cauchy en toute dimension.

Ceci appelle plusieurs remarques. D'abord, notons qu'il n'en va pas ainsi pour n fini : par exemple, pour $n = 1$,

$$\frac{\langle A_1 \rangle}{\langle L_1 \rangle^2} = \frac{1}{16} < \frac{1}{4\pi}. \quad (9.38)$$

L'aire enclose est dans ce cas moindre que celle du disque ayant même périmètre. Or, l'isotropie du problème, dans tous les cas, pourrait nous conduire à nous attendre à ce que l'enveloppe convexe moyenne de n chemins browniens soit une courbe à symétrie circulaire, sans direction privilégiée — c'est-à-dire un cercle⁵. Le fait que ce ne soit le cas que pour n grand suggère qu'il est nécessaire pour cela que l'augmentation du nombre de chemins vienne en quelque sorte « lisser » l'enveloppe convexe globale.

Second point à noter, la forme asymptotique du préfacteur pour l'aire moyenne de l'enveloppe convexe de n chemins browniens ouverts, à savoir $\ln n$, est identique non seulement à celle de l'aire moyenne de l'enveloppe convexe de n points indépendants choisis selon une distribution gaussienne $\rho(x, y) \propto e^{-\frac{x^2+y^2}{2}}$, mais aussi à celle de la valeur moyenne du nombre de sites distincts visités par n marcheurs aléatoires sur un réseau⁶. Attardons-nous sur ce point. Le comportement asymptotique du nombre de sites distincts visités par n marcheurs aléatoires « non malveillants » sur un réseau (par exemple \mathbb{Z}^2) a été étudié de manière systématique par H. Larralde *et al.* [136, 162] (on trouvera une synthèse très claire de ces résultats et d'autres dans [137]). Pour n marches sur un réseau ($n \gg 1$) de k pas chacune (chaque pas se faisant vers un site voisin sur le réseau), Larralde *et al.* ont identifié trois régimes selon la valeur de k :

⁵Remarquons en effet que, par isotropie, $\langle M(\theta) \rangle$ ne dépend pas de θ . Et souvenons-nous aussi que dans un cas non isotrope, celui d'une excursion dans le demi-plan supérieur (traité au chapitre 6), $\langle M(\theta) \rangle$ (éq. (6.61)) a la même forme que la fonction support d'un cercle posé sur l'origine (p. 102). Ceci attire notre attention sur le fait qu'une courbe plane peut avoir la même fonction support qu'un cercle sans être un cercle. En mathématiques, c'est le problème des *courbes de largeur constante* dont les polygones de Reuleaux sont un exemple : ces courbes ont la même largeur sous tous les angles (et donc le même périmètre qu'un cercle de diamètre cette largeur, comme le démontra justement Barbier à l'aide de la formule de Cauchy [145]) mais ne sont pas circulaires et enclosent une aire moindre que le cercle de même diamètre.

⁶Nous remercions M. Hernán Larralde d'avoir attiré notre attention sur ce point.

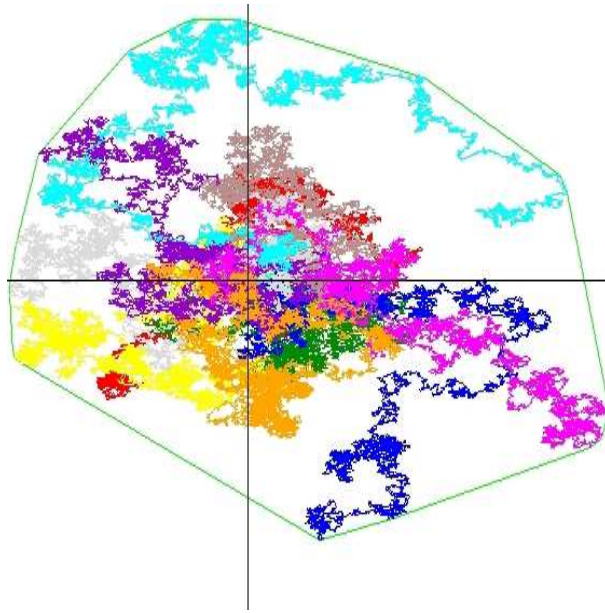


FIG. 9.6 – Enveloppe convexe de 10 chemins browniens ouverts. (Les chemins sont indépendants et issus de l'origine.)

- Régime I : pour k petit (disons $k \ll k_c$), les superpositions de marcheurs sur les sites déjà visités sont si nombreuses que tous les voisins de ces sites sont visités au pas suivant, si bien que le nombre moyen de sites distincts visités $\langle S_n(k) \rangle$ croît comme l'aire du disque de rayon k , c'est-à-dire proportionnellement à k^2 .
- Régime II : pour k « intermédiaire » ($k_c \ll k \ll k'_c$), le système entre dans une phase plus proprement diffusive, si bien que le nombre moyen de sites distincts visités $\langle S_n(k) \rangle$ croît comme l'aire du disque de rayon \sqrt{k} , soit proportionnellement à k , avec **un préfacteur dépendant de n , en l'occurrence $\ln n$** . Dans ce régime, Acedo et Yuste [137] décrivent le territoire exploré comme « une couronne de nature dendritique [zone dans laquelle les trajectoires se superposent peu] entourant un noyau [disque à l'intérieur duquel le taux de superposition est élevé]. » (*cf.* fig. 9.7).
- Régime III : pour k grand, les marcheurs sont loin les uns des autres et leurs trajectoires se superposent peu. Le nombre moyen de sites dis-

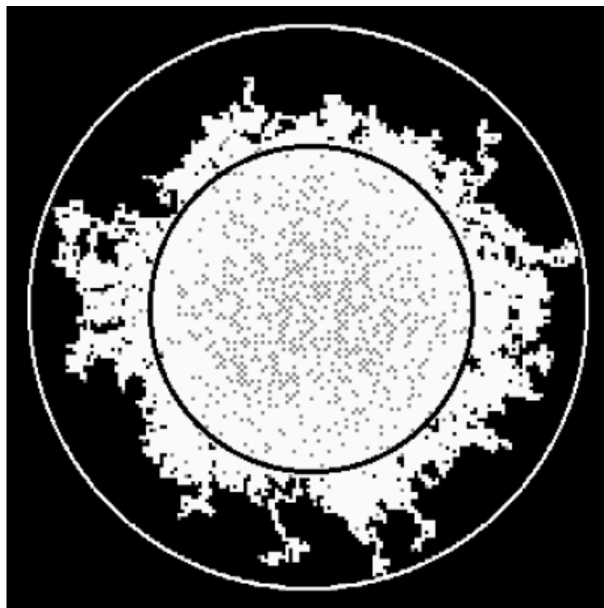


FIG. 9.7 – Figure de Acedo et Yuste [137] montrant les sites visités par 1000 marcheurs sur un réseau plan. Les sites visités sont en blanc, les sites non visités en noir et les points gris représentent la position des marcheurs. Les deux cercles sont centrés sur l’origine (point de départ des marcheurs), le blanc a pour rayon la distance du site visité le plus éloigné et le noir celle du site non visité le moins éloigné de l’origine.

tincts visités par les n marcheurs croît alors comme n fois le nombre moyen de sites distincts visités par un seul marcheur en k pas : $\langle S_n(k) \rangle \sim n \langle S_1(k) \rangle$, le comportement de $\langle S_1(k) \rangle$ pour k grand étant donné par $\frac{k}{\ln k}$.

Les temps k_c et k'_c de transition d’un régime à l’autre sont donnés (en dimension 2) par $k_c \sim \ln n$ et $k'_c \sim e^n$.

Le parallèle entre ce modèle de marcheurs sur réseau et le nôtre (n mouvements browniens plans de durée fixée) demande beaucoup de prudence⁷, mais nous pouvons remarquer que, dans le régime II, l’ensemble des sites distincts visités croît comme un cercle d’un rayon identique à celui de l’enveloppe

⁷En particulier du fait du passage à la fois à un temps continu (c’est-à-dire un nombre de pas infini) et à un espace continu (la constante de réseau devant tendre vers 0).

convexe de n mouvement browniens plans.

Pour conclure cette discussion, l'évocation d'un territoire exploré nous ramène à l'exemple d'application que nous citons au début de ce chapitre comme motivation de notre étude, à savoir le calcul du périmètre et de l'aire moyens du « domaine vital » d'un groupe d'animaux — ces grandeurs revêtant beaucoup d'importance lors de la création d'une réserve écologique. En effet, le coût économique et social de la désignation d'un territoire comme réserve naturelle vient souvent peser à la baisse sur la taille de celle-ci, tandis que la volonté que puissent y survivre les espèces présentes impose que la réserve soit dimensionnée de manière à pouvoir accueillir des populations viables des espèces en question. La manière dont la taille du domaine vital d'un groupe d'animaux croît avec le nombre d'individus dans le groupe joue donc un rôle crucial dans ce contexte. Or, nos résultats tendent à indiquer que cette croissance se fait logarithmiquement — c'est-à-dire plutôt lentement : si un groupe passe de 10 à 100 individus, son domaine vital se contentera de doubler, du moins dans les conditions simples compatibles avec notre modèle, à savoir l'indépendance des chemins individuels et l'absence de limitations (autres que spatiales) concernant par exemple la quantité de nourriture disponible.

Signalons, avant de conclure cette seconde partie, que les formules de Cauchy existent dans des dimensions supérieures à 2. Nous donnons en annexe (§ F) le calcul de la surface⁸ moyenne de l'enveloppe convexe en dimension 3. Les résultats que nous obtenons illustrent la manière dont on peut, à l'aide des formules de Cauchy et des statistiques de valeurs extrêmes en dimension 1, « gravir » l'échelle des dimensions dans l'étude de l'enveloppe convexe du mouvement brownien.

⁸En dimension 3, l'enveloppe convexe d'un chemin brownien est un polytope.

Conclusion

LES travaux présentés dans ce mémoire nous ont permis d'obtenir un certain nombre de résultats nouveaux relatifs aux statistiques d'extrêmes du mouvement brownien et à leurs applications.

- Dans le contexte de mouvements browniens linéaires d'abord, nous avons obtenu les résultats suivants :
 - la distribution du **temps d'atteinte du maximum sur un intervalle de temps fixé** pour des mouvements contraints : l'**excursion brownienne**, éq. (2.39) ; et le **méandre brownien**, éq. (2.50). Nous avons aussi calculé les comportements asymptotiques des expressions exactes obtenues (éq. (2.58), (2.60) et (2.65)) .
 - la distribution du **temps d'atteinte du maximum sur un intervalle de durée variable**, ainsi que ses comportements asymptotiques : maximum d'un **mouvement brownien libre avec ou sans dérive**, issu d'un point positif et observé **jusqu'à son premier passage en zéro** (éq. (3.12), (3.13), (3.15) et § 3.4).
 - la distribution du **temps d'atteinte du maximum de n mouvements browniens linéaires indépendants**, libres ou ponts, sur un intervalle de durée fixée (éq. (4.7) et (4.11))⁹.
- Relativement à l'enveloppe convexe du mouvement brownien plan, les résultats nouveaux sont les suivants :
 - dans le **cas classique d'un seul mouvement brownien plan : l'aire moyenne de l'enveloppe convexe d'un chemin fermé** (éq. (9.8)) ; le périmètre et l'aire moyens de l'enveloppe convexe d'un chemin fermé confiné au demi-plan supérieur¹⁰, ainsi que la distribution exacte de sa fonction support $M(\theta)$ pour tout $\theta \in [0, 2\pi]$ (éq.(6.34)).
 - dans le **contexte lui-même nouveau de plusieurs mouvements browniens plans**, tous issus de l'origine mais indépendants : les **valeurs moyennes du périmètre et de l'aire de l'enveloppe**

⁹Les résultats contenus dans ce mémoire permettent d'écrire également la distribution du temps d'atteinte du maximum pour n excursions browniennes indépendantes ou n méandres browniens indépendants, comme nous le signalions au § 4.3.

¹⁰Rappelons que dans ce cas les résultats sont les mêmes que pour un chemin brownien fermé non confiné, comme nous l'avons expliqué au § 9.3

convexe de n chemins, ouverts ou fermés, pour tout $n \geq 1$, ainsi que leurs comportements asymptotiques (éq. (9.12), (9.15), (9.21), (9.24), (9.13), (9.19), (9.28) et (9.29))¹¹.

Les premiers de ces résultats viennent compléter l'étude du temps d'atteinte du maximum pour un mouvement brownien sans dérive sur un intervalle de temps fixé. Seuls avaient été traités jusqu'à présent les cas du mouvement libre et du pont. Après ceux du méandre et de l'excursion, traités ici, pourront être étudiés les cas de mouvements avec dérive, ce qui permettra d'affiner les modèles en fonction des contextes, comme l'ont fait par exemple Bouchaud et Majumdar [36] pour un mouvement libre avec dérive sur un intervalle de temps fixé, dans un contexte de modélisation financière, et comme nous l'avons fait pour un mouvement considéré jusqu'à son premier passage en 0.

Les résultats relatifs à n mouvements browniens linéaires apportent quant à eux, en généralisant les précédents, de nouveaux outils de modélisation. Le temps nous a manqué pour en développer d'autres applications que celle présentée ici, où ils jouent néanmoins un rôle central dans l'étude de l'enveloppe convexe du mouvement brownien plan. De ce point de vue, une piste intéressante serait de voir comment se traduisent les différents types de contraintes imposées aux mouvements linéaires en termes de mouvements plans et d'enveloppe convexe de ceux-ci, et vice-versa. Par exemple, n mouvements plans indépendants mais tous contraints à demeurer à l'intérieur d'un disque de rayon R correspondront en projection à n mouvements linéaires indépendants demeurant entre $-R$ et R . Les maxima de ces n mouvements linéaires seront indépendants, et partageront la même distribution à support fini — outre l'intérêt en termes d'applications¹², il serait intéressant du point de vue théorique de voir comment les grandeurs relatives à l'enveloppe convexe des n chemins plans ainsi contraints se comparent à celles de l'enve-

¹¹Dans une note présentée en annexe (§ F), nous calculons aussi la surface moyenne de l'enveloppe convexe (qui est alors un polytope) de n mouvements browniens tridimensionnels indépendants, ouverts ou fermés.

¹²Dans le contexte écologique de l'application que nous détaillons au chapitre 9, il pourrait s'agir par exemple de calculer les grandeurs géométriques relatives au domaine vital d'un groupe d'animaux vivant sur une île, ou dans une zone entourée de montagnes.

loppe convexe de n points choisis indépendamment selon une distribution de type Weibull.

Les résultats de la première partie sous-tendent ceux de la seconde partie, à travers la méthode que nous avons développée, qui met au jour le lien étroit existant entre l'enveloppe convexe du mouvement brownien plan et les statistiques d'extrêmes du mouvement brownien linéaire (*via* les distributions du maximum et du temps d'atteinte du maximum) — tout comme elle met en évidence le lien entre statistiques des valeurs extrêmes de variables aléatoires standard et enveloppes convexes de nuages de points choisis de manière aléatoire et indépendante dans le plan. De ce point de vue, l'étude des enveloppes convexes de mouvements browniens plans développe parallèlement un cas particulier de statistique des extrêmes des processus aléatoires [95], pouvant être aisément comparé à la statistique des extrêmes « classique ». Cette dernière concerne en effet des variables aléatoires indépendantes identiquement distribuées, c'est-à-dire des points sur l'axe réel ; or le cas du maximum de n mouvements browniens indépendants peut être traité comme la donnée de n variables aléatoires indépendantes — les n maxima — distribuées selon une même loi — celle du maximum d'un mouvement brownien linéaire. C'est d'ailleurs pourquoi les valeurs moyennes du périmètre et de l'aire de l'enveloppe convexe de n mouvements browniens libres de durée T sont identiques¹³ à celles de l'enveloppe convexe d'un nuage de n points choisis indépendamment selon une distribution gaussienne bivariée ayant des marginales à moyenne nulle et de variance T . Néanmoins, la statistique des valeurs extrêmes dans le cas de processus aléatoires tels que le mouvement brownien s'enrichit de la question du temps d'atteinte d'un extremum global. Et nous avons vu l'importance que revêt précisément la distribution de ce temps d'atteinte pour le calcul de l'aire moyenne des enveloppes convexes de mouvements browniens plans.

Principale application présentée ici des résultats relatifs aux statistiques d'extrêmes des mouvements browniens linéaires, l'étude des valeurs moyennes du périmètre et de l'aire de l'enveloppe convexe de n mouvements browniens

¹³Rappelons que la distribution du maximum M d'un mouvement brownien libre de durée T est donnée par (*cf.* éq. (2.9)) : $\rho_1(M) = \sqrt{\frac{2}{\pi T}} e^{-\frac{M^2}{2T}}$.

plans nous semble en outre ouvrir une perspective relativement nouvelle, peu de travaux, à notre connaissance, ayant jusqu'ici porté sur plusieurs mouvements browniens plans simultanés et distincts, malgré les applications existantes, telle que celle présentée dans ce mémoire sur l'évaluation du *domaine vital* d'un groupe d'animaux. Le modèle traité ici est bien sûr, à bien des égards, très simple et nécessite d'être développé. Il serait également pertinent de savoir si l'ensemble des mesures nécessaires pour déterminer la constante de diffusion caractéristique d'une phase de déplacement des individus d'une espèce donnée sont moins contraignantes, moins coûteuses et donnent accès à des résultats plus précis (à travers nos formules) que celles consistant à faire la moyenne des périmètres et aires des polygones convexes obtenus à partir de l'enregistrement des trajectoires de plusieurs individus.

La considération de son application à l'écologie invite également à complexifier le modèle que nous proposons ici pour tenir compte des contraintes topographiques¹⁴ et des interactions entre les membres du groupe lors des phases de déplacement de type diffusif : il s'agit donc d'introduire des corrélations non triviales¹⁵ entre les n mouvements browniens plans considérés pour reproduire les phénomènes de mouvement collectif¹⁶.

Notre étude rejoint ici des sujets suscitant actuellement un intérêt grandissant parmi les physiciens [164, 165] — et plaçant une fois encore le mouvement brownien à la croisée de chemins biologiques, mathématiques et physiques.

¹⁴Par exemple, comme nous le mentionnions plus haut, pour des groupes d'animaux vivant en milieu insulaire, on peut considérer des chemins browniens restreints à une partie du plan.

¹⁵Par exemple en s'inspirant du modèle proposé par Vicsek *et al.* [163], dans lequel des particules autopropulsées se déplaçant à vitesse constante, réorientent à chaque pas de temps la direction de leur mouvement en fonction de la moyenne des directions suivies par leurs voisins et d'un terme aléatoire.

¹⁶Il sera alors en particulier intéressant de voir dans quelle mesure la croissance relativement lente de l'aire moyenne du domaine, en $\ln n$, demeure valide, selon les types d'interaction entre marcheurs.

Annexes

Annexe A

Note historique

LE MOUVEMENT BROWNIEN, s'il savait faire montre d'une quelconque fierté, pourrait se targuer d'être à la croisée de chemins scientifiques et mathématiques multiples.

Observée très tôt après l'invention du microscope au XVII^e siècle, cette agitation erratique, perpétuelle et sans cause apparente de particules en suspension dans un liquide inspira diverses hypothèses : manifestation fondamentale du vivant, convection due au pouvoir calorifère de l'éclairage utilisé sous le microscope, courants liés à l'évaporation... Le botaniste britannique Robert Brown, par les expériences qu'il mena en 1827 non seulement sur des particules contenues dans des grains de pollen mais aussi et surtout sur des particules à *coup sûr* inertes¹ [24], fit un pas décisif sur le chemin qui devait conduire à une compréhension physique du mouvement qui porte aujourd'hui son nom².

Cette compréhension est généralement datée de 1905, l'*annus mirabilis* d'Albert Einstein qui publia cette année-là, outre ses travaux fondateurs consacrés à la relativité restreinte et à l'effet photoélectrique, un article intitulé « *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*³ » [23]. Il

¹À l'époque, la possibilité que les grains de pollen pussent être animés n'était pas exclue.

²Une reconstitution des observations réalisées par Brown semble avoir mis un terme, en 1992, à la controverse sur la nature réellement « brownienne » du mouvement décrit par Brown. [166]

³Titre que l'on peut traduire ainsi : « Sur le mouvement, induit par la théorie cinétique

y établissait les fondations de la théorie physique du mouvement brownien, dans laquelle la probabilité $P(\vec{x}, t)$ de trouver une particule brownienne en un point \vec{x} au temps t obéit à l'équation de la diffusion :

$$\frac{\partial P}{\partial t} = D\nabla^2 P. \quad (\text{A.1})$$

L'article d'Einstein contenait, relativement à cette équation, une formule propre à marquer les esprits (ce qu'elle ne manqua pas de faire) :

$$D = \frac{RT}{\mathcal{N}} \frac{1}{6\pi\eta a}, \quad (\text{A.2})$$

D désignant donc le coefficient de diffusion du mouvement, R la constante des gaz parfaits, η la viscosité du liquide, a le rayon des particules en suspension, T la température et \mathcal{N} le nombre d'Avogadro — à une mesure expérimentale duquel l'équation d'Einstein ouvrait enfin la voie, près d'un siècle après qu'Avogadro en eut proposé l'existence [167]. C'est sur cette voie que s'engagea Jean Perrin dans ses expériences menées en 1908 [25, 168] et demeurées célèbres car, en confirmant la théorie d'Einstein (généralisée dans un second article paru en 1906 [169]), les résultats de Perrin ne faisaient rien moins qu'établir expérimentalement l'antique hypothèse atomiste de Leucippe et Démocrite [170].

Si ses travaux furent peut-être les plus décisifs, ce n'est pas seul qu'Einstein ouvrit la voie empruntée par Perrin. Un physicien australien, William Sutherland, parvenait lui aussi, en 1905 également [171], et de manière indépendante, à l'équation (A.2), de même que le physicien polonais Marian Smoluchowski qui publiait en 1906 deux articles [172, 173] dans lesquels il suivait un autre raisonnement, plus probabiliste, faisant le lien entre *marches aléatoires* et mouvement brownien⁴. Smoluchowski publiera plus tard un des premiers articles de synthèse sur le sujet, en 1916 [174], huit ans après la

moléculaire de la chaleur, des particules en suspension dans un liquide au repos ».

⁴Tous les points que nous évoquons ici sont relatés en détails dans l'article consacré par Bertrand Duplantier au mouvement brownien à l'occasion du centenaire de l'*annus mirabilis* d'Einstein, en 2005, année mondiale de la physique [18]. Nous nous en inspirons largement et librement, ainsi que, plus loin, du dialogue de Jean-Pierre Kahane entre un faux naïf et un vrai mathématicien au sujet du mouvement brownien [22].

démonstration par Paul Langevin [175] du fait que les approches de Smoluchowski et d'Einstein donnaient bien les mêmes résultats. Mais si le nom de Langevin fait indiscutablement partie de ceux qui sont indissociables de la théorie du mouvement brownien, c'est surtout parce qu'il eut l'idée lumineuse d'écrire, pour une particule brownienne, une équation dynamique certes newtonienne (au moins en apparence), mais intégrant un terme aléatoire, une force complémentaire *stochastique* [175].

Tant d'agitation ne pouvait passer longtemps inaperçue aux yeux des mathématiciens. Dès 1900, Louis Bachelier s'intéressait dans sa thèse [176] à la description des cours de la bourse de Paris par un modèle de marches aléatoires, avant de faire le lien avec l'équation de la diffusion. Mais surtout, les écrits de Perrin, et en particulier ses remarques sur l'impossibilité de donner une tangente à la courbe brownienne, inspirèrent au moins en partie les travaux de Norbert Wiener qui devait aboutir à la description mathématique du mouvement brownien, connue sous le nom de *processus de Wiener*, ainsi qu'à la *mesure de Wiener*, qui se révéla plus tard d'une belle efficacité dans la reformulation par Richard Feynman de la mécanique quantique en termes d'intégrales de chemins [177]. Entre temps, Leonard Ornstein et George Uhlenbeck, entre autres, avaient contribué au développement de la théorie du mouvement brownien [178, 179], ainsi que le mathématicien français Paul Lévy, qui publia plusieurs articles sur le sujet, et son livre intitulé *Processus stochastiques et mouvement brownien* [84].

Lévy s'intéressa à de nombreux aspects du mouvement brownien, tant linéaire⁵ que plan⁶, dont il découvrit de nombreuses propriétés. Un des résultats les plus célèbres qu'il obtint, la loi dite de l'arcsinus, donne ainsi la densité de probabilité du temps auquel un mouvement brownien linéaire, observé pendant une durée fixée, atteint sa valeur maximum sur la droite [40]. Lévy fut également un des premiers à envisager l'étude de l'enveloppe convexe⁷ du mouvement brownien plan, s'interrogeant en particulier sur la

⁵Position d'une particule brownienne sur une droite.

⁶Position d'une particule brownienne dans le plan, ses coordonnées étant deux mouvements browniens linéaires indépendants.

⁷Rappelons qu'un ensemble est dit convexe si, pour toute paire de points appartenant à cet ensemble, le segment les reliant appartient également à cet ensemble. En outre,

continuité et la dérivabilité du bord de cette enveloppe. Il entrerait ce faisant dans le champ des ensembles convexes aléatoires, c'est-à-dire des ensembles à coup sûr convexes mais dont la donnée est aléatoire : par exemple, jetant trois points au hasard dans le plan, le triangle les ayant pour sommets est un ensemble convexe aléatoire, ou encore, jetant une infinité de droites au hasard à travers le plan, les cellules ainsi dessinées composeront une mosaïque d'ensembles convexes aléatoires recouvrant le plan.

L'étude des ensembles convexes aléatoires, et plus généralement des ensembles aléatoires est liée, mathématiquement et historiquement, au développement de la probabilité géométrique et de la géométrie intégrale⁸. Le traitement du problème de l'aiguille de Buffon (ou jeu du joint couvert) par le mathématicien français E. Barbier [145] est un exemple de la première (la probabilité recherchée y étant déterminée en « comptant » le nombre moyen de droites d'une direction donnée intersectant une courbe) qui rejoint la seconde, par la démonstration d'un cas particulier d'une formule due à A. Cauchy [143] permettant de calculer une grandeur géométrique (en l'occurrence le périmètre ou l'aire) au moyen d'une intégration. Ces développements furent poursuivis en particulier par E. Barbier et J. Bertrand en France, et par J. J. Sylvester⁹ et M. W. Crofton outre-Manche [182].

La théorie des ensembles aléatoires s'est ensuite substantiellement développée au cours de la seconde moitié du XX^e siècle, à partir des travaux de G. Choquet [183] puis de D. G. Kendall [184] et de G. Matheron [180], et sous la pression stimulante des applications pratiques et expérimentales :

pour un ensemble quelconque, l'enveloppe convexe est le plus petit convexe le contenant intégralement, ce que l'on peut écrire, mathématiquement, comme l'intersection de tous les convexes le contenant.

⁸Dans sa préface au livre de G. Matheron [180], G. S. Watson définit d'ailleurs la probabilité géométrique comme la théorie des ensembles aléatoires. Nous nous inspirons ici librement de cette préface et de celle de Matheron.

⁹Sylvester est d'ailleurs l'auteur du problème suivant (qui porte son nom) : si l'on choisit quatre points selon une distribution uniforme sur un disque D du plan, quelle est la probabilité p_4 que ces quatre points soient les sommets d'un quadrilatère convexe [181] ? H.-J. Hilhorst *et al.* ont récemment montré que le problème de Sylvester, généralisé à n points, était lié au problème dit de l'accélération aléatoire, et ont utilisé des résultats récents concernant ce dernier pour déterminer le comportement asymptotique de p_n lorsque n devient grand.

étude des milieux poreux [185, 186], géologie, pétrographie et imagerie médicale à travers les techniques tomographiques et stéréologiques [187, 188]. Dans ce dernier cas, ce sont les possibilités offertes par la géométrie intégrale d'accéder indirectement (*via* une projection sur un espace de dimension inférieure) à des paramètres quantitatifs tels que le périmètre, la surface ou le nombre de composantes connexes, qui jouent un rôle crucial¹⁰.

¹⁰Les formules de Cauchy que nous utilisons dans cette thèse en sont un exemple.

Annexe B

Calcul d'une somme faisant intervenir une fonction de Bessel modifiée

I_ν étant la fonction de Bessel modifiée de première espèce d'ordre ν , nous cherchons à calculer, pour n entier, la somme :

$$\sum_{k=1}^{\infty} k^2 I_{kn}(u) \quad (\text{B.1})$$

Rappelons que $I_\nu(x)$ est solution de :

$$I_\nu'' + \frac{1}{x} I_\nu' - \left(1 + \frac{\nu^2}{x^2}\right) I_\nu = 0, \quad (\text{B.2})$$

ce que nous pouvons écrire sous la forme :

$$I_\nu'' + \frac{1}{x} I_\nu' - I_\nu = \frac{\nu^2}{x^2} I_\nu. \quad (\text{B.3})$$

Al-Jarrah *et al.* [189] ont par ailleurs démontré le résultat suivant :

$$\sum_{k=0}^{\infty} I_{nk}(u) = \frac{1}{2} I_0(u) + \frac{1}{2n} \sum_{k=0}^{n-1} e^{u \cos(2\pi \frac{k}{n})}. \quad (\text{B.4})$$

Partant de ce résultat, nous faisons passer dans le membre de droite le terme $k = 0$ de la somme à gauche puis dérivons par rapport à u :

$$\sum_{k=1}^{\infty} I'_{nk}(u) = -\frac{1}{2}I'_0(u) + \frac{1}{2n} \sum_{k=0}^{n-1} \cos\left(2\pi\frac{k}{n}\right) e^{u \cos(2\pi\frac{k}{n})} \quad (\text{B.5})$$

$$\sum_{k=1}^{\infty} I''_{nk}(u) = -\frac{1}{2}I''_0(u) + \frac{1}{2n} \sum_{k=0}^{n-1} \cos^2\left(2\pi\frac{k}{n}\right) e^{u \cos(2\pi\frac{k}{n})}. \quad (\text{B.6})$$

En combinant ces relations avec (B.3), nous allons pouvoir calculer (B.1) :

$$\begin{aligned} \sum_{k=1}^{\infty} \left(I''_{nk} + \frac{1}{u} I'_{nk} - I_{nk} \right) &= -\frac{1}{2} \left(I''_0 + \frac{1}{u} I'_0 - I_0 \right) \\ &+ \frac{1}{2n} \sum_{k=0}^{n-1} \left[\cos^2\left(2\pi\frac{k}{n}\right) + \frac{1}{u} \cos\left(2\pi\frac{k}{n}\right) - 1 \right] e^{u \cos(2\pi\frac{k}{n})} \end{aligned} \quad (\text{B.7})$$

Le premier terme du membre de droite est nul, par propriété de I_0 (éq. (B.3)). Puis, en utilisant (B.3) dans le membre de gauche et en réarrangeant les termes, nous obtenons :

$$\boxed{\sum_{k=1}^{\infty} k^2 I_{kn}(u) = \frac{u^2}{2n^3} \sum_{k=0}^{n-1} \left[\frac{1}{u} \cos \frac{2\pi k}{n} - \sin^2 \frac{2\pi k}{n} \right] e^{u \cos \frac{2\pi k}{n}}.} \quad (\text{B.8})$$

Signalons qu'il est possible d'obtenir ce résultat par un calcul¹ à peine plus long n'utilisant que les propriétés « standard » des fonctions de Bessel (c'est-à-dire sans avoir recours au théorème d'Al-Jarrah *et al.*).

¹Nous remercions M. Kirone Mallick de nous l'avoir indiqué.

Annexe C

Formules de Cauchy

Nous donnons ici une « démonstration » rapide des formules de Cauchy¹ :

$$L = \int_0^{2\pi} M(\theta) d\theta \quad (\text{C.1})$$

$$A = \frac{1}{2} \int_0^{2\pi} \left(M^2(\theta) - (M'(\theta))^2 \right) d\theta \quad (\text{C.2})$$

Nous considérons une courbe polygonale et, sans perte de généralité, examinons les intégrales apparaissant dans les formules de Cauchy pour une portion de la courbe correspondant à la situation représentée sur la figure C. Sur l'intervalle $\theta \in [-\phi_1, \phi_2]$, la valeur de la fonction support de la courbe polygonale sera donnée par A . En notant R la distance entre l'origine O et le sommet A , nous avons donc :

$$M(\theta) = R \cos \theta \quad (\text{C.3})$$

La première formule de Cauchy donne alors :

$$L = \int_{-\phi_1}^{\phi_2} M(\theta) d\theta = R(\sin \phi_1 + \sin \phi_2), \quad (\text{C.4})$$

ce qui correspond bien à la longueur de la courbe (à savoir $H_1A + AH_2$) entre H_1 et H_2 .

¹Nous remercions M. Deepak Dhar de nous avoir suggéré l'idée de cette démonstration.

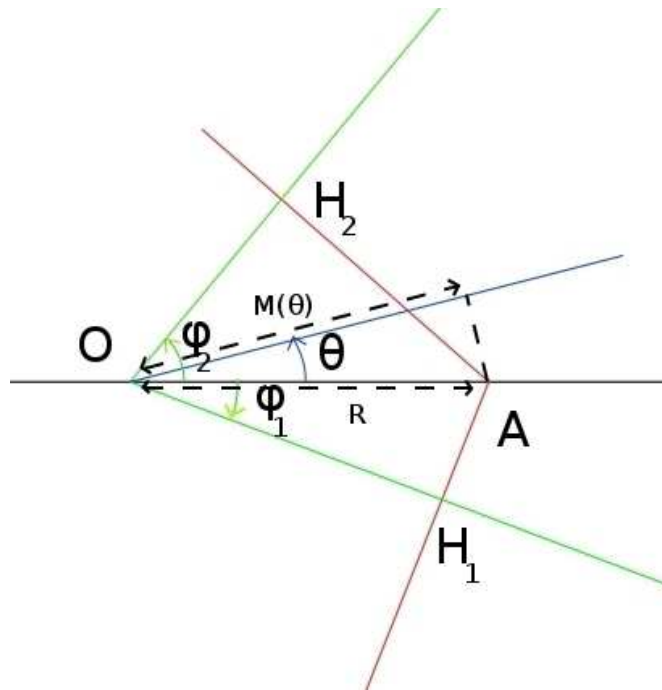


FIG. C.1 – Portion d’une courbe polygonale (en rouge) autour d’un de ses sommets (A), avec les droites passant par l’origine et perpendiculaires à la courbe avant A et après A (en vert), ainsi que la droite de direction θ (en bleu)

La seconde formule de Cauchy donne quant à elle :

$$\begin{aligned}
 A &= \frac{1}{2} \int_{-\phi_1}^{\phi_2} \left([M(\theta)]^2 - [M'(\theta)]^2 \right) d\theta \\
 &= \frac{R^2}{2} \int_{-\phi_1}^{\phi_2} [\cos^2 \theta - \sin^2 \theta] d\theta \\
 &= \frac{R^2}{2} (\sin \phi_2 \cos \phi_2 + \sin \phi_1 \cos \phi_1)
 \end{aligned}$$

ce qui correspond bien à l’aire du polygone OH_1AH_2 .

Ceci démontre les formules pour des polygones convexes contenant l’origine. Si l’origine est placée à l’extérieur, on peut voir que les signes des différents termes feront que les formules demeureront valides. Enfin, un passage à la limite continue permet d’obtenir le résultat pour les courbes lisses.

Annexe D

Chemins browniens multiples : analyse asymptotique

D.1 Chemins ouverts : périmètre moyen

M_n désignant l'abscisse maximum de n chemins browniens plans de durée T , nous notons $\text{Prob}(M_n \leq M) \equiv F_n(M)$. Celle-ci s'exprime en termes de la fonction de répartition $F(M)$ du maximum d'un mouvement brownien linéaire standard (éq. (2.15)) :

$$\begin{aligned} F_n(M) &= [F(M)]^n \\ &= \left[\frac{2}{\sqrt{\pi}} \int_0^{\frac{M}{\sqrt{2T}}} e^{-u^2} du \right]^n \\ &= \left[1 - \frac{2}{\sqrt{\pi}} \int_{\frac{M}{\sqrt{2T}}}^{\infty} e^{-u^2} du \right]^n \end{aligned} \tag{D.1}$$

Dans la limite où n et M deviennent très grands, (D.1) devient :

$$\begin{aligned} F_n(M) &= \exp \left[n \ln \left(1 - \frac{2}{\sqrt{\pi}} \int_{\frac{M}{\sqrt{2T}}}^{\infty} e^{-u^2} du \right) \right] \\ &\sim \exp \left[-n \frac{2}{\sqrt{\pi}} \int_{\frac{M}{\sqrt{2T}}}^{\infty} e^{-u^2} du \right] \end{aligned} \tag{D.2}$$

Une intégration par parties donne alors :

$$\begin{aligned} \int_{\frac{M}{\sqrt{2T}}}^{\infty} e^{-u^2} du &= \int_{\frac{M}{\sqrt{2T}}}^{\infty} \frac{1}{2u} 2ue^{-u^2} du \\ &= \sqrt{\frac{T}{2M^2}} e^{-\frac{M^2}{2T}} + o\left(\sqrt{\frac{T}{2M^2}} e^{-\frac{M^2}{2T}}\right) \end{aligned}$$

En insérant ceci dans l'équation (D.2) :

$$F_n(M) \sim e^{-\frac{2nTe^{-M^2/2T}}{M\sqrt{2\pi}}} \quad (\text{D.3})$$

$$\sim e^{-e^{-\frac{1}{2T}(M^2-2T \ln n)}} \quad (\text{D.4})$$

Ici nous écrivons :

$$\delta \equiv M - \sqrt{2T \ln n},$$

δ étant supposé disparaître à l'infini, hypothèse que nous pourrions vérifier *a posteriori*. Il s'ensuit que :

$$M^2 = 2T \ln n \left(1 + \frac{\delta}{\sqrt{2T \ln n}}\right)^2 \sim 2T \ln n \left(1 + \frac{2\delta}{\sqrt{2T \ln n}}\right)$$

D'où :

$$M^2 - 2T \ln n \sim 2\delta\sqrt{2T \ln n} = 2\sqrt{2T \ln n} (M - \sqrt{2T \ln n})$$

En insérant ceci dans (D.4), il vient :

$$F_n(M) \sim e^{-e^{-\sqrt{\frac{2 \ln n}{T}}(M - \sqrt{2T \ln n})}} \quad (\text{D.5})$$

Pour calculer le comportement asymptotique de la valeur moyenne de M_n quand n est grand, remarquons d'abord que :

$$\begin{aligned} \langle M_n \rangle &= \int_0^{\infty} M F_n'(M) dM \\ &\sim \int_A^{\infty} M F_n'(M) dM \end{aligned} \quad (\text{D.6})$$

où $A \gg 1$ est appelé à disparaître lors d'une étape ultérieure du calcul.

En combinant ceci à l'éq. D.5 :

$$\langle M_n \rangle \sim \int_A^{\infty} M \sqrt{\frac{2 \ln n}{T}} e^{-\sqrt{\frac{2 \ln n}{T}}(M - \sqrt{2T \ln n})} e^{-e^{-\sqrt{\frac{2 \ln n}{T}}(M - \sqrt{2T \ln n})}} dM \quad (\text{D.7})$$

En posant :

$$y = \sqrt{\frac{2 \ln n}{T}} \left(M - \sqrt{2T \ln n} \right)$$

nous obtenons :

$$\langle M_n \rangle \sim \int_{\sqrt{\frac{2 \ln n}{T}} (A - \sqrt{2T \ln n})}^{\infty} \left(\sqrt{2T \ln n} + y \sqrt{\frac{T}{2 \ln n}} \right) e^{-y} e^{-e^{-y}} dy \quad (\text{D.8})$$

Dans la limite où $n \rightarrow \infty$, ceci donne :

$$\begin{aligned} \langle M_n \rangle &\sim \int_{-\infty}^{\infty} \sqrt{2T \ln n} e^{-y} e^{-e^{-y}} dy \\ &\sim \sqrt{2T \ln n} \end{aligned} \quad (\text{D.9})$$

Ainsi, pour le périmètre moyen de l'enveloppe convexe de n chemins browniens plans (chacun de durée T), on obtient le comportement asymptotique suivant :

$$\langle L_n \rangle \sim 2\pi \sqrt{2T \ln n} \quad (\text{D.10})$$

comme nous l'avons écrit dans (9.13).

D.2 Chemins ouverts : aire moyenne

Nous voulons ici calculer le comportement asymptotique de l'aire moyenne de l'enveloppe convexe de n chemins browniens plans ouverts, indépendants et de durée T , lorsque n devient grand. Nous employons la formule de Cauchy (9.3) dans le cadre de mouvements isotropes :

$$\langle A_n \rangle = \pi \left(\langle M_n^2 \rangle - \langle [M'_n]^2 \rangle \right). \quad (\text{D.11})$$

Le premier terme à droite se calcule aisément à partir du paragraphe précédent (il suffit d'écrire M^2 au lieu de M dans l'éq. (D.7)) et donne :

$$\langle M_n^2 \rangle \sim 2T \ln n \quad (\text{D.12})$$

Pour parvenir au résultat (9.19), il reste ensuite à montrer que $\langle A_n \rangle$ est dominé par $\langle M_n^2 \rangle$.

Rappelons l'équation (9.15) :

$$\langle A_n \rangle = 4nT \sqrt{\pi} \int_0^\infty du u [\operatorname{erf}(u)]^{n-1} \left(u e^{-u^2} - g(u) \right), \quad (\text{D.13})$$

avec

$$g(u) = \frac{1}{2\sqrt{\pi}} \int_0^1 \frac{e^{-u^2/\tau} d\tau}{\sqrt{\tau(1-\tau)}}.$$

L'intégrale apparaissant dans l'éq. (D.13) est dominée par la contribution venant des grandes valeurs de u . Nous examinons donc dans cette limite l'intégrale définissant $g(u)$:

$$\int_0^1 d\tau \frac{1}{\sqrt{\tau(1-\tau)}} e^{-\frac{u^2}{\tau}} \quad (\text{D.14})$$

En posant $\tau = 1 - y$:

$$\int_0^1 d\tau \frac{1}{\sqrt{\tau(1-\tau)}} e^{-\frac{u^2}{\tau}} = \int_0^1 dy \frac{1}{\sqrt{y(1-y)}} e^{-\frac{u^2}{1-y}} \quad (\text{D.15})$$

$$\sim \int_0^1 dy \frac{1}{\sqrt{y(1-y)}} e^{-u^2(1+y)} \quad (\text{D.16})$$

$$\sim e^{-u^2} \int_0^1 dy \frac{1}{\sqrt{y(1-y)}} e^{-yu^2} \quad (\text{D.17})$$

Nous écrivons maintenant :

$$z = u\sqrt{y}.$$

Ceci conduit à :

$$\begin{aligned} \int_0^1 d\tau \frac{1}{\sqrt{\tau(1-\tau)}} e^{-\frac{u^2}{\tau}} &\sim e^{-u^2} \int_0^u \frac{2z dz}{u^2} \frac{1}{\sqrt{\frac{z^2}{u^2} \left(1 - \frac{z^2}{u^2}\right)}} e^{-z^2} \\ &\sim \frac{2e^{-u^2}}{u} \int_0^u dz \frac{1}{\sqrt{1 - \frac{z^2}{u^2}}} e^{-z^2} \\ &\sim \frac{2e^{-u^2}}{u} \int_0^u dz \left(1 + \frac{z^2}{2u^2}\right) e^{-z^2} \\ &\sim \frac{e^{-u^2}}{u} \sqrt{\pi} \operatorname{erf}(u) + \frac{e^{-u^2}}{u} \int_0^u dz \frac{z^2 e^{-z^2}}{u^2} \end{aligned}$$

Une intégration par partie montre que le second terme du côté droit est $o\left(\frac{e^{-u^2}}{u^2}\right)$, d'où il suit que :

$$\int_0^1 d\tau \frac{1}{\sqrt{\tau(1-\tau)}} e^{-\frac{u^2}{\tau}} \sim \frac{\sqrt{\pi} e^{-u^2}}{u}. \quad (\text{D.18})$$

En revenant à l'éq. (D.13), où l'intégrale est dominée par la contribution des grandes valeurs de u , nous avons alors :

$$\langle A_n \rangle = 4nT \sqrt{\pi} \int_0^\infty du u [\text{erf}(u)]^{n-1} \left(ue^{-u^2} - g(u) \right) \quad (\text{D.19})$$

$$\sim 4nT \sqrt{\pi} \int_0^\infty du u [\text{erf}(u)]^{n-1} \left(ue^{-u^2} - \frac{\sqrt{\pi} e^{-u^2}}{u} \right) \quad (\text{D.20})$$

$$\sim 4nT \sqrt{\pi} \int_0^\infty du u^2 [\text{erf}(u)]^{n-1} e^{-u^2} \quad (\text{D.21})$$

$$\sim \pi \langle M_n^2 \rangle \quad (\text{D.22})$$

$$\sim 2\pi T \ln n. \quad (\text{D.23})$$

D.3 Chemins fermés : périmètre moyen

Le calcul est analogue à celui de la section D.1, à ceci près que la fonction de répartition F_n du maximum M de n ponts browniens linéaires est donnée (*cf.* éq. (4.3) et (2.21)) par :

$$F_n(M) = \left[1 - e^{-\frac{2M^2}{T}} \right]^n \quad (\text{D.24})$$

Dans la limite où n et M deviennent très grands, nous avons alors :

$$\begin{aligned} F_n(M) &\sim \exp \left[n \ln \left(1 - e^{-\frac{2M^2}{T}} \right) \right] \\ &\sim e^{-e^{-\frac{2}{T}(M^2 - \frac{T}{2} \ln n)}} \end{aligned} \quad (\text{D.25})$$

Nous voyons que nous retrouvons ici l'équation (D.4), en remplaçant T par $\frac{T}{4}$. Nous en déduisons le résultat donné en (9.28) :

$$\langle L_n \rangle \sim \pi \sqrt{2T \ln n}. \quad (\text{D.26})$$

D.4 Chemins fermés : aire moyenne

Nous voulons ici calculer le comportement asymptotique de l'aire moyenne de l'enveloppe convexe de n chemins browniens plans fermés, indépendants et de durée T , lorsque n devient grand. Le calcul est analogue à celui de la section D.2. Nous appliquons la formule de Cauchy dans le contexte d'échantillons isotropes (9.3) :

$$\langle A_n \rangle = \pi \left(\langle M_n^2 \rangle - \langle [M_n']^2 \rangle \right). \quad (\text{D.27})$$

Le premier terme dans le membre de droite se calcule aisément à partir des sections D.3 et D.1 :

$$\langle M_n^2 \rangle \sim \frac{T}{2} \ln n \quad (\text{D.28})$$

Nous montrons maintenant que $\langle A_n \rangle$ est dominé par $\langle M_n^2 \rangle$. L'équivalent de l'éq. (D.13) pour des chemins fermés est :

$$\langle A_n \rangle = \frac{2nT}{\sqrt{\pi}} \int_0^\infty du u^2 \left(1 - e^{-u^2}\right)^{n-1} \left(ue^{-u^2} - g(u)\right), \quad (\text{D.29})$$

avec :

$$g(u) = \frac{1}{8} \int_0^1 d\tau \frac{e^{-\frac{u^2}{4\tau(1-\tau)}}}{\sqrt{\tau(1-\tau)}}.$$

L'intégrale apparaissant dans l'éq. (D.29) est dominée par les contributions venant des grandes valeurs de u . Nous examinons donc $g(u)$ dans cette limite.

$$g(u) = \frac{1}{8} \int_0^1 d\tau \frac{e^{-\frac{u^2}{4\tau(1-\tau)}}}{\sqrt{\tau(1-\tau)}} \quad (\text{D.30})$$

$$= \frac{1}{4} \int_0^{\frac{1}{2}} d\tau \frac{e^{-\frac{u^2}{4\tau(1-\tau)}}}{\sqrt{\tau(1-\tau)}} \quad (\text{D.31})$$

Nous écrivons maintenant :

$$\tau = \frac{1}{2} - z.$$

Ce qui conduit à :

$$\frac{1}{4} \int_0^{\frac{1}{2}} d\tau \frac{e^{-\frac{u^2}{4\tau(1-\tau)}}}{\sqrt{\tau(1-\tau)}} = \frac{1}{2} \int_0^{\frac{1}{2}} dz \frac{e^{-\frac{u^2}{1-4z^2}}}{\sqrt{1-4z^2}}$$

En posant :

$$v = 2uz,$$

il vient :

$$\begin{aligned} \frac{1}{2} \int_0^{\frac{1}{2}} dz \frac{e^{-\frac{u^2}{1-4z^2}}}{\sqrt{1-4z^2}} &= \frac{1}{4} \int_0^u \frac{dv}{u} \frac{e^{-\frac{u^2}{1-\frac{v^2}{u^2}}}}{\sqrt{1-\frac{v^2}{u^2}}} \\ &\sim \frac{1}{4} \frac{e^{-u^2}}{u} \int_0^u dv e^{-v^2} \left(1 + \frac{v^2}{2u^2}\right) \\ &\sim \frac{e^{-u^2}}{8u} \sqrt{\pi} \operatorname{erf}(u) + o\left(\frac{e^{-u^2}}{u^2}\right) \end{aligned}$$

En revenant à l'éq. (D.29), où l'intégrale est dominée par la contribution des grandes valeurs de u , nous avons alors :

$$\begin{aligned} \langle A_n \rangle &= \frac{2nT}{\sqrt{\pi}} \int_0^\infty du u^2 \left(1 - e^{-u^2}\right)^{n-1} \left(ue^{-u^2} - g(u)\right) \\ &\sim \frac{2nT}{\sqrt{\pi}} \int_0^\infty du u^2 \left(1 - e^{-u^2}\right)^{n-1} \left(ue^{-u^2} - \frac{\sqrt{\pi}e^{-u^2}}{8u}\right) \\ &\sim \frac{2nT}{\sqrt{\pi}} \int_0^\infty du u^3 \left(1 - e^{-u^2}\right)^{n-1} e^{-u^2} \\ &\sim \pi \langle M_n^2 \rangle \\ &\sim \frac{\pi T}{2} \ln n. \end{aligned} \tag{D.32}$$

Annexe E

Simulations numériques

Les simulations de mouvement brownien utilisées pour illustrer nos résultats sont réalisées à l'aide d'un algorithme élémentaire. Nous partons de l'équation de Langevin du mouvement brownien libre de constante de diffusion $D = \frac{1}{2}$:

$$\frac{dx}{dt} = \eta(t), \quad (\text{E.1})$$

où η est un bruit blanc gaussien, de moyenne nulle et de corrélation δ :

$$\langle \eta(t) \rangle = 0 \quad (\text{E.2})$$

$$\langle \eta(t)\eta(t') \rangle = \delta(t - t'), \quad (\text{E.3})$$

où δ est la fonction *delta* de Dirac.

L'idée pour la simulation est d'utiliser un générateur aléatoire nous fournissant une suite de nombres aléatoires r distribués selon une gaussienne normalisée (nous avons utilisé la routine « gasdev » de [190]) et, bien sûr, de discrétiser le mouvement :

$$x(t + \Delta t) = x(t) + r\sqrt{\Delta t}. \quad (\text{E.4})$$

Le choix de Δt et de la durée totale T du mouvement fixe le nombre de pas de la marche aléatoire ainsi générée, et plus Δt est petit, plus la marche « ressemble » à un mouvement brownien¹. Dans les simulations numériques de ce mémoire, nous prenons la plupart du temps $\Delta t = 10^{-4}$ et $T = 1$.

¹Il peut être nécessaire, bien sûr, de quantifier cette « ressemblance ».

Il est ensuite possible, à partir de ce mouvement fondamental, de générer les différents types de mouvements linéaires conditionnés, par exemple des ponts browniens, avec la construction classique :

$$y(t) = x(t) - t \frac{x(T)}{T} \quad (\text{E.5})$$

ou des excursions browniennes, avec la construction de Vervaat [44] à partir d'un pont brownien, telle que représentée sur la figure E.1.

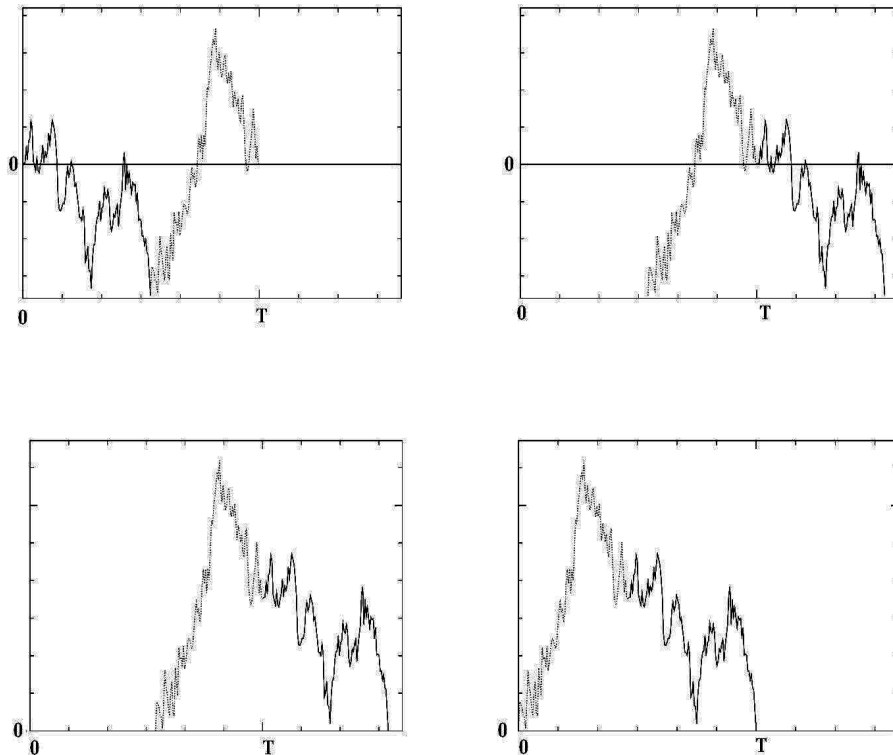


FIG. E.1 – La construction de Vervaat : du pont brownien à l'excursion brownienne

Le balayage de Graham

Pour calculer les enveloppes convexes des chemins browniens générés numériquement, nous avons implémenté un des premiers algorithmes valides, proposé en 1976 par R. Graham [70].

Du point de vue du calcul de l'enveloppe convexe, une marche aléatoire peut être envisagée comme un simple nuage de points, et c'est d'ailleurs dans ce cadre général qu'opère l'algorithme de Graham.

La première étape consiste à ordonner les points, en fonction d'un point P_0 choisi tel qu'il sera à coup sûr un des sommets de l'enveloppe convexe. Il suffit pour cela de prendre par exemple le point le plus bas, celui ayant la plus petite ordonnée. Faisant passer par P_0 une droite horizontale, les $n - 1$ autres points sont ensuite étiquetés en fonction de l'angle que forme avec cette droite la droite les reliant à P_0 — par exemple dans l'ordre croissant. L'idée est alors de balayer l'ensemble des points, en suivant leur indexation, et d'éliminer les points responsables de concavités. Ceci se fait en examinant la direction du « virage » dessiné par trois points consécutifs $P_i P_{i+1} P_{i+2}$ (en commençant par P_0, P_1, P_2) — mathématiquement, il s'agit d'examiner le signe du produit vectoriel $\vec{P_i P_{i+1}} \wedge \vec{P_i P_{i+2}}$ — :

- si le virage est à gauche, les trois points en question ne dessinent pas une concavité et nous pouvons avancer d'un cran pour examiner $P_{i+1} P_{i+2} P_{i+3}$;
- s'il est orienté à droite, il dessine une concavité et nous pouvons éliminer P_{i+1} : celui-ci ne sera pas sur l'enveloppe convexe du nuage. Il faut alors revenir partiellement en arrière et examiner le virage dessiné par $P_{i-1} P_i P_{i+2}$ — en effet celui-ci pourrait être concave alors que $P_{i-1} P_i P_{i+1}$ ne l'était pas.

Pour un nombre fini de points, nous parcourons l'ensemble de l'échantillon, n'ayant plus à la fin que les sommets de l'enveloppe convexe du nuage, telle qu'illustré sur les figures de la page suivante où, pour simplifier, le nuage ne comprend que cinq points, étiquetés a, b, c, d, e .

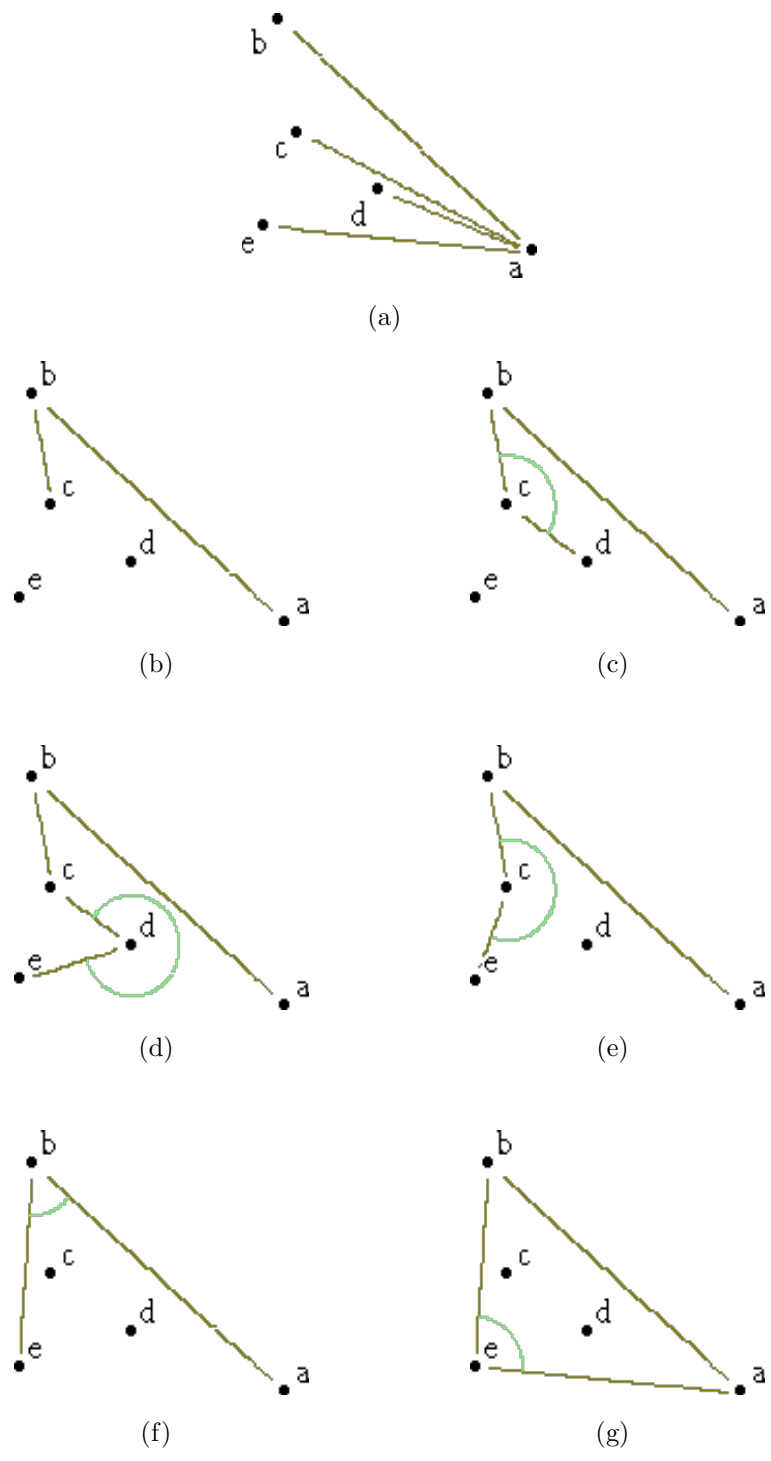


FIG. E.2 – L’algorithme de Graham sur un nuage de cinq points : (a) on réordonne les points, puis (b)–(g) on examine les points par triplets consécutifs ; (b), (c), (f) et (g) : s’ils forment une configuration convexe, on avance d’un cran ; (d) et (e) : s’ils forment une configuration concave, on abandonne le point médian et on revient un cran en arrière. Enfin (g), on obtient ainsi l’enveloppe convexe de l’échantillon.

Annexe F

Note : enveloppes convexes aléatoires en 3 dimensions

L'approche reposant sur la fonction support pour l'étude des enveloppes convexes aléatoires peut être étendue à des dimensions supérieures à 2 en utilisant les formules de Cauchy correspondantes. Nous décrivons ici cette transposition et donnons un aperçu de quelques résultats nouveaux obtenus par ce moyen.

F.1 Formule de Cauchy en trois dimensions

Dans son article original [143], Cauchy donne la formule suivante pour les surfaces bidimensionnelles plongées dans l'espace à trois dimensions.

Soient

- S l'aire d'une surface \mathcal{S} fermée et convexe de l'espace à 3 dimensions (par exemple un polyèdre ou un cube),
- $A(\phi, \theta)$ l'aire de la projection de \mathcal{S} sur le plan $\mathcal{H}_{\phi, \theta}$ ayant pour vecteur orthogonal $(\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \phi)$

Alors :

$$S = \frac{1}{\pi} \int_{-\pi}^{\pi} \int_0^{\pi} A(\phi, \theta) \sin \phi \, d\phi \, d\theta \quad (\text{F.1})$$

L'idée est ensuite d'appliquer une formule de Cauchy « bidimensionnelle » (7.11) à la projection de \mathcal{S} et de l'associer à (F.1) afin d'obtenir S .

Pour ce faire, nous choisissons une base du plan $\mathcal{H}_{\phi,\theta}$:

$$\begin{aligned}\mathbf{e}_1 &= (\cos \phi \cos \theta, \cos \phi \sin \theta, -\sin \phi) \\ \mathbf{e}_2 &= (-\sin \theta, \cos \theta, 0).\end{aligned}$$

L'action de la projection sur le plan envoie le vecteur $(x, y, z) \in \mathbb{R}^3$ sur le point :

$$(x \cos \phi \cos \theta + y \cos \phi \sin \theta - z \sin \phi, y \cos \theta - x \sin \theta) \in \mathcal{H}_{\phi,\theta}.$$

En posant maintenant :

$$\begin{aligned}M_{\phi,\theta}(\psi) = \\ \max_{i \in I} \{ (x_i \cos \phi \cos \theta + y_i \cos \phi \sin \theta - z_i \sin \phi) \cos \psi + (y_i \cos \theta - x_i \sin \theta) \sin \psi \},\end{aligned}\tag{F.2}$$

où I paramétrise les points de \mathcal{S} , nous pouvons effectivement associer (F.1) et (7.11) ; nous obtenons :

$$S = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \int_0^{\pi} d\phi \sin \phi \int_0^{2\pi} d\psi \left(M_{\phi,\theta}^2(\psi) - \left[\frac{dM_{\phi,\theta}}{d\psi} \right]^2 \right)\tag{F.3}$$

Si nous illustrons ceci par un exemple simple, la sphère de rayon r centrée sur l'origine :

$$M_{\phi,\theta}(\psi) = r \Rightarrow S = 4\pi r^2,$$

nous retrouvons bien le résultat attendu.

F.2 Application de la formule de Cauchy « tridimensionnelle » aux enveloppes convexes aléatoires

Il n'est pas difficile de montrer que l'enveloppe convexe de la projection d'un objet tridimensionnel sur un plan, enveloppe qui est une courbe sur une surface bidimensionnelle, est elle-même la frontière de la projection de l'enveloppe convexe (qui est une surface) de l'objet initial.

Dès lors, nous procédons pour des échantillons tridimensionnels d'une manière analogue à ce que nous avons fait dans le cas d'échantillons bivariés et nous prenons les moyennes sur l'ensemble des réalisations à l'intérieur de la formule de Cauchy :

$$\langle S \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \int_0^{\pi} d\phi \sin \phi \int_0^{2\pi} d\psi \left(\langle M_{\phi,\theta}^2(\psi) \rangle - \left\langle \left[\frac{dM_{\phi,\theta}}{d\psi} \right]^2 \right\rangle \right). \quad (\text{F.4})$$

De nouveau, à l'instar de ce qui se passait en dimension 2, les expressions se simplifient pour des échantillons dont la distribution est à symétrie sphérique :

$$\langle S \rangle = 4\pi \left(\langle M_{0,0}^2(0) \rangle - \left\langle \left[\frac{dM_{0,0}}{d\psi}(0) \right]^2 \right\rangle \right) \quad (\text{F.5})$$

$$= 4 \langle S^{(2)} \rangle, \quad (\text{F.6})$$

avec $S^{(2)}$ l'aire de l'enveloppe convexe de l'échantillon bidimensionnel défini par les coordonnées x et y des points de l'échantillon tridimensionnel.

F.3 Résultats

Pour des échantillons isotropes tels que des chemins browniens simples, ouverts ou fermés, dans l'espace, nous obtenons immédiatement, à partir des résultats des sections précédentes, l'aire moyenne du plus petit polytope convexe contenant les chemins tridimensionnels :

- Pour 1 chemin ouvert de durée T en 3 dimensions :

$$\boxed{\langle S \rangle = 2\pi T}, \quad (\text{F.7})$$

qui s'obtient à partir du résultat bidimensionnel $\langle S^{(2)} \rangle = \frac{\pi T}{2}$ (éq. (9.6)).

- Pour 1 chemin fermé de durée T en 3 dimensions :

$$\boxed{\langle S \rangle = \frac{4}{3}\pi T} \quad (\text{F.8})$$

qui s'obtient à partir du résultat bidimensionnel $\langle S^{(2)} \rangle = \frac{\pi T}{3}$ (éq. (9.8)).
et plus généralement :

- Pour n chemins ouverts de durée T en 3 dimensions :

$$\boxed{\langle S_n \rangle = 16nT \sqrt{\pi} \int_0^\infty du u [\operatorname{erf}(u)]^{n-1} \left(ue^{-u^2} - g(u) \right)} \quad (\text{F.9})$$

avec

$$g(u) = \frac{1}{2\sqrt{\pi}} \int_0^1 \frac{e^{-u^2/t} dt}{\sqrt{t(1-t)}}.$$

(s'obtient à partir du résultat bidimensionnel (éq. (9.15))

- Pour n chemins fermés de durée T en 3 dimensions :

$$\boxed{\langle S_n^{(c)} \rangle = 2\pi T \left[\sum_{k=1}^n \frac{1}{k} - \frac{n}{3} + \frac{1}{2} \sum_{k=2}^n (-1)^k f(k) \right]} \quad (\text{F.10})$$

avec

$$f(k) = \binom{n}{k} (k-1)^{-3/2} \left(k \tan^{-1}(\sqrt{k-1}) - \sqrt{k-1} \right).$$

(s'obtient à partir du résultat bidimensionnel (éq. (9.24))

Nous trouvons tout aussi facilement les comportements asymptotiques suivant pour n grand :

- Pour n chemins ouverts de durée T en 3 dimensions :

$$\boxed{\langle S_n \rangle \sim 8\pi T \ln n} \quad (\text{F.11})$$

- Pour n chemins fermés de durée T en 3 dimensions :

$$\boxed{\langle S_n^{(c)} \rangle \sim 2\pi T \ln n} \quad (\text{F.12})$$

Nous trouvons donc à nouveau une croissance relativement lente (logarithmique) de la surface de l'enveloppe convexe de n chemins browniens dans l'espace. Notons cependant que l'équivalent de cette surface pour des chemins plans est le périmètre de l'enveloppe convexe bidimensionnel et non son aire ; or ce périmètre croît, lui, en $\sqrt{\ln n}$ (éq. (9.13)) — donc moins vite que son équivalent tridimensionnel.

Annexe G

Articles

Distribution of the time at which the deviation of a Brownian motion is maximum before its first-passage time

Julien Randon-Furling and Satya N Majumdar

Laboratoire de Physique Théorique et Modèles Statistiques,
Université Paris-Sud, CNRS UMR 8626, 91405 Orsay Cedex, France
E-mail: julien.randon-furling@lptms.u-psud.fr and majumdar@lptms.u-psud.fr

Received 16 August 2007

Accepted 20 September 2007

Published 12 October 2007

Online at stacks.iop.org/JSTAT/2007/P10008

[doi:10.1088/1742-5468/2007/10/P10008](https://doi.org/10.1088/1742-5468/2007/10/P10008)

Abstract. We calculate analytically the probability density $P(t_m)$ of the time t_m at which a continuous-time Brownian motion (with and without drift) attains its maximum before passing through the origin for the first time. We also compute the joint probability density $P(M, t_m)$ of the maximum M and t_m . In the driftless case, we find that $P(t_m)$ has power-law tails: $P(t_m) \sim t_m^{-3/2}$ for large t_m and $P(t_m) \sim t_m^{-1/2}$ for small t_m . In the presence of a drift towards the origin, $P(t_m)$ decays exponentially for large t_m . The results from numerical simulations are in excellent agreement with our analytical predictions.

Keywords: Brownian motion, first-passage problems, extreme value problems

Contents

1. Introduction	2
2. Driftless case	4
3. In the presence of a negative drift	9
4. Summary and conclusion	12
Acknowledgments	12
References	12

1. Introduction

In this paper, we derive the probability distribution of a random variable associated with a Brownian motion, namely the time at which a Brownian motion attains its maximum value before it crosses the origin for the first time. This random variable appears quite naturally in different problems, such as in queueing theory and in the evolution of stock prices in finance.

Let us first consider, for example, a single-server discrete-time queueing process, modelled as a simple random walk [1, 2] via

$$l_n = l_{n-1} + \xi_n,$$

where l_n is the length of the queue at time n and the ξ_n are independent and identically distributed random variables each taking values $+1$ with probability p (signifying the arrival of a new customer), -1 with probability q (indicating the departure of an already served customer) or 0 with probability $(1 - p - q)$. In the queueing language, this is referred to as the Geo/Geo/1 queue [2].

Given l_0 , one calls the *busy period* the period at the end of which the queue becomes empty for the first time (see figure 1): during such a period, the server always has some customers to serve. It is then natural to enquire about the time at which the queue is at its longest during the busy period. In the random walk model where the queue length l_n is the position of the walker at time step n , this amounts to investigating the time at which the position of the walker (initially positive) is farthest from the origin before it crosses the origin for the first time.

Another area where the same variable appears quite naturally is in the evolution of stock prices in finance. The evolution of a stock price S_n with time n is often modelled by the exponential of a random walk [3, 4]. Starting from its initial value S_0 the price evolves with time stochastically. A natural question for an agent holding this stock is: what is a suitable time for selling this stock? If the stock price goes below a threshold, say R , it is too risky to wait any longer. Thus an agent can wait at most up to the time at which the ratio S_n/R crosses the level 1 from above. Within this time, the ratio S_n/R will achieve its maximum at some intermediate time which is clearly the best time to sell the stock. Assuming that the random variable $l_n = \ln(S_n/R)$ performs a random walk starting from its initial value $l_0 = \ln(S_0/R) > 0$, one then wants to calculate the probability distribution

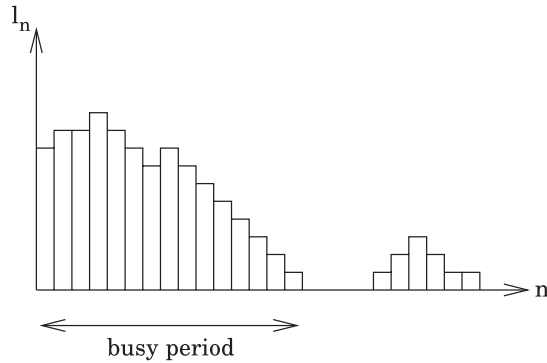


Figure 1. Queue with busy period.

of the time at which the random walk is farthest from the origin till its first-passage time through the origin, i.e., till the time at which $S_n/R = 1$, i.e., $\ln(S_n/R) = 0$ indicating the first-passage through the origin.

In this paper, we will consider a further simplified case, namely a continuous-time Brownian motion as opposed to the discrete-time random walk in the above two problems. For a continuous-time Brownian motion we calculate explicitly, using path-integral methods, the probability density of the time t_m at which a Brownian motion $x(t)$ (starting from $x(0) = x_0 > 0$) is farthest from the origin before it crosses the origin for the first time. Algorithmically speaking, for each sample of the Brownian motion starting at x_0 we stop when it crosses the origin for the first time, say at time $t = t_f$, and locate the time $0 \leq t_m < t_f$ at which the Brownian motion achieves its maximum value. Note that both t_f (the first-passage time) and t_m vary from one sample to another. We repeat this many times and then construct a histogram of the t_m values, which gives its probability density function $P(t_m)$. Even though the discrete-time problem is more relevant, we expect the continuous-time result to provide the right asymptotics for the discrete problem. As we will see below, the continuous-time problem, though still non-trivial, is easier to handle analytically.

Note that for a Brownian motion or a Brownian bridge over a *fixed* time interval $[0, T]$, the probability density $P(\tau, T)$ of the time τ at which the process attains its maximum is well known [5]. For example, for a zero-drift Brownian motion over $[0, T]$ starting at the origin, the probability density $P(\tau, T) = (1/T)g(\tau/T)$, where $g(x) = 1/[\pi\sqrt{x(1-x)}]$ for $x \in [0, 1]$ [5]. On the other hand, for a Brownian bridge over the fixed interval $[0, T]$ and starting at the origin, the probability density of τ is uniform, $P(\tau, T) = 1/T$ for $0 \leq \tau \leq T$ [5]. In contrast, in our case, the Brownian motion is not over a fixed time interval, but rather over a variable time interval $[0, t_f]$ where the upper edge t_f is the first-passage time which itself is a random variable [6] and hence varies from sample to sample.

The statistical properties of the functionals (such as the area, the maximum, etc) of a Brownian motion or its variants (such as a bridge, excursion, meander, etc) over a fixed time interval have many applications in physics, graph theory, and computer science, and they have been studied extensively (for recent reviews on Brownian functionals see [7, 8]). In particular, the area under a Brownian excursion or meander has found many recent applications in problems as diverse as fluctuating interfaces [9], graph enumeration [10],

lengths of internal paths in rooted planar trees [8, 11] or cost functions in data storage via the ‘linear probing with hashing’ algorithm [8, 12]. Similarly, the statistical properties of functionals of Brownian motion restricted up to its first-passage time (usually referred to as ‘first-passage functionals’) also have various applications, and have appeared recently in many different contexts [1, 15, 16], including the computation of the time period of oscillation of an undamped particle in a random potential [14] and the determination of the distribution of the lifetime of a comet in the solar system [8, 13]. The probability density of the area swept by an initially positive Brownian motion till its first-passage time was computed exactly in [15], with an application to queuing theory. In this paper, our focus is on the random variable t_m which, though not quite a functional in the strict sense, is an important random variable associated with such a Brownian motion restricted up to its first-passage time.

In [15], the authors also computed directly the probability density $P(M)$ of the maximum M of a Brownian motion (starting at $x_0 > 0$) before its first-passage time through the origin, via a ‘backward’ Fokker–Planck method and showed that it has a power-law behaviour $P(M) = x_0/M^2$ where $M \geq x_0$. In this paper, we extend this work using a path decomposition method that allows us to obtain the joint probability density $P(M, t_m)$ of the maximum M and the time t_m at which the maximum occurs before the first-passage time. By integrating over M , we then get the ‘marginal’ $P(t_m)$, i.e., the probability density of t_m . We calculate $P(t_m)$ explicitly both for a driftless and drifted Brownian motion. We also compare the results of numerical simulations to our analytical predictions and find excellent agreement.

2. Driftless case

We consider a continuous-time Brownian motion evolving via $dx/dt = \xi(t)$, where $\xi(t)$ is a white noise with $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$. We start by recalling the quick derivation of $P(M)$ given in [15].

Let $q(x)$ be the probability that a Brownian particle starting from $x \in [0, M]$ exits the interval for the first time through 0, i.e., the probability that the maximum before the first-passage time is less than or equal to M . Writing $\phi_{\Delta t}(\Delta x)$ for the distribution function of a Brownian displacement in the time interval Δt , we have

$$q(x) = \int q(x + \Delta x)\phi_{\Delta t}(\Delta x) d\Delta x. \quad (1)$$

Expanding $q(x + \Delta x)$ for small values of Δx , and using the fact that in the absence of drift the mean value of Δx is 0, one finds that q satisfies

$$\frac{d^2 q}{dx^2} = 0, \quad q(0) = 1, \quad q(M) = 0, \quad (2)$$

whose solution is

$$q(x) = 1 - \frac{x}{M}. \quad (3)$$

As mentioned in its definition, it can easily be seen that $q(x)$ also corresponds to the probability that the maximum before the first-passage time is less than or equal to M ; therefore, differentiating equation (3) with respect to M gives the probability

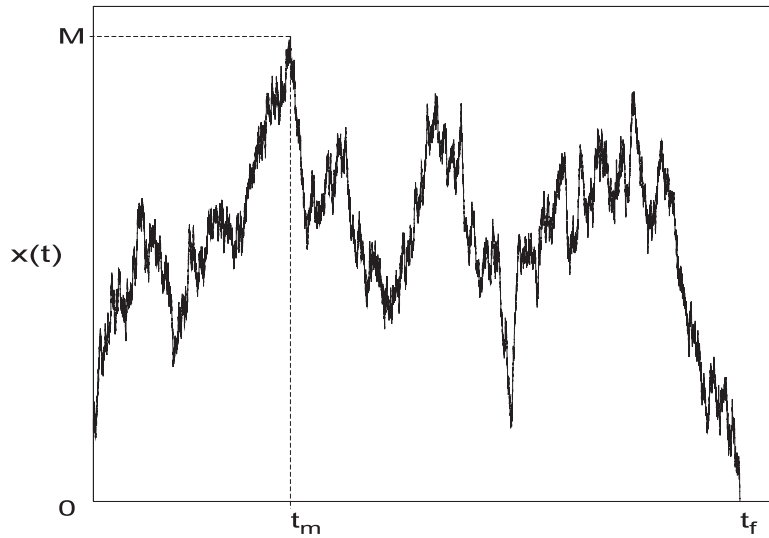


Figure 2. Maximum before the first passage through the origin for the zero-drift case.

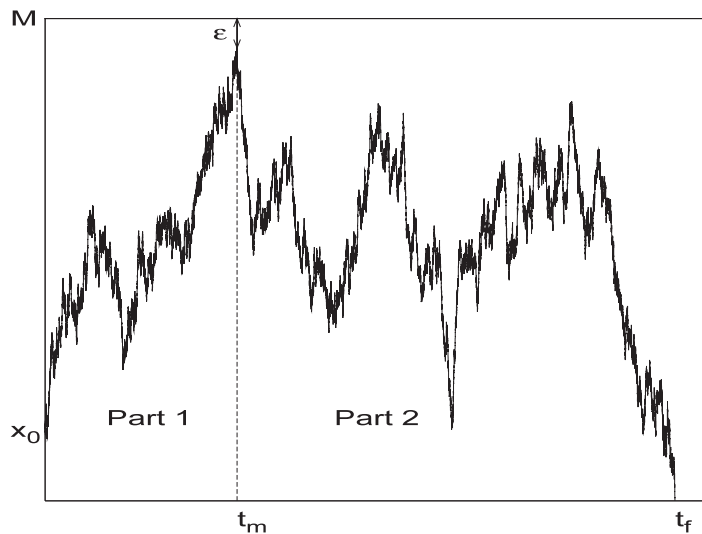


Figure 3. Splitting probabilities.

density of M :

$$P(M) = \frac{x}{M^2}. \tag{4}$$

To compute the joint probability density $P(M, t_m)$ we proceed as follows. We first assume that the maximum occurs at t_m and then we split the Brownian path into two parts (before/after t_m , as shown in figure 2) and determine the weights of a path's left-hand side and right-hand side separately. Note that due to the *Markovian* property of the Brownian path, once the position of the walker is specified at t_m , the weights of the left and the right parts become completely *independent* and the total weight is just proportional to the product of the weights of the two separate parts. For the left part,

we have a process that propagates from x_0 at $t = 0$ to M at $t = t_m$ without crossing the level M in $[0, t_m]$ (since M is the maximum) and the level 0 (the origin). For the right part, the process propagates from the value M at $t = t_m$ to 0 at $t = t_f$, where $t_f \geq t_m$ without crossing the level M and the level 0 in between. We need to be careful, however, because, as pointed out in [9], a Brownian walker that crosses a given level once crosses it infinitely many times immediately after the first crossing. It is therefore impossible to enforce the constraint $x(t_m) = M$ and simultaneously force the motion to stay below M before or after t_m (for a lattice walk, one does not have this problem since the lattice constant provides a natural cut-off). Following the method used in [9], we introduce a cut-off ϵ by imposing $x(t_m) = M - \epsilon$ and consider all paths having a maximum less than or equal to M and passing through $M - \epsilon$ at $t = t_m$ (as shown in figure 3). We compute their weight and then let ϵ go to 0 eventually.

On the right side of $t = t_m$: we have to determine the weight of a path that starts at $M - \epsilon$ and exits for the first time the interval $[0, M]$ through 0. This is given by equation (3):

$$q(M - \epsilon) = \frac{\epsilon}{M}. \quad (5)$$

On the left side of $t = t_m$: we use a path-integral treatment with the Feynman–Kac formula (as in [9]) giving the weight of a path in terms of the propagator $\langle x_0 | e^{-\hat{H}t_m} | M - \epsilon \rangle$, where $\hat{H} = -\frac{1}{2}(\partial^2/\partial x^2) + V(x)$ with $V(x)$ a square well having infinite barriers at $x = 0$ and M , and $V(x) = 0$ for $0 < x < M$ (the infinite barriers at $x = 0$ and M enforce the condition that the path can penetrate neither at $x = 0$ nor at $x = M$). The normalized eigenfunctions of \hat{H} labelled by the integer $n = 1, 2, 3, \dots$ are $\psi_n(x) = \sqrt{2/M} \sin(n\pi x/M)$ with the associated eigenvalues $E_n = n^2\pi^2/2M^2$. The eigenfunction $\psi_n(x)$ vanishes at both ends $x = 0$ and M of the box. The propagator can be easily evaluated in this eigenbasis, $\langle x_0 | e^{-\hat{H}t_m} | M - \epsilon \rangle = \sum_{n=1}^{\infty} \psi_n(x_0)\psi_n(M - \epsilon) e^{-E_n t_m}$ and one gets:

$$\langle x_0 | e^{-\hat{H}t_m} | M - \epsilon \rangle = \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \epsilon)}{M}\right) \exp\left(-\frac{n^2\pi^2}{2M^2}t_m\right). \quad (6)$$

In the limit when $\epsilon \rightarrow 0$, we get to leading order

$$\langle x_0 | e^{-\hat{H}t_m} | M - \epsilon \rangle = \frac{2\pi}{M^2} \epsilon \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) \exp\left(-\frac{n^2\pi^2}{2M^2}t_m\right) + O(\epsilon^2). \quad (7)$$

Taking the product of equations (5) and (7), we get the total weight of the path, to leading order in small ϵ ,

$$P(M, t_m; \epsilon) \propto \epsilon^2 \frac{\pi}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) \exp\left(-\left(\frac{n^2\pi^2}{2M^2}\right)t_m\right). \quad (8)$$

The proportionality constant is set by using the normalization condition:

$$\int_{x_0}^{\infty} dM \int_0^{\infty} dt_m P(M, t_m; \epsilon) = 1.$$

It is easy to show that the proportionality constant $A(\varepsilon) = \varepsilon^{-2}$. Thus, in the limit $\varepsilon \rightarrow 0$, we finally obtain

$$P(M, t_m) = \frac{\pi}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) \exp\left(-\frac{n^2 \pi^2}{2M^2} t_m\right). \quad (9)$$

As a first check, let us show that $\int_0^{\infty} dt_m P(M, t_m) = x_0/M^2$, thus recovering the marginal $P(M)$ of the maximum in equation (4). Integrating over t_m , we get

$$P(M) = \frac{2}{\pi M} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sin\left(\frac{n\pi x_0}{M}\right) = \frac{x_0}{M^2}, \quad (10)$$

where the last identity can be found (and derived easily) in [17].

Finally, from equation (9) an integration over M (note that M varies from x_0 to ∞) yields the desired marginal $P(t_m)$:

$$\begin{aligned} P(t_m) &= \pi \int_{x_0}^{\infty} \frac{dM}{M^3} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) \exp\left(-\frac{n^2 \pi^2}{2M^2} t_m\right) \\ &= \frac{1}{\pi t_m} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \int_0^{n\pi} du \cos(u) \exp\left(-\frac{u^2}{2x_0^2} t_m\right). \end{aligned} \quad (11)$$

The sum in equation (11) can be expressed in terms of a known special function, and we get

$$P(t_m) = \frac{1}{2\pi t_m} \left[\pi - \int_0^{\pi} \vartheta_4\left(\frac{y}{2}, e^{-y^2(t_m/2x_0^2)}\right) dy \right], \quad (12)$$

where $\vartheta_4(z, q)$ is the fourth of Jacobi's theta functions [19]. Subsequently one can obtain the large and small t_m asymptotics of $P(t_m)$ from the exact expression in equation (12). *Large- t_m asymptote.* We first consider the case when $t_m \gg x_0^2$. Changing variables in equation (11) through $z = \sqrt{t_m/2x_0^2} u$ and letting $z \rightarrow 0$ gives for $t_m \gg x_0^2$:

$$P(t_m) \approx \frac{x_0 \log 2}{t_m^{3/2}} \sqrt{\frac{1}{2\pi}}. \quad (13)$$

Small- t_m asymptote. In the opposite limit $t_m \ll x_0^2$, we start from equation (9) and first take a Laplace transform:

$$\begin{aligned} \int_0^{\infty} dt_m e^{-st_m} P(M, t_m) &= \frac{2}{\pi M} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{n}{(n^2 + 2M^2 s/\pi^2)} \sin\left(\frac{n\pi x_0}{M}\right) \\ &= \frac{\sinh(x_0 \sqrt{2s})}{M \sinh(M \sqrt{2s})}, \end{aligned}$$

where the sum of the series can be found in [17]. Letting s become much larger than x_0^{-2} and M^{-2} , we obtain

$$\int_0^{\infty} dt_m e^{-st_m} P(M, t_m) \approx \frac{e^{-\sqrt{2s}(M-x_0)}}{M},$$

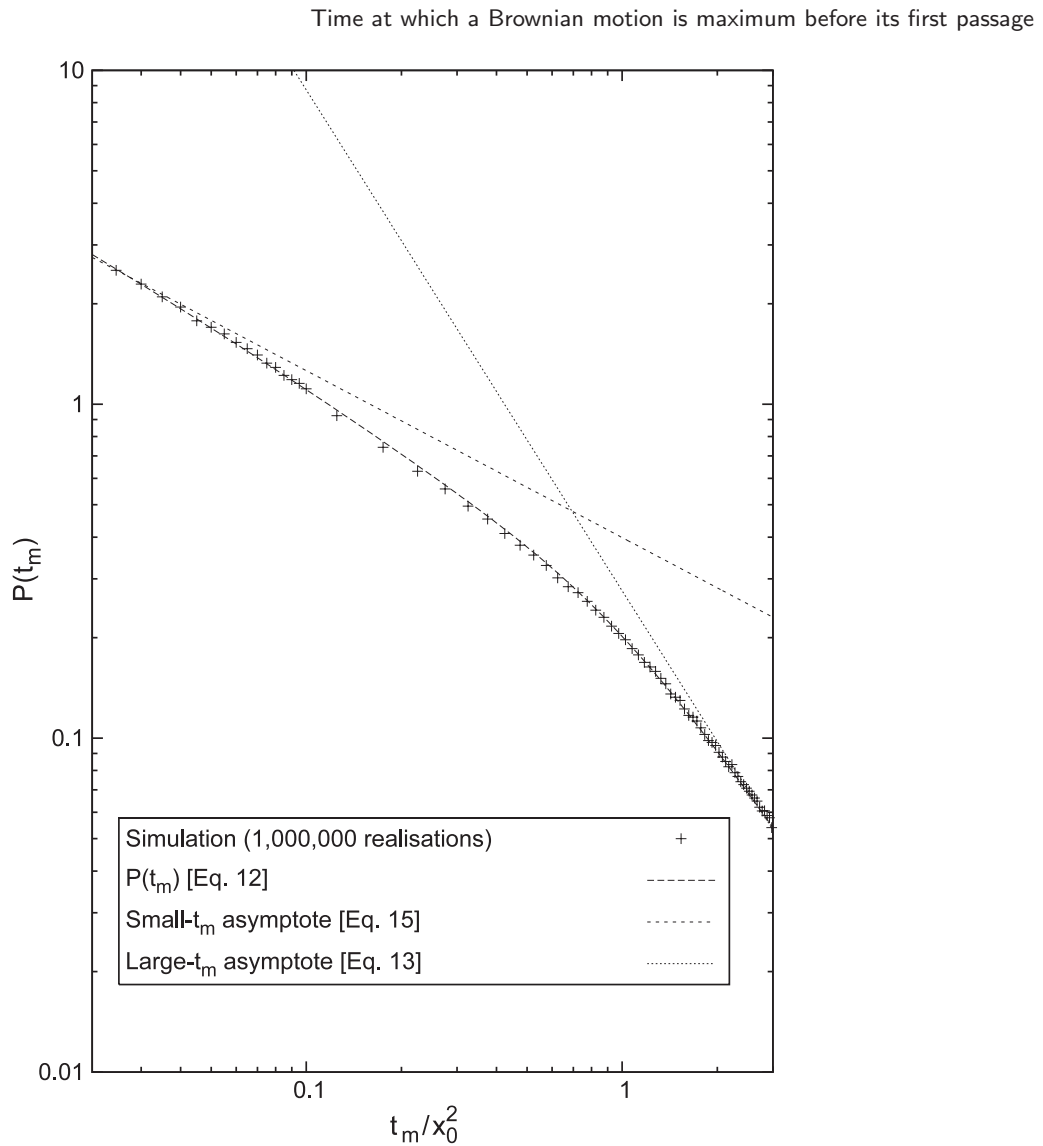


Figure 4. The probability density $P(t_m)$ and its asymptotes for the driftless case. In the simulation, x_0 was set to 1.

which, after the Laplace inversion [18] yields

$$P(M, t_m) \approx \frac{t_m^{-3/2}}{\sqrt{2\pi}} \frac{(M - x_0)}{M} \exp\left(-\frac{(M - x_0)^2}{2t_m}\right). \quad (14)$$

Integrating over M gives for $t_m \ll x_0^2$

$$P(t_m) \approx \frac{1}{x_0 \sqrt{2\pi t_m}}. \quad (15)$$

Thus, $P(t_m)$ has power-law behaviour at both large and small tails. For large t_m , the probability density falls off as $P(t_m) \sim t_m^{-3/2}$, whereas for small t_m it diverges as $P(t_m) \sim t_m^{-1/2}$. The exact analytical form of $P(t_m)$ and its asymptotes from equations (12), (13), (15) are plotted (using Mathematica) in figure 4 together with the

points obtained from the numerical simulation (with 1 000 000 realizations). They are in good agreement with each other.

3. In the presence of a negative drift

We now consider a Brownian motion in the presence of a drift μ . For $\mu > 0$, it is clear that the walker will escape to ∞ with a nonzero probability. This means that with a finite probability $t_m \rightarrow \infty$. Therefore this case is not much of interest in the present context. Instead, we focus here on the opposite case, where the drift is towards the origin, i.e., $\mu < 0$. The Langevin equation describing the motion becomes

$$\frac{dx}{dt} = -|\mu| + \xi(t),$$

where $\xi(t)$ is the Gaussian white noise with $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$. We use the same strategy as in the driftless case, i.e., splitting the motion into two independent parts (before and after t_m) and introducing a small cut-off ε .

On the right-hand side. Letting, as in the driftless case, $q(x)$ be the probability that a Brownian particle starting from $x \in [0, M]$ exits the interval for the first time through 0, we have as before

$$q(x) = \int q(x + \Delta x) \phi_{\Delta t}(\Delta x) d\Delta x. \quad (16)$$

In the presence of a drift, the mean value of Δx is no longer 0 and one can easily show that the analogue of equation (2) now reads

$$q''(x) - 2|\mu|q(x) = 0, \quad q(0) = 1, \quad q(M) = 0. \quad (17)$$

The solution is

$$q(x) = \frac{\sinh(|\mu|(M - x))}{\sinh(|\mu|M)} e^{|\mu|x}, \quad (18)$$

and so we have

$$q(M - \varepsilon) = \frac{\sinh(|\mu|\varepsilon)}{\sinh(|\mu|M)} e^{|\mu|(M - \varepsilon)}. \quad (19)$$

As in the driftless case, the probability density of M can be obtained by differentiation of equation (18) with respect to M , as was done in [15]:

$$P_d(M) = \frac{|\mu| \sinh(|\mu|x)}{\sinh^2(|\mu|M)} e^{|\mu|x}, \quad (20)$$

where we have added the subscript 'd' to indicate that the density corresponds to the drifted case.

On the left-hand side. We use the same path-integral method as in the driftless case. The weight of a path is now proportional to

$$\begin{aligned} & \exp \left[-\frac{1}{2} \int_0^{t_m} d\tau \left(\frac{dx}{d\tau} + |\mu| \right)^2 \right] \\ & = \exp \left[-\frac{|\mu|^2}{2} t_m - |\mu| \int_0^{t_m} d\tau \frac{dx}{d\tau} \right] \exp \left[-\frac{1}{2} \int_0^{t_m} d\tau \left(\frac{dx}{d\tau} \right)^2 \right]. \end{aligned} \quad (21)$$

The position of the Brownian particle at $t = 0$ and t_m is known, so we can substitute $(M - \varepsilon) - x_0$ for $\int_0^{t_m} d\tau (dx/d\tau)$ in the first exponential factor on the right-hand side of equation (21).

The propagator for the drifted case will therefore be equal to that for the driftless case (given in equation (6)) multiplied by the factor $\exp[|\mu|x_0 - (|\mu|^2/2)t_m - |\mu|(M - \varepsilon)]$, and will be given by

$$\exp\left[|\mu|x_0 - \frac{|\mu|^2}{2}t_m - |\mu|(M - \varepsilon)\right] \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \varepsilon)}{M}\right) \exp\left(-\frac{n^2\pi^2}{2M^2}t_m\right). \quad (22)$$

As in the driftless case, we multiply the weights of the left and right side of t_m derived above (equations (19) and (22)), and take the $\varepsilon \rightarrow 0$ limit to obtain

$$P_d(M, t_m) = \frac{|\mu|Me^{|\mu|x_0 - (|\mu|^2/2)t_m}}{\sinh(|\mu|M)} P(M, t_m), \quad (23)$$

where $P(M, t_m)$ is the joint density for the driftless case given in equation (9). Once again, by integrating over t_m , one can recover the marginal probability density of the maximum $P_d(M)$ derived originally in [15]. On the other hand, integrating over M gives the marginal $P_d(t_m)$. We were not able to derive a compact expression for $P_d(t_m)$ as in the driftless case, though the asymptotes of $P_d(t_m)$ can be derived explicitly, as shown below.

Small- t_m asymptote. From equation (23), we can derive very quickly the behaviour of $P_d(t_m)$ when $t_m \ll x_0^2$. Substituting in equation (23) the asymptotic result for the driftless case from equation (14) and integrating over M , we get

$$P_d(t_m) \sim \frac{|\mu|e^{|\mu|x_0 - (|\mu|^2/2)t_m}}{\sinh(|\mu|x_0)\sqrt{2\pi t_m}}. \quad (24)$$

Thus for small t_m , $P_d(t_m)$ diverges as $t_m^{-1/2}$, as in the driftless case.

Large- t_m asymptote. To study the behaviour of $P_d(t_m)$ when $t_m \gg x_0^2$, we start from the following expression for $P_d(t_m)$:

$$\begin{aligned} P_d(t_m) &= \int_{x_0}^{\infty} dM P_d(M, t_m) \\ &= \int_{x_0}^{\infty} dM \frac{|\mu|Me^{|\mu|x_0 - (|\mu|^2/2)t_m}}{\sinh(\mu M)} P(M, t_m) \\ &= \int_{x_0}^{\infty} dM \frac{|\mu|\pi e^{|\mu|x_0 - (|\mu|^2/2)t_m}}{\sinh(|\mu|M)M^2} \sum_{n=1}^{\infty} (-1)^{n+1} n \sin\left(\frac{n\pi x_0}{M}\right) \exp\left(-\frac{n^2\pi^2}{2M^2}t_m\right). \end{aligned} \quad (25)$$

The series in equation (25) is dominated by the first term ($n = 1$) for large t_m . Hence, retaining only the $n = 1$ term and making a change of variable $y = 1/M$ in the integral, we get

$$\pi|\mu|e^{|\mu|x_0 - (|\mu|^2/2)t_m} \int_0^{1/x_0} dy \frac{\sin(\pi x_0 y)}{\sinh(|\mu|/y)} \exp\left(-\frac{\pi^2 y^2 t_m}{2}\right).$$

Time at which a Brownian motion is maximum before its first passage

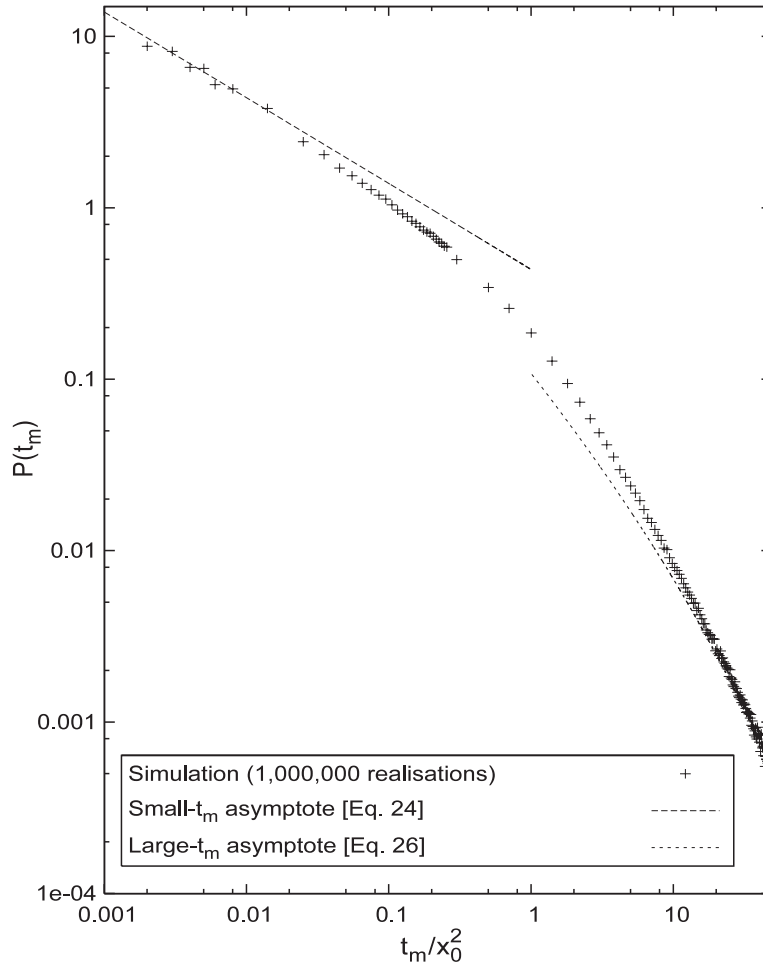


Figure 5. The simulated probability density $P(t_m)$ and its asymptotes in the presence of a drift towards the origin ($|\mu| = 0.1$). In the simulation, x_0 was set to 1.

For large t_m , the most important contribution to the integral comes from the small- y regime. Expanding the sin and sinh functions and keeping only the leading-order term reduces the integral to

$$2\pi^2 x_0 |\mu| e^{|\mu|x_0 - (|\mu|^2/2)t_m} \int_0^{1/x_0} dy y \exp\left(-t_m \left(\frac{\pi^2}{2} y^2 + \frac{|\mu|}{yt_m}\right)\right).$$

Letting $h(y) = (\pi^2/2)y^2 + |\mu|/yt_m$, we next use the saddle point method to obtain the leading term via minimizing the function h and get

$$P_d(t_m) \sim \left[2\sqrt{\frac{2}{3}}\pi^{5/6} x_0 |\mu|^{4/3} e^{|\mu|x_0}\right] t_m^{-5/6} \exp\left(-\frac{|\mu|^2}{2}t_m - \frac{3}{2}(|\mu|\pi)^{2/3} t_m^{1/3}\right). \quad (26)$$

Thus, as expected, the density $P_d(t_m)$ has an exponential decay for large t_m in the presence of a negative drift. Figure 5 shows a plot of the asymptotes (equations (24) and (26)) together with the data from numerical simulation (1000 000 realizations with $|\mu| = 0.1$).

4. Summary and conclusion

In summary, we have obtained an exact expression for the probability density of the time at which a Brownian motion attains its maximum before passing through the origin for the first time, and studied the tails of this probability density both for the driftless and for the drifted Brownian motion. This was done by first computing the joint distribution $P(M, t_m)$ of the maximum M attained and the time t_m at which it is attained. In the context of the queuing theory, the result that $P(t_m)$ decreases monotonically with increasing t_m suggests that the beginning of a busy period is more likely to be the time at which a queue is at its longest.

It would be interesting to derive the explicit results, obtained here by the path-integral method, from the general theory of filtrations in Brownian motion developed recently in [20, 21].

It would also be of interest to extend this calculation to the discrete-time random walk case, which remains a real challenge.

Acknowledgments

We thank A Comtet, M J Kearney and P L Krapivsky for useful discussions. We also thank M Yor for useful discussions and for pointing out references [20, 21].

References

- [1] Kearney M J, 2004 *J. Phys. A: Math. Gen.* **37** 8421
- [2] Asmussen S, 2003 *Applied Probability and Queues* 2nd edn (New York: Springer)
- [3] Williams R J, 2006 *Introduction to the Mathematics of Finance* (Providence, RI: AMS)
- [4] Yor M, 2000 *Exponential Functionals of Brownian Motion and Related Topics* (Berlin: Springer)
See also Comtet A, Monthus C and Yor M, 1998 *J. Appl. Probab.* **35** 255
- [5] Feller W, 1968 *An Introduction to Probability Theory and its Applications* (New York: Wiley)
- [6] Redner S, 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
- [7] Comtet A, Desbois J and Texier C, 2005 *J. Phys. A: Math. Gen.* **38** R341
- [8] For a short review on Brownian functionals and their applications see Majumdar S N, 2005 *Curr. Sci.* **89** 2075 [[cond-mat/0510064](https://arxiv.org/abs/cond-mat/0510064)]
- [9] Majumdar S N and Comtet A, 2005 *J. Stat. Phys.* **119** 777
Majumdar S N and Comtet A, 2004 *Phys. Rev. Lett.* **92** 225501
- [10] For an extensive review on the area under Brownian motion and its variants, see Janson S, 2007 *Prob. Surveys* **4** 80
- [11] Takács L, 1991 *Adv. Appl. Probab.* **23** 557
Takács L, 1995 *J. Appl. Probab.* **32** 375
- [12] Flajolet P, Poblete P and Viola A, 1998 *Algorithmica* **22** 490
- [13] Hammersley J M, 1961 *Proc. 4th Berkeley Symp. on Math. Stat. and Prob.* vol 3 (Berkeley: University of California Press) p 17
- [14] Dean D S and Majumdar S N, 2001 *J. Phys. A: Math. Gen.* **34** L697
- [15] Kearney M J and Majumdar S N, 2005 *J. Phys. A: Math. Gen.* **38** 4097
- [16] Kearney M J, Majumdar S N and Martin R J, 2007 *Preprint* [0706.2038](https://arxiv.org/abs/0706.2038)
- [17] Gradshteyn I S and Ryzhik I M, 2000 *Table of Integrals, Series and Products* 6th edn (New York: Academic)
- [18] Bateman H, 1954 *Tables of Integral Transforms* (New York: McGraw-Hill)
- [19] Abramowitz M and Stegun I A, 1973 *Handbook of Mathematical Functions* (New York: Dover)
- [20] Mansuy R and Yor M, 2006 *Lecture Notes in Mathematics* vol 1873 (Berlin: Springer)
- [21] Nikeghbali A and Yor M, 2006 *Ill. J. Math.* **50** 791

On the time to reach maximum for a variety of constrained Brownian motions

Satya N Majumdar¹, Julien Randon-Furling¹, Michael J Kearney²
and Marc Yor³

¹ Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud. Bât. 100, 91405, Orsay Cedex, France

² Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK

³ Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VI et VII, 4 Place Jussieu-Case 188, F-75252, Paris Cedex 05, France

Received 11 March 2008, in final form 18 June 2008

Published 1 August 2008

Online at stacks.iop.org/JPhysA/41/365005

Abstract

We derive $P(M, t_m)$, the joint probability density of the maximum M and the time t_m at which this maximum is achieved, for a class of constrained Brownian motions. In particular, we provide explicit results for excursions, meanders and reflected bridges associated with Brownian motion. By subsequently integrating over M , the marginal density $P(t_m)$ is obtained in each case in the form of a doubly infinite series. For the excursion and meander, we analyse the moments and asymptotic limits of $P(t_m)$ in some detail and show that the theoretical results are in excellent accord with numerical simulations. Our primary method of derivation is based on a path-integral technique; however, an alternative approach is also outlined which is founded on certain ‘agreement formulae’ that are encountered more generally in probabilistic studies of Brownian motion processes.

PACS numbers: 02.50.-r, 05.40.-a, 05.40.Jc

1. Introduction

Brownian motion (the Wiener process) is the most important and widely studied continuous-time stochastic process and, as such, has generated a huge literature. Despite this attention, however, it is still possible to identify problems relating to Brownian motion which are relatively easy to pose but not that well understood. Such problems are often directly linked to areas of application in the physical or social sciences, wherein their solution is of immediate relevance.

Within this overall context, there has been a recent renewal of interest in studying functionals of *constrained* Brownian motion. This has been driven by questions which arise

quite naturally in, e.g., financial transactions [1], data storage in computer science [2], queueing dynamics [3], interface fluctuations [4] and extreme statistics in time series analysis [5]. These have proven to be of interest to physicists and mathematicians alike; for an overview, see [6–8]. The issue we seek to address in this paper is motivated by general considerations rather than any one specific topic and finds its roots in the following classic problem. Given a Brownian motion, $x(\tau)$, in the interval $[0, t]$, subject to $x(0) = 0$ but otherwise unconstrained, at what time, t_m , does $x(\tau)$ reach its maximum value, M ? More precisely, what is the probability density, $P(t_m)$, associated with t_m ?

The answer to this is well known; $P(t_m) = \frac{1}{\pi} t_m^{-1/2} (t - t_m)^{-1/2}$, or equivalently the cumulative distribution is given by $\Pr(t_m \leq x) = \frac{2}{\pi} \sin^{-1}[\sqrt{x/t}]$. This is Lévy’s famous ‘arcsine law’ [9]. It is somewhat counterintuitive; the density has a minimum at the midpoint $t_m = t/2$ and peaks at the end points $t_m = 0$ and $t_m = t$, showing that Brownian motion is inherently ‘stiff’ [10, 11]. The corresponding result for a Brownian bridge (which has the additional constraint that $x(t) = 0$) is also known, namely $P(t_m) = 1/t$, which one may call the ‘uniform law’ [10]. Recently, an expression for $P(t_m)$ for a Brownian motion up to its first-passage time was also presented [12], adding to the results on first-passage Brownian functionals given in [13, 14]. The main focus of the present paper is to derive $P(t_m)$ for three other cases: (i) a Brownian excursion, (ii) a Brownian meander and (iii) a reflected Brownian bridge. We include the latter since it appears naturally in the context of certain probability laws related to the excursion and the meander (see section 4).

A Brownian excursion in the interval $[0, t]$ is defined as a Brownian motion, $x(\tau)$, constrained so that $x(0) = 0$, $x(t) = 0$ with $x(\tau) > 0$ for $0 < \tau < t$. A Brownian meander in the interval $[0, t]$ is the same except there is no constraint on the value of $x(t)$, other than it is positive. A reflected Brownian bridge in the interval $[0, t]$ is defined as the absolute value $|x(\tau)|$ of a Brownian motion constrained such that $x(0) = 0$ and $x(t) = 0$. For the excursion, on basic dimensional grounds one has that $P(t_m) = t^{-1} f(t_m/t)$, where the function $f(x)$ satisfies the normalization condition $\int_0^1 f(x) dx = 1$. Similarly, in relation to the meander and reflected bridge one can define corresponding functions $g(x)$ and $h(x)$ which are likewise normalized. It follows that the interval length t is only a trivial scaling factor and one can interpret the scaling functions $f(x)$, $g(x)$ and $h(x)$ as being the relevant probability densities for the respective motions in the interval $[0, 1]$. Our primary aim is to compute these functions explicitly. For convenience, we summarize the main findings here;

$$f(x) = 3 \sum_{m,n=1}^{\infty} (-1)^{m+n} \frac{m^2 n^2}{[n^2 x + m^2(1-x)]^{5/2}} \quad \text{Excursion} \quad (1)$$

$$g(x) = 2 \sum_{m=0,n=1}^{\infty} (-1)^{n+1} \frac{n^2}{[n^2 x + (2m+1)^2(1-x)]^{3/2}} \quad \text{Meander} \quad (2)$$

$$h(x) = 2 \sum_{m,n=0}^{\infty} (-1)^{m+n} \frac{(2m+1)(2n+1)}{[(2n+1)^2 x + (2m+1)^2(1-x)]^{3/2}} \quad \text{Reflected bridge.} \quad (3)$$

The primary method we employ to derive these results is based on a path-integral technique (in essence, the Feynman–Kac formula). In section 2 we describe how the approach leads naturally to expressions for the joint probability density $P(M, t_m)$. By subsequently integrating over M one can then obtain expressions for the marginal densities $P(t_m)$, and hence obtain the functions $f(x)$, $g(x)$ and $h(x)$ defined above. In section 3 we concentrate on analysing $f(x)$ and $g(x)$ in terms of their moments and asymptotic tails, and show that the results are

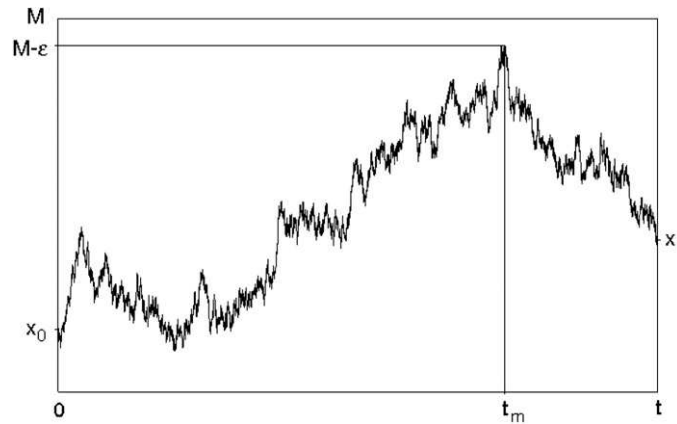


Figure 1. Schematic showing a typical Brownian path $x(\tau)$ constrained so that $x(0) = x_0$, $x(t) = x_t$ and $x(t_m) = M - \epsilon$, with $0 < x(\tau) < M$ almost surely for $0 \leq \tau \leq t$.

in excellent accord with numerical simulations. In section 4 we return to the results obtained for $P(M, t_m)$ to show that they may also be obtained by considering certain probabilistic laws known as ‘agreement formulae’. These laws are associated with random variables which have been defined and analysed by probability theorists in the study of Brownian motion processes. Finally, in section 5, overall conclusions are drawn.

2. Deriving the probability densities

The basic approach outlined in this section has been described in detail in [6, 7], so here we present a simplified overview. At a fundamental level, the processes being studied are represented by the Langevin equation $dx(\tau)/d\tau = \xi(\tau)$, where $\xi(\tau)$ is a Gaussian white noise source with correlator $\langle \xi(\tau)\xi(\tau') \rangle = \delta(\tau - \tau')$. For the purpose of simulation this means that the realization of sample paths can be achieved through a suitable limiting process of a discrete random walk (see the following section). From a theoretical perspective, a powerful tool in many instances is the use of Fokker–Planck equations, but it is not always easy to handle the given constraints on the process in question using this approach [6, 7]. The path-integral technique is a powerful alternative to the Fokker–Planck method. The central idea is that the probability measure $P[x(\tau)]$ associated with an unconstrained Brownian path $x(\tau)$ over the time interval $0 \leq \tau \leq t$ satisfies

$$P[x(\tau)] \propto \exp \left[-\frac{1}{2} \int_0^t \left(\frac{dx}{d\tau} \right)^2 d\tau \right]. \tag{4}$$

From this observation one can systematically construct solutions to the problems of interest by interpreting the ‘Lagrangian path-integral’ formalism of (4) in terms of an equivalent ‘Hamiltonian propagator’ formalism within which the constraints on the motion may be accommodated quite naturally. We proceed on a case-by-case basis.

2.1. Brownian excursion

Let us first consider the case of the Brownian excursion. With reference to figure 1, we are interested in those paths $x(\tau)$ which have $x(0) = 0$ and $x(t) = 0$ with $x(\tau) > 0$ for $0 < \tau < t$.

Note that a continuous-time Brownian motion, starting at $x(0) = 0$ at $\tau = 0$, will recross the origin an infinite number of times in the time interval $[0, \delta]$ for all $\delta > 0$ [10]. Hence it is impossible to maintain the constraint $x(\tau) > 0$ for $\tau > 0$ if we insist that $x(0) = 0$ from the outset. This problem can be circumvented by the following procedure [6, 7]. We assume that the process starts at $x(0) = x_0 > 0$, then impose the constraint $x(\tau) > 0$ for $\tau > 0$ without any problem, and only take the limit $x_0 \rightarrow 0$ at the appropriate stage in the calculation. Similarly, we assume the process ends at $x(t) = x_t > 0$, eventually taking the limit $x_t \rightarrow 0$. For computational convenience one can set $x_t = x_0$.

Next, one considers the time t_m at which the excursion reaches its (almost surely unique) maximum M . Again, we treat this as a limiting process by fixing the value of $x(t_m)$ to be $M - \varepsilon$ whilst imposing the constraint that the actual maximum is less than M , with the limit $\varepsilon \rightarrow 0$ only being taken at the appropriate stage. With these caveats, and with t_m and M assumed fixed, one can decompose a given path $x(\tau)$ into a left-hand segment, for which $0 \leq \tau \leq t_m$, and a right-hand segment, for which $t_m \leq \tau \leq t$, wherein for both $0 < x(\tau) < M$ almost surely (see figure 1). The statistical weight of, say, the left-hand segment is proportional to the propagator $\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle$, where the Hamiltonian $\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x)$. The potential $V(x)$ has infinite barriers at $x = 0$ and $x = M$; this ensures that the process is constrained to satisfy $0 \leq x(\tau) \leq M$ for $0 \leq \tau \leq t$. The normalized eigenfunctions of \hat{H} are simply $\psi_n(x) = \sqrt{\frac{2}{M}} \sin\left(\frac{n\pi x}{M}\right)$, whilst the corresponding eigenvalues are given by $E_n = n^2\pi^2/2M^2$. One can easily evaluate the propagator in this eigenbasis

$$\langle x_0 | e^{-\hat{H}t_m} | M - \varepsilon \rangle = \frac{2}{M} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x_0}{M}\right) \sin\left(\frac{n\pi(M - \varepsilon)}{M}\right) e^{-\frac{n^2\pi^2}{2M^2}t_m}. \quad (5)$$

Similarly, the statistical weight of the right-hand segment is proportional to the propagator $\langle M - \varepsilon | e^{-\hat{H}(t-t_m)} | x_0 \rangle$, which may be written down by analogy. With t_m and M fixed, the Markovian nature of the Brownian motion process means that the statistical weight of the right-hand segment is independent of the statistical weight of the left-hand segment. It follows that the joint probability density, $P(M, t_m)$, after taking the limits $\varepsilon \rightarrow 0$ and $x_0 \rightarrow 0$, satisfies

$$P(M, t_m) = A \frac{4\pi^4 \varepsilon^2 x_0^2}{M^6} \sum_{m,n=1}^{\infty} (-1)^{m+n} m^2 n^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{m^2\pi^2}{2M^2}(t-t_m)} + \dots \quad (6)$$

where \dots denotes the higher order terms in ε and x_0 . The amplitude (i.e. constant of proportionality) A , which is a function of ε and x_0 , may be determined by the normalization condition $\int_0^\infty \int_0^t P(M, t_m) dt_m dM = 1$. The integrals are straightforward to evaluate and to complete the calculation we make use of the following results;

$$\lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^{m+n} \frac{m+n}{mn} = \ln 2, \quad \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^{m+n} \frac{1}{m+n} = \ln 2 - \frac{1}{2}. \quad (7)$$

These results are simple to derive by appropriately differentiating or integrating the basic geometric series $\sum_{n=1}^{\infty} \alpha^n = \alpha/(1 - \alpha)$ for $|\alpha| < 1$. This idea of introducing α and letting $\alpha \rightarrow -1$ is not just a useful computational aid; it is an important regularization procedure which gives a precise meaning to certain sums which arise in the analysis. With the help of (7) one therefore finds that, for a Brownian excursion, the joint probability density is given by,

$$P(M, t_m) = \sqrt{2} \frac{\pi^{9/2} t^{3/2}}{M^6} \sum_{m,n=1}^{\infty} (-1)^{m+n} m^2 n^2 e^{-\frac{n^2\pi^2}{2M^2}t_m - \frac{m^2\pi^2}{2M^2}(t-t_m)}. \quad (8)$$

We are primarily interested in this paper in the marginal density $P(t_m)$. This may be obtained by integrating (8) over M , i.e. $P(t_m) = \int_0^\infty P(M, t_m) dM$. Before doing so,

however, we make a brief detour by considering the other marginal density associated with the maximum of a Brownian excursion, namely $P(M) = \int_0^t P(M, t_m) dt_m$. This is well known in the literature [15–17],

$$P(M) = \sqrt{2}\pi^{5/2}t^{3/2} \frac{d}{dM} \left\{ \frac{1}{M^3} \sum_{n=1}^{\infty} n^2 e^{-\frac{n^2\pi^2}{2M^2}t} \right\} \tag{9}$$

with moments given by,

$$\langle M^k \rangle = \frac{k(k-1)t^{k/2}}{2^{k/2}} \Gamma\left(\frac{k}{2}\right) \zeta(k) \tag{10}$$

where $\zeta(k)$ is the Riemann zeta function. By integrating (8) over t_m one should obtain an expression for $P(M)$ which is equivalent to (9). Interestingly, by doing so one obtains a representation which is quite different:

$$P(M) = \frac{2^{3/2}\pi^{5/2}t^{3/2}}{M^4} \sum_{m,n=1}^{\infty} (-1)^{m+n} \frac{m^2n^2}{m^2-n^2} \left[e^{-\frac{n^2\pi^2}{2M^2}t} - e^{-\frac{m^2\pi^2}{2M^2}t} \right]. \tag{11}$$

It is by no means obvious that (9) and (11) are equivalent, but in the appendix we shall prove that this is the case. It follows that the moments calculated using (11) must agree with (10). One quickly establishes that $\langle M \rangle = \sqrt{\pi t}/2$ and $\langle M^2 \rangle = \pi^2 t/6$, as required, and with a little more effort one can also verify that $\langle M^4 \rangle = \pi^4 t^2/30$. For the third moment, however, using (11) and comparing with (10) one obtains (when suitably regularized) an unusual and interesting identity,

$$\lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^{m+n} \frac{m^2n^2}{m+n} \left(\frac{\ln m - \ln n}{m-n} \right) = \frac{3\zeta(3)}{8\pi^2} \tag{12}$$

where $\zeta(3)$ is Apéry’s constant. We have checked this numerically to high precision and it is correct. We have been unable to find a shorter, more direct proof of (12).

Let us now return to considering the marginal density $P(t_m)$ for a Brownian excursion. By integrating (8) over M one obtains

$$P(t_m) = 3t^{3/2} \sum_{m,n=1}^{\infty} (-1)^{m+n} \frac{m^2n^2}{[n^2t_m + m^2(t-t_m)]^{5/2}} \equiv \frac{1}{t} f\left(\frac{t_m}{t}\right) \tag{13}$$

where the scaling function $f(x)$ is given by

$$f(x) = 3 \sum_{m,n=1}^{\infty} (-1)^{m+n} \frac{m^2n^2}{[n^2x + m^2(1-x)]^{5/2}}. \tag{14}$$

This is our first main result. Specifically, $f(x)$ is the probability density for the time to reach maximum for a Brownian excursion in the interval $[0, 1]$. We note that $f(x)$ is invariant under the interchange $x \rightarrow 1-x$; i.e. $f(x)$ is symmetric about $x = 1/2$. One may easily check that (14) is correctly normalized over this interval, i.e. $\int_0^1 f(x) dx = 1$. Thus,

$$\begin{aligned} \int_0^1 f(x) dx &= 3 \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^{m+n} m^2n^2 \int_0^1 \frac{dx}{[n^2x + m^2(1-x)]^{5/2}} \\ &= 2 \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^{m+n} \frac{m^2 + mn + n^2}{mn(m+n)} = 1 \end{aligned} \tag{15}$$

where we have used the results in (7) when making the final step. In the following section we will consider the low-order moments and asymptotics of $f(x)$, and also make comparison with numerical simulations.

2.2. *Brownian meander*

Turning now to the case of a Brownian meander, it is straightforward to adapt the above analysis for the excursion and we therefore only present the outline details. The key difference between the meander and the excursion is that there is no constraint on the final co-ordinate of the motion, x_t , other than $0 \leq x_t \leq M$ (see figure 1). Thus one must integrate over this co-ordinate using the result

$$\int_0^M \sin\left(\frac{m\pi x_t}{M}\right) dx_t = \frac{M}{m\pi} [1 - (-1)^m]. \tag{16}$$

Proceeding exactly as before one then derives in the limit $\varepsilon \rightarrow 0$ and $x_0 \rightarrow 0$

$$P(M, t_m) = B \frac{4\pi^2 \varepsilon^2 x_0}{M^4} \sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] n^2 e^{-\frac{n^2 \pi^2}{2M^2} t_m - \frac{m^2 \pi^2}{2M^2} (t-t_m)} + \dots \tag{17}$$

Again the unknown amplitude B may be determined by the normalization condition $\int_0^\infty \int_0^t P(M, t_m) dt_m dM = 1$. A useful result in this regard is

$$\lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^n \frac{n}{m(m+n)} = \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} \alpha^n \left[\frac{1}{m} - \frac{1}{m+n} \right] = -\frac{1}{2} \ln 2. \tag{18}$$

This may be proved by representing the sum in (18) as an integral by first using the identities $m^{-1} \equiv \int_0^\infty e^{-my} dy$ and $(m+n)^{-1} \equiv \int_0^\infty e^{-(m+n)y} dy$ and then interchanging the order of summation and integration. The final result for the joint probability density for a Brownian meander is given by

$$P(M, t_m) = \frac{\sqrt{2\pi} 5/2 t^{1/2}}{M^4} \sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] n^2 e^{-\frac{n^2 \pi^2}{2M^2} t_m - \frac{m^2 \pi^2}{2M^2} (t-t_m)}. \tag{19}$$

From this result, one can obtain in passing an expression for the marginal density $P(M)$ by integrating over t_m

$$P(M) = \frac{2^{3/2} \pi^{1/2} t^{1/2}}{M^2} \sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{m^2 - n^2} [e^{-\frac{n^2 \pi^2}{2M^2} t} - e^{-\frac{m^2 \pi^2}{2M^2} t}]. \tag{20}$$

There is a standard, well-known expression for $P(M)$, namely [18, 19],

$$P(M) = 2^{3/2} \sqrt{\pi} t^{1/2} \frac{d}{dM} \left\{ \frac{1}{M} \sum_{n=0}^{\infty} e^{-\frac{(2n+1)^2 \pi^2}{2M^2} t} \right\} \tag{21}$$

with moments [20]

$$\langle M^k \rangle = k 2^{k/2} (1 - 2^{1-k}) t^{k/2} \Gamma\left(\frac{k}{2}\right) \zeta(k). \tag{22}$$

As before, we prove in the appendix that the two representations (20) and (21) are equivalent, as indeed they must be. Using (20) one can determine that the second moment is given by $\langle M^2 \rangle = \pi^2 t / 3$, which agrees with (22). More interestingly, using (20) to determine the first moment and comparing with the result given by (22), namely $\langle M \rangle = \sqrt{2\pi t} \ln 2$, one obtains the following identity;

$$\lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} [\alpha^{m+n} - \alpha^n] \frac{n^2}{m+n} \left(\frac{\ln m - \ln n}{m-n} \right) = \frac{1}{2} \ln 2. \tag{23}$$

Once more, we have been unable to find a shorter, more direct proof of (23), but again we have checked it numerically to high precision and are satisfied that it is correct.

Returning to the main theme, to obtain the marginal density $P(t_m)$ for a Brownian meander we integrate (19) over M to give

$$P(t_m) = t^{1/2} \sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{[n^2 t_m + m^2(t - t_m)]^{3/2}} \equiv \frac{1}{t} g\left(\frac{t_m}{t}\right) \quad (24)$$

where

$$g(x) = \sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{[n^2 x + m^2(1 - x)]^{3/2}}. \quad (25)$$

One can write this in a slightly neater form by noting that only the terms where m is odd contribute, i.e. $(-1)^{m+n} - (-1)^n = 0$ when m is even. Thus

$$g(x) = 2 \sum_{m=0,n=1}^{\infty} (-1)^{n+1} \frac{n^2}{[n^2 x + (2m + 1)^2(1 - x)]^{3/2}}. \quad (26)$$

This is our second main result. Specifically, $g(x)$ is the probability density for the time to reach maximum for a Brownian meander in the interval $[0, 1]$. Again it is useful to check using (25) that $g(x)$ is correctly normalized;

$$\begin{aligned} \int_0^1 g(x) dx &= \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} [\alpha^{m+n} - \alpha^n] n^2 \int_0^1 \frac{dx}{[n^2 x + m^2(1 - x)]^{3/2}} \\ &= 2 \lim_{\alpha \rightarrow -1} \sum_{m,n=1}^{\infty} [\alpha^{m+n} - \alpha^n] \frac{n}{m(m+n)} = 1 \end{aligned} \quad (27)$$

where we have used the results in (7) and (18) in the final step. In the following section we will consider the low-order moments and asymptotics of $g(x)$, together with $f(x)$, with comparisons made against numerical simulations. It is perhaps worth stressing in advance that, unlike the function $f(x)$ which is *symmetric* about $x = 1/2$, the function $g(x)$ is manifestly *asymmetric*; indeed it diverges as $x \rightarrow 1$.

2.3. Reflected Brownian bridge

One can easily adapt the above path-integral method to calculate $P(M, t_m)$ for other Brownian motion processes. Anticipating the discussion in section 4, there are good mathematical reasons for studying the reflected Brownian bridge alongside the excursion and meander. The only significant modification to the calculation involves considering a Brownian motion $x(\tau)$ which is constrained to lie in the box $-M < x(\tau) < M$ for $0 \leq \tau \leq t$. We do not give the details of the derivation here, but simply present the results. One finds that,

$$P(M, t_m) = \sqrt{2} \frac{\pi^{5/2} t^{1/2}}{M^4} \sum_{m,n=0}^{\infty} (-1)^{m+n} \left(m + \frac{1}{2}\right) \left(n + \frac{1}{2}\right) e^{-\frac{(n+1/2)^2 \pi^2}{2M^2} t_m - \frac{(m+1/2)^2 \pi^2}{2M^2} (t - t_m)}. \quad (28)$$

By integrating (28) over t_m one therefore has that

$$P(M) = \frac{2^{3/2} \pi^{1/2} t^{1/2}}{M^2} \sum_{m,n=0}^{\infty} (-1)^{m+n} \frac{\left(m + \frac{1}{2}\right) \left(n + \frac{1}{2}\right)}{\left(m + \frac{1}{2}\right)^2 - \left(n + \frac{1}{2}\right)^2} \left[e^{-\frac{(n+1/2)^2 \pi^2}{2M^2} t} - e^{-\frac{(m+1/2)^2 \pi^2}{2M^2} t} \right] \quad (29)$$

which is equivalent (see the appendix) to the conventional expression [20]

$$P(M) = \sqrt{2\pi t} \frac{d}{dM} \left\{ \frac{1}{M} \sum_{n=0}^{\infty} e^{-\frac{(2n+1)^2 \pi^2}{8M^2} t} \right\}. \quad (30)$$

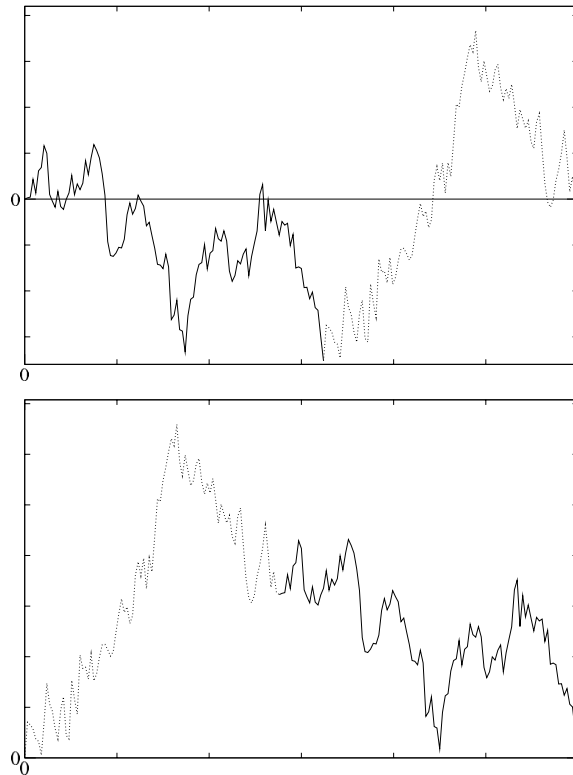


Figure 2. Schematic of the Vervaat construction: the upper diagram depicts a Brownian bridge; the lower diagram depicts the Brownian excursion constructed from that bridge.

More pertinently, by integrating (28) over M one obtains the marginal density $P(t_m) = t^{-1}h(t_m/t)$ where the scaling function $h(x)$ is given by,

$$h(x) = 2 \sum_{m,n=0}^{\infty} (-1)^{m+n} \frac{(2m+1)(2n+1)}{[(2n+1)^2x + (2m+1)^2(1-x)]^{3/2}}. \quad (31)$$

This is our third main result. It is straightforward to check that $h(x)$ is correctly normalized, i.e. $\int_0^1 h(x) dx = 1$, and to see that $h(x)$ is invariant under the interchange $x \rightarrow 1-x$; i.e. symmetric about $x = 1/2$.

3. Analysis and numerical simulations

In this section we concentrate on analysing two of the key results in this paper; the probability density $f(x)$ for the time to reach maximum for a Brownian excursion in the interval $[0, 1]$, given by (14), and the probability density $g(x)$ for the time to reach maximum for a Brownian meander in the interval $[0, 1]$, given by (26). We begin by considering the excursion.

To simulate a Brownian excursion in the interval $[0, 1]$, one starts with an ordinary Brownian motion, $x(\tau)$, represented as the appropriate limit of a discrete random walk. From this a Brownian bridge may be constructed using the transformation $x(\tau) \rightarrow x(\tau) - \tau x(1)$. Finally, from this Brownian bridge one can obtain a Brownian excursion using the Vervaat construction [21]. This is best illustrated graphically. In figure 2 the basic idea becomes clear;

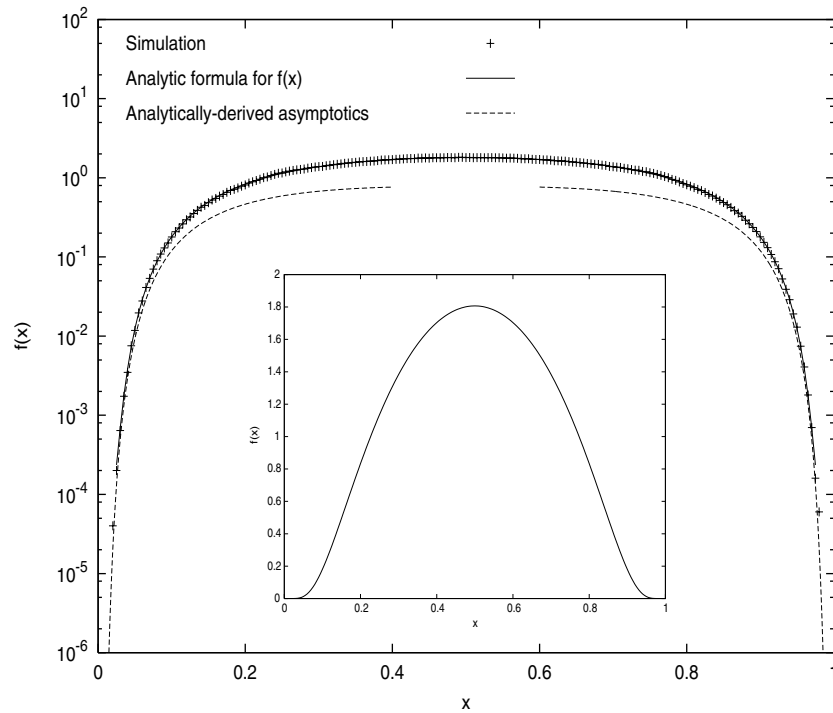


Figure 3. Comparison of simulated results with theoretical predictions for the probability density function $f(x)$. The inset shows the theoretical curve on a linear scale.

by cutting a given Brownian bridge at its minimum point, and rearranging and translating the two segments, one can construct a Brownian path with the desired characteristics, namely $x(0) = 0, x(1) = 0$ with $x(\tau) > 0$ for $0 < \tau < 1$. In figure 3 we plot the function $f(x)$ alongside the results of numerical simulations of the excursion process (based on 10^7 samples). The results are indistinguishable. We can determine the asymptotic behaviour of $f(x)$ as $x \rightarrow 0$ as follows. First we use the identity

$$\frac{1}{a^{5/2}} \equiv \frac{1}{\Gamma(\frac{5}{2})} \int_0^\infty y^{3/2} e^{-ay} dy \tag{32}$$

to write $f(x)$ as given by (14) in an equivalent form

$$f(x) = \frac{4}{\sqrt{\pi}} \int_0^\infty y^{3/2} \left\{ \sum_{n=1}^\infty (-1)^n n^2 e^{-n^2 xy} \right\} \left\{ \sum_{m=1}^\infty (-1)^m m^2 e^{-m^2(1-x)y} \right\} dy. \tag{33}$$

By such means one achieves a convenient factorization, although it comes at a price in that one has to carry out at some stage the parametric integration over the dummy variable y . Next, to study the limit $x \rightarrow 0$ we substitute $y = z/\sqrt{x}$ in (33) to give

$$f(x) = \frac{4}{\sqrt{\pi}} \frac{1}{x^{5/4}} \int_0^\infty z^{3/2} \left\{ \sum_{n=1}^\infty (-1)^n n^2 e^{-n^2 \sqrt{x}z} \right\} \left\{ \sum_{m=1}^\infty (-1)^m m^2 e^{-m^2(1-x)z/\sqrt{x}} \right\} dz. \tag{34}$$

If we consider the second summation in (34) first, in the limit $x \rightarrow 0$ this is dominated by the $m = 1$ term, thus,

$$\sum_{m=1}^{\infty} (-1)^m m^2 e^{-m^2(1-x)z/\sqrt{x}} \sim -e^{-z/\sqrt{x}}. \tag{35}$$

To deal with the first summation in (34) we use a key identity known from the theory of Jacobi theta functions [22]

$$1 + 2 \sum_{n=1}^{\infty} (-1)^n e^{-n^2 z} = 2 \left(\frac{\pi}{z}\right)^{1/2} \sum_{n=0}^{\infty} e^{-\pi^2(n+1/2)^2/z}. \tag{36}$$

Differentiating both sides of (36) with respect to z , replacing $z \rightarrow \sqrt{x}z$ and retaining only the dominant term on the right-hand-side gives the following asymptotic behaviour as $x \rightarrow 0$

$$\sum_{n=1}^{\infty} (-1)^n n^2 e^{-n^2 \sqrt{x}z} \sim -\frac{\pi^{5/2}}{4x^{5/4}z^{5/2}} e^{-\pi^2/(4\sqrt{x}z)}. \tag{37}$$

This means that by substituting (35) and (37) into (34) one obtains

$$f(x) \sim \frac{\pi^2}{x^{5/2}} \int_0^{\infty} \frac{1}{z} \exp\left[-\frac{1}{\sqrt{x}} \left(\frac{\pi^2}{4z} + z\right)\right] dz. \tag{38}$$

Using Laplace’s (saddle point) method to approximate this integral as $x \rightarrow 0$ one therefore finds the asymptotic behaviour

$$f(x) \sim 2^{1/2} \frac{\pi^2}{x^{9/4}} e^{-\pi/\sqrt{x}}. \tag{39}$$

This is also plotted in figure 3, and again the agreement with the numerical simulations is excellent. By symmetry, the asymptotic behaviour of $f(x)$ as $x \rightarrow 1$ is given by (39) with the replacement $x \rightarrow 1 - x$.

One can also consider the moments of $f(x)$. We have calculated the first three moments explicitly using (14) and obtain

$$\langle x \rangle = \frac{1}{2}; \quad \langle x^2 \rangle = \frac{15 - \pi^2}{18} = 0.28502\dots, \quad \langle x^3 \rangle = 1 - \frac{\pi^2}{12} = 0.17753\dots \tag{40}$$

The first moment follows on the grounds of symmetry, and the third moment may be deduced from the first two moments on the grounds of symmetry. From the simulation results we obtain the numerical estimates: $\langle x \rangle \approx 0.500\dots$, $\langle x^2 \rangle \approx 0.285\dots$ and $\langle x^3 \rangle \approx 0.177\dots$ which are fully consistent. One can calculate higher order moments in principle but this becomes an increasingly laborious task.

Turning now to the case of the meander, the simulation approach is slightly different. Again one starts with an ordinary Brownian motion, $x(\tau)$, represented as a discrete random walk, but this time only those sample paths for which $x(\tau) > 0$ for $0 < \tau < 1$ are retained; in the appropriate limit one has a good approximation to sample paths for the meander. In figure 4 we plot the function $g(x)$ alongside the results of numerical simulations of the meander process (based on 10^7 samples). Once again the agreement is excellent. This time, the asymptotic behaviour as $x \rightarrow 0$ is quite different from the behaviour as $x \rightarrow 1$, where $g(x)$ diverges. To investigate these limiting behaviours, we start with (26) and adapt the previous

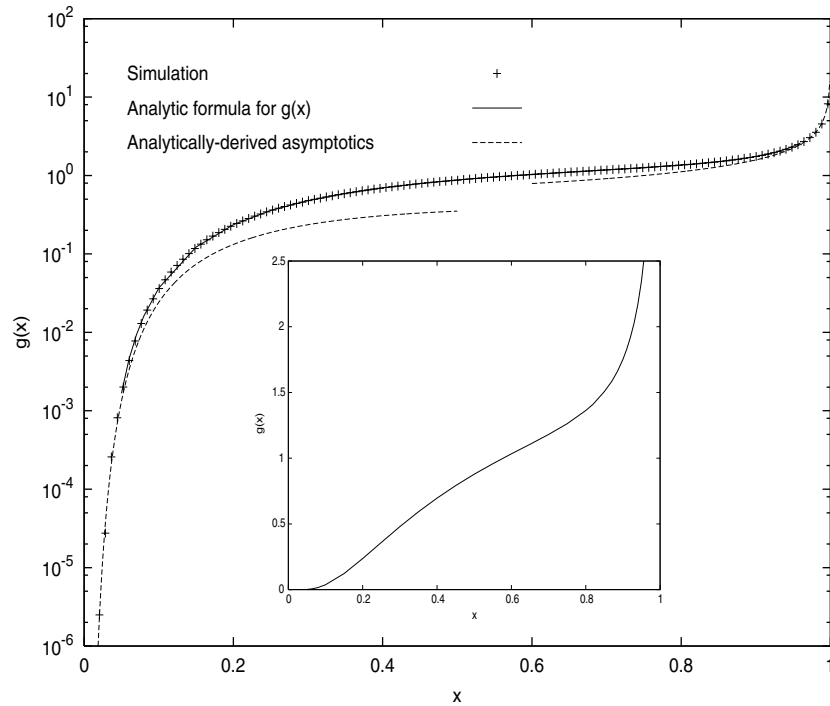


Figure 4. Comparison of simulated results with theoretical predictions for the probability density function $g(x)$. The inset shows the theoretical curve on a linear scale.

strategy to first rewrite $g(x)$ in the form

$$g(x) = \frac{4}{\sqrt{\pi}} \int_0^\infty y^{1/2} \left\{ \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 xy} \right\} \left\{ \sum_{m=0}^\infty e^{-(2m+1)^2(1-x)y} \right\} dy. \quad (41)$$

To study the limit $x \rightarrow 0$ we substitute $y = z/\sqrt{x}$ in (41) to give

$$g(x) = \frac{4}{\sqrt{\pi} x^{3/4}} \int_0^\infty z^{1/2} \left\{ \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 \sqrt{x} z} \right\} \left\{ \sum_{m=0}^\infty e^{-4(m+1/2)^2(1-x)z/\sqrt{x}} \right\} dz. \quad (42)$$

As $x \rightarrow 0$, to approximate the first summation in (42) we can use the result (37), whilst for the second summation in (42) we need retain only the $m = 0$ term so that

$$\sum_{m=0}^\infty e^{-4(m+1/2)^2(1-x)z/\sqrt{x}} \sim e^{-z/\sqrt{x}}. \quad (43)$$

This means that

$$g(x) \sim \frac{\pi^2}{x^2} \int_0^\infty \frac{1}{z^2} \exp \left[-\frac{1}{\sqrt{x}} \left(\frac{\pi^2}{4z} + z \right) \right] dz. \quad (44)$$

Once again by using Laplace's method to approximate the integral as $x \rightarrow 0$ we obtain the asymptotic behaviour of $g(x)$ for $x \rightarrow 0$

$$g(x) \sim 2^{3/2} \frac{\pi}{x^{7/4}} e^{-\pi/\sqrt{x}}. \quad (45)$$

This is plotted in figure 4 and agrees very well with the simulations. Turning now to the limit $x \rightarrow 1$, we cannot rely on symmetry arguments as was the case for the excursion, so we substitute $y = z/\sqrt{1-x}$ into (41) to give

$$g(x) = \frac{4}{\sqrt{\pi}(1-x)^{3/4}} \int_0^\infty z^{1/2} \left\{ \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 x z / \sqrt{1-x}} \right\} \left\{ \sum_{m=0}^\infty e^{-4(m+1/2)^2 \sqrt{1-x} z} \right\} dz. \tag{46}$$

Considering the second summation in (46) first, we can manipulate the fundamental theta function identity (36) to show that in the limit $x \rightarrow 1$

$$\sum_{m=0}^\infty e^{-4(m+1/2)^2 \sqrt{1-x} z} \sim \frac{\sqrt{\pi}}{4} \left(\frac{1}{1-x} \right)^{1/4} \frac{1}{z^{1/2}}. \tag{47}$$

For the first summation in (46) one has to be careful; it turns out that *all* the terms in it contribute to the integral to the same order as $x \rightarrow 1$. Thus one has that

$$g(x) \sim \frac{1}{(1-x)} \int_0^\infty \sum_{n=1}^\infty (-1)^{n+1} n^2 e^{-n^2 z / \sqrt{1-x}} dz = \frac{1}{\sqrt{1-x}} \sum_{n=1}^\infty (-1)^{n+1}. \tag{48}$$

The summation requires proper regularization for its correct interpretation, and gives

$$\sum_{n=1}^\infty (-1)^{n+1} = \lim_{\alpha \rightarrow -1} \sum_{n=1}^\infty \alpha^{n+1} = \lim_{\alpha \rightarrow -1} \frac{\alpha^2}{1-\alpha} = \frac{1}{2}. \tag{49}$$

Thus the asymptotic behaviour of $g(x)$ as $x \rightarrow 1$ turns out to be very simple

$$g(x) \sim \frac{1}{2\sqrt{1-x}}. \tag{50}$$

Again this accords well with the simulations, as shown in figure 4. The square root divergence is the same, except for the pre-factor, as that found for the case of unconstrained Brownian motion, as discussed in the introduction. One can understand why this might be at a qualitative level quite easily; for the unconstrained Brownian motion the dominant sample paths contributing to the $x \rightarrow 1$ behaviour are those which are largely positive throughout their journey and hence resemble meanders.

For completeness, we have also calculated the first two moments of $g(x)$ explicitly using (25), with the result that,

$$\langle x \rangle = \frac{\pi^2 - 4}{8} = 0.73370\dots, \quad \langle x^2 \rangle = \frac{19\pi^2 - 60}{216} = 0.59038\dots \tag{51}$$

From the simulation results we obtain the numerical estimates $\langle x \rangle \approx 0.733\dots$ and $\langle x^2 \rangle \approx 0.590\dots$, which again are fully consistent.

For the reflected Brownian bridge, one can follow similar procedures. In figure 5 we show a plot of the theoretical function given by (31). One notes that the function is almost constant except near $x = 0$ and $x = 1$, and it is worth comparing this with the density profile for the ordinary Brownian bridge which is constant everywhere (as discussed in the introduction). Through the use of additional Jacobi theta function identities one can show from (31) that as $x \rightarrow 0$,

$$h(x) \sim \frac{\pi}{x^{5/4}} e^{-\pi/2\sqrt{x}} \tag{52}$$

with the behaviour as $x \rightarrow 1$ given by (52) with the replacement $x \rightarrow 1-x$, due to symmetry. We omit the details, since the derivation of $h(x)$ presented in the previous section serves primarily to make a connection with what follows.

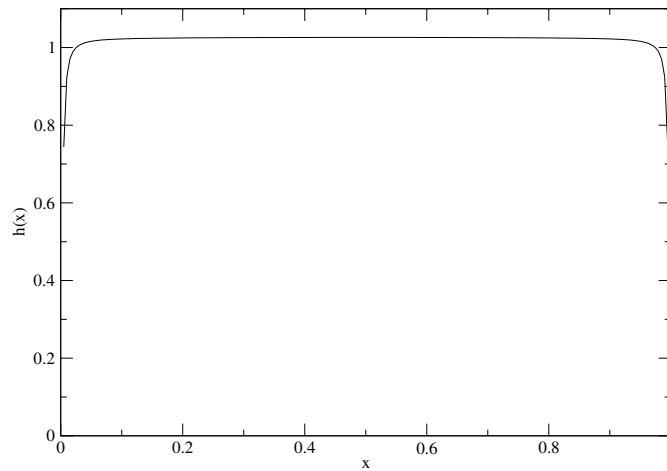


Figure 5. The theoretical curve for the probability density function $h(x)$ for the reflected Brownian bridge.

4. An alternative derivation via ‘agreement formulae’

The results derived above for the joint probability density $P(M, t_m)$ for the excursion (8), meander (19) and reflected bridge (28) have a particular mathematical structure which can be understood from a different perspective. In this section we show, in outline form only, that these results are manifestations of probabilistic laws associated with three random variables which have previously been studied in the Brownian motion literature; see [11] and, in particular, [23–25] and references therein. Such work builds upon the initial path decomposition work of Williams [26], which in turn provides another interpretation of the findings presented in [12] for the maximum of a Brownian motion up to its first-passage time, and upon the results presented in [27]. By such means one can provide an alternative, although less physically intuitive, method of derivation of the main results in this paper. First consider three variables S, T and C , which are characterized by the Laplace transform of their respective probability densities $P_S(u), P_T(u)$ and $P_C(u)$ [25, 28]

$$E[e^{-\lambda S}] = \frac{\sqrt{2\lambda}}{\sinh(\sqrt{2\lambda})} = 2 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{n^2 \pi^2}{n^2 \pi^2 + 2\lambda} \tag{53}$$

$$E[e^{-\lambda T}] = \frac{\tanh(\sqrt{2\lambda})}{\sqrt{2\lambda}} = 2 \sum_{n=0}^{\infty} \frac{1}{(n + \frac{1}{2})^2 \pi^2 + 2\lambda} \tag{54}$$

$$E[e^{-\lambda C}] = \frac{1}{\cosh(\sqrt{2\lambda})} = 2 \sum_{n=0}^{\infty} (-1)^n \frac{(n + \frac{1}{2})\pi}{(n + \frac{1}{2})^2 \pi^2 + 2\lambda} \tag{55}$$

where $E[e^{-\lambda S}] \equiv \int_0^\infty P_S(u) e^{-\lambda u} du$, etc and the series expansions on the right-hand-side of (53), (54) and (55) are well-known identities [29]. From these identities it follows that the probability densities can be written as

$$P_S(u) = \pi^2 \sum_{n=1}^{\infty} (-1)^{n+1} n^2 e^{-\frac{n^2 \pi^2}{2} u} \tag{56}$$

$$P_T(u) = \sum_{n=0}^{\infty} e^{-\frac{(n+1/2)^2 \pi^2}{2} u} \tag{57}$$

$$P_C(u) = \pi \sum_{n=0}^{\infty} (-1)^n \left(n + \frac{1}{2}\right) e^{-\frac{(n+1/2)^2 \pi^2}{2} u}. \tag{58}$$

With this preamble, the structure of (8), (19) and (28) can be deduced by considering certain so-called ‘agreement formulae’ [23–25]. These are identities in law between two-dimensional random variables which relate to various fundamental processes defined in the interval [0, 1] and which are valid for an *arbitrary* function f

$$E[f(M^2, t_m)] = \sqrt{\frac{\pi}{2}} E \left[f \left(\frac{1}{S + S'}, \frac{S}{S + S'} \right) \sqrt{S + S'} \right] \quad \text{Excursion} \tag{59}$$

$$E[f(M^2, t_m)] = \sqrt{\frac{\pi}{8}} E \left[f \left(\frac{1}{S + \frac{1}{4}T}, \frac{S}{S + \frac{1}{4}T} \right) \frac{1}{\sqrt{S + \frac{1}{4}T}} \right] \quad \text{Meander} \tag{60}$$

$$E[f(M^2, t_m)] = \sqrt{\frac{\pi}{2}} E \left[f \left(\frac{1}{C + C'}, \frac{C}{C + C'} \right) \frac{1}{\sqrt{C + C'}} \right] \quad \text{Reflected bridge.} \tag{61}$$

In these formulae, the variables S , T and C are the random variables described above with probability densities given by (56), (57) and (58) respectively, whilst M and t_m have the same meaning (in relation to the named process) as they have had throughout the paper. All three results, namely (59), (60) and (61), are of the same generic form

$$E[f(M^2, t_m)] = \alpha E \left[f \left(\frac{1}{U + V}, \frac{U}{U + V} \right) (U + V)^\mu \right] \tag{62}$$

where α and μ are chosen accordingly and, crucially, on the right-hand-side the random variables U and V are *independent*. The first (59) and third (61) of these results are particular cases of a Bessel bridge process of dimension $d = 2(1 + \mu)$ represented in terms of two independent Bessel processes considered up to their first hitting times of 1, and placed ‘back-to-back’ [24, 25]; see [11] for further references. The second result (60) is not precisely found in the literature, but may be obtained as a consequence of (61) (the details will be presented in another publication). For now, if we denote the probability density of U by $h(u)$ and the probability density of V by $k(v)$, one can use (62) to obtain a relationship between $P(M, t_m)$, the joint probability density of (M, t_m) , and the pair (h, k) . Again we skip the details, but in summary one can show that (62) implies that the function $P(*, *)$ must satisfy

$$\frac{1}{2\alpha} \frac{1}{(u + v)^{\mu+5/2}} P \left(\frac{1}{\sqrt{u + v}}, \frac{u}{u + v} \right) = h(u)k(v). \tag{63}$$

Letting $u = t_m/M^2$ and $v = (1 - t_m)/M^2$, and then exploiting the scaling properties of the processes to consider the general interval $[0, t]$ rather than just the interval $[0, 1]$, it follows that the joint probability densities $P(M, t_m)$ for the excursion, meander and reflected bridge are all of the form,

$$P(M, t_m) = 2\alpha \frac{t^{\mu+1}}{M^{2\mu+5}} h \left(\frac{t_m}{M^2} \right) k \left(\frac{t - t_m}{M^2} \right). \tag{64}$$

With reference to (59), (60) and (61), it therefore follows that

$$P(M, t_m) = \sqrt{2\pi} \frac{t^{3/2}}{M^6} P_S \left(\frac{t_m}{M^2} \right) \times P_S \left(\frac{t - t_m}{M^2} \right) \quad \text{Excursion} \tag{65}$$

$$P(M, t_m) = \sqrt{\frac{\pi}{2}} \frac{t^{1/2}}{M^4} P_S\left(\frac{t_m}{M^2}\right) \times 4P_T\left(\frac{4(t-t_m)}{M^2}\right) \quad \text{Meander} \quad (66)$$

$$P(M, t_m) = \sqrt{2\pi} \frac{t^{1/2}}{M^4} P_C\left(\frac{t_m}{M^2}\right) \times P_C\left(\frac{t-t_m}{M^2}\right) \quad \text{Reflected bridge.} \quad (67)$$

One may easily check using (56), (57) and (58) that (65), (66) and (67) reproduce in full the earlier results derived using the path-integral method. An expanded version of the discussion in this section will be given in a subsequent publication.

5. Conclusions

We have succeeded in deriving expressions for the probability density $P(t_m)$ for the time to reach maximum for a Brownian excursion (14), a Brownian meander (26) and a reflected Brownian bridge (31). This has been achieved first by using a path-integral technique, suitably adapted to each case in turn, with the key feature of introducing appropriate cut-offs which are then allowed to tend to zero. The derivation is reasonably transparent and, of course, can be adapted to give comparatively simple derivations of the ‘arcsine law’ and the ‘uniform law’ mentioned in the introduction. Indeed, this was one of the earliest applications of the Feynman–Kac formula. In passing, we have also derived in each case an expression for the probability density $P(M)$ associated with the distribution of the maximum. The representations for $P(M)$ thus obtained look quite different to the standard representations found in the literature but we have been able to prove their equivalence (see the appendix). By considering the moments, therefore, this leads to new, non-trivial identities (such as (12) and (23)) which we have verified numerically to high precision. For the excursion and meander, we have also analysed the moments and asymptotic limits of $P(t_m)$ in some detail and demonstrated that the theoretical results are in complete accord with numerical simulations. Finally, the ‘agreement formulae’ (59), (60) and (61) provide an alternative route to the derivation of the main results. At a fundamental level, this points to fascinating and deep connections with other problems and is a promising avenue for further study.

Appendix. Establishing the equivalence of certain probability densities

The expression for the probability density $P(M)$ for the maximum of a Brownian excursion obtained using the path-integral method, (9), looks quite different to the conventional expression, (11). Here we establish the equivalence. After simplifying both (9) and (11) this is tantamount to having to prove that

$$2 \sum_{m,n=1}^{\infty} (-1)^{m+n} \frac{m^2 n^2}{m^2 - n^2} [e^{-n^2 x} - e^{-m^2 x}] \stackrel{?}{=} -3 \sum_{n=1}^{\infty} n^2 e^{-n^2 x} + 2x \sum_{n=1}^{\infty} n^4 e^{-n^2 x}. \quad (A.1)$$

It is expedient (temporarily) to separate out the $m = n$ term on the left-hand-side (which is evaluated by a limiting procedure) since it exactly cancels the second term on the right-hand-side. By further integrating both sides with respect to x this means that (A.1) reduces to

$$2 \sum_{\substack{m,n=1 \\ m \neq n}}^{\infty} \frac{(-1)^{m+n}}{m^2 - n^2} [n^2 e^{-m^2 x} - m^2 e^{-n^2 x}] \stackrel{?}{=} 3 \sum_{n=1}^{\infty} e^{-n^2 x}. \quad (A.2)$$

Next we take the Laplace transform of both sides of (A.2) with respect to x to give

$$-2 \sum_{\substack{m,n=1 \\ m \neq n}}^{\infty} (-1)^{m+n} \left[\frac{(m^2 + n^2 + s)}{(m^2 + s)(n^2 + s)} \right] \stackrel{?}{=} 3 \sum_{n=1}^{\infty} \frac{1}{n^2 + s}. \quad (\text{A.3})$$

It is now helpful to ‘add back’ the $m = n$ term on the left-hand-side and to the right-hand-side also. After some simple algebra (A.3) then reduces to

$$2 \sum_{m,n=1}^{\infty} (-1)^{m+n} \left[\frac{(m^2 + n^2 + s)}{(m^2 + s)(n^2 + s)} \right] \stackrel{?}{=} \sum_{n=1}^{\infty} \frac{1}{n^2 + s} - 2s \sum_{n=1}^{\infty} \frac{1}{(n^2 + s)^2}. \quad (\text{A.4})$$

One can further simplify the left-hand-side by writing

$$\frac{(m^2 + n^2 + s)}{(m^2 + s)(n^2 + s)} \equiv \frac{1}{(n^2 + s)} + \frac{1}{(m^2 + s)} - \frac{s}{(m^2 + s)(n^2 + s)} \quad (\text{A.5})$$

whereupon the overall task condenses down to demonstrating that

$$-2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + s} - 2s \left(\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + s} \right)^2 \stackrel{?}{=} \sum_{n=1}^{\infty} \frac{1}{n^2 + s} - 2s \sum_{n=1}^{\infty} \frac{1}{(n^2 + s)^2}. \quad (\text{A.6})$$

The remaining steps needed to establish that (A.6) is a rigorous equality simply require one to use the identities [29]

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + s} = \frac{\pi}{2\sqrt{s}} \coth(\pi\sqrt{s}) - \frac{1}{2s} \quad (\text{A.7})$$

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + s} = \frac{\pi}{2\sqrt{s}} \operatorname{cosech}(\pi\sqrt{s}) - \frac{1}{2s} \quad (\text{A.8})$$

$$\sum_{n=1}^{\infty} \frac{1}{(n^2 + s)^2} = -\frac{1}{2s^2} + \frac{\pi}{4s^{3/2}} \coth(\pi\sqrt{s}) + \frac{\pi^2}{4s} \operatorname{cosech}^2(\pi\sqrt{s}) \quad (\text{A.9})$$

where (A.9) can be deduced from (A.7) by differentiating both sides with respect to s . It is now elementary to show that the left-hand-side and the right-hand-side of (A.6) are equal. It follows that since the Laplace transforms of (9) and (11) are equal, then the functions themselves are equal, and the proof is complete.

Similarly for the meander, the expression for the probability density of the maximum obtained using the path-integral method, (20), looks quite different to the conventional expression, (21). To establish the equivalence we proceed in the same manner as above. First, it is a simple matter to reduce the task to one of proving that

$$\sum_{\substack{m,n=1 \\ m \neq n}}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{m^2 - n^2} [e^{-n^2x} - e^{-m^2x}] \stackrel{?}{=} - \sum_{n=0}^{\infty} e^{-(2n+1)^2x}. \quad (\text{A.10})$$

Again we take Laplace transforms with respect to x so as to reduce (A.10) to

$$\sum_{\substack{m,n=1 \\ m \neq n}}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{(n^2 + s)(m^2 + s)} \stackrel{?}{=} - \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s}. \quad (\text{A.11})$$

By ‘adding back’ the $m = n$ term one can then reduce (A.11) to

$$\sum_{m,n=1}^{\infty} [(-1)^{m+n} - (-1)^n] \frac{n^2}{(n^2 + s)(m^2 + s)} \stackrel{?}{=} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s} - 2s \sum_{n=0}^{\infty} \frac{1}{((2n + 1)^2 + s)^2}. \tag{A.12}$$

Further manipulations of the left-hand-side of (A.12) yield additional simplifications until finally the task is to establish the following;

$$\left\{ \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + s} \right\} \left\{ \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s} \right\} \stackrel{?}{=} - \sum_{n=0}^{\infty} \frac{1}{((2n + 1)^2 + s)^2}. \tag{A.13}$$

This is easily done using the results (A.7) and (A.8). Thus by subtracting (A.8) from (A.7) one obtains

$$\sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s} = \frac{\pi}{4\sqrt{s}} [\coth(\pi\sqrt{s}) - \operatorname{cosech}(\pi\sqrt{s})] = \frac{\pi}{4\sqrt{s}} \tanh\left(\frac{\pi\sqrt{s}}{2}\right). \tag{A.14}$$

By further differentiating (A.14) with respect to s one obtains

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{1}{((2n + 1)^2 + s)^2} &= \frac{\pi}{8s^{3/2}} [\coth(\pi\sqrt{s}) - \operatorname{cosech}(\pi\sqrt{s})] \\ &\quad + \frac{\pi^2}{8s} \operatorname{cosech}^2(\pi\sqrt{s}) [1 - \cosh(\pi\sqrt{s})]. \end{aligned} \tag{A.15}$$

The proof then follows by direct substitution.

Finally, for the reflected bridge, to establish the equivalence between (29) and (30) one can follow the above procedure, which requires one to show that

$$2 \sum_{\substack{m,n=0 \\ m \neq n}}^{\infty} (-1)^{m+n} \frac{(2m + 1)(2n + 1)}{((2m + 1)^2 + s)((2n + 1)^2 + s)} \stackrel{?}{=} - \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s} \tag{A.16}$$

which is the same as showing that

$$2 \left[\sum_{n=0}^{\infty} (-1)^n \frac{(2n + 1)}{(2n + 1)^2 + s} \right]^2 \stackrel{?}{=} - \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2 + s} + 2 \sum_{n=0}^{\infty} \frac{(2n + 1)^2}{((2n + 1)^2 + s)^2}. \tag{A.17}$$

The right-hand-side of this expression may be rewritten using (A.14) and (A.15). The left-hand-side may then be shown to be equivalent using the result [29]

$$\sum_{n=0}^{\infty} (-1)^n \frac{(2n + 1)}{(2n + 1)^2 + s} = \frac{\pi}{4} \operatorname{sech}\left(\frac{\pi\sqrt{s}}{2}\right). \tag{A.18}$$

References

[1] Yor M 2001 *Exponential Functionals of Brownian Motion and Related Processes* (Berlin: Springer)
 [2] Flajolet P, Poblete P and Viola A 1998 *Algorithmica* **22** 490
 [3] Kearney M J 2004 *J. Phys. A: Math. Gen.* **37** 8421
 [4] Majumdar S N and Comtet A 2004 *Phys. Rev. Lett.* **92** 225501
 [5] Burkhardt T W, Gyorgyi G, Moloney N R and Racz Z 2007 *Phys. Rev. E* **76** 041119
 [6] Majumdar S N and Comtet A 2005 *J. Stat. Phys.* **119** 777
 [7] Majumdar S N 2005 *Curr. Sci.* **89** 2076 Preprint cond-mat/0510064
 [8] Janson S 2007 *Probab. Surv.* **4** 80

- [9] Lévy P 1939 *Comp. Math.* **7** 283
- [10] Feller W 1968 *An Introduction to Probability Theory and its Applications* (New York: Wiley)
- [11] Revuz D and Yor M 1999 *Continuous Martingales and Brownian Motion* 3rd edn (Berlin: Springer)
- [12] Randon-Furling J and Majumdar S N 2007 *J. Stat. Mech.* [P10008](#)
- [13] Kearney M J and Majumdar S N 2005 *J. Phys. A: Math. Gen.* **38** 4097
- [14] Kearney M J, Majumdar S N and Martin R J 2007 *J. Phys. A: Math. Theor.* **40** F863
- [15] Chung K L 1975 *Bull. Am. Math. Soc.* **81** 742
- [16] Chung K L 1976 *Ark. Mat. Math. Soc.* **14** 155
- [17] Kennedy D P 1976 *J. Appl. Probab.* **13** 371
- [18] Durrett R T, Iglehart D L and Miller D R 1977 *Ann. Probab.* **5** 117
- [19] Durrett R T and Iglehart D L 1977 *Ann. Probab.* **5** 130
- [20] Pitman J and Yor M 1999 *Electron. J. Probab.* **4** 15
- [21] Vervaat W 1979 *Ann. Probab.* **7** 143
- [22] Abramowitz M and Stegun I A 1973 *Handbook of Mathematical Functions* (New York: Dover)
- [23] Biane P and Yor M 1987 *Bull. Sci. Math.* **111** 23
- [24] Pitman J and Yor M 1996 *Ito's Stochastic Calculus and Probability Theory* (New York: Springer) pp 293–310
- [25] Biane P, Pitman J and Yor M 2001 *Bull. Am. Math. Soc.* **38** 435
- [26] Williams D 1974 *Proc. Lond. Math. Soc.* **28** 738
- [27] Csaki E, Foldes A and Salminen P 1987 *Ann. Inst. Henri Poincaré* **23** 179
- [28] Pitman J and Yor M 2003 *Can. J. Math.* **55** 292
- [29] Prudnikov A P, Marichev O I and Brychkov Yu A 1989 *Integrals and Series* (New York: Gordon and Breach)

Exact Distribution of the Maximal Height of p Vicious Walkers

Grégory Schehr,¹ Satya N. Majumdar,² Alain Comtet,² and Julien Randon-Furling²

¹Laboratoire de Physique Théorique (UMR du CNRS 8627), Université de Paris-Sud, 91405 Orsay Cedex, France

²Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, Bâtiment 100, 91405 Orsay Cedex, France

(Received 3 July 2008; published 8 October 2008)

Using path-integral techniques, we compute exactly the distribution of the maximal height H_p of p nonintersecting Brownian walkers over a unit time interval in one dimension, both for excursions p watermelons with a wall, and bridges p watermelons without a wall, for all integer $p \geq 1$. For large p , we show that $\langle H_p \rangle \sim \sqrt{2p}$ (excursions) whereas $\langle H_p \rangle \sim \sqrt{p}$ (bridges). Our exact results prove that previous numerical experiments only measured the preasymptotic behaviors and not the correct asymptotic ones. In addition, our method establishes a physical connection between vicious walkers and random matrix theory.

DOI: 10.1103/PhysRevLett.101.150601

PACS numbers: 05.40.-a, 02.50.-r, 05.70.Np

Introduction.—Since the pioneering work of de Gennes [1], followed up by Fisher [2], the subject of vicious (non-intersecting) random walkers has attracted a lot of interest among physicists. It has been studied in the context of wetting and melting [2], networks of polymers [3] and fibrous structures [1], persistence properties in nonequilibrium systems [4] and stochastic growth models [5,6]. There also exist connections between the vicious walker problem and the random matrix theory (RMT) [7–9], including for instance Dyson’s Brownian motion [10]. These connections to RMT have rekindled recent interest in the vicious walker problem and have led to new interesting questions. However, despite extensive recent mathematical literature on the subject, the connections to RMT have so far been established using mostly combinatorial approaches. Given the nonintersection constraint in the vicious walker problem, it is natural to expect a free Fermion approach to make its connection to RMT physically more explicit. The aim of this Letter is to present such an approach which, in addition, allows us to derive a variety of new exact results in the vicious walker problem.

Physically, one-dimensional vicious walkers play an important role in describing the elementary topological excitations in the $p \times 1$ commensurate adsorbed phases close to the commensurate-incommensurate (C-IC) transition [11]. In the commensurate phase the elementary excitations are pairs of dislocations at a given distance with p nonintersecting domain walls emerging from one and terminating at the other. This is just a “watermelon” pattern configuration of p nonintersecting Brownian bridges [see Fig. 1(b)]. The sizes of such defects and their fluctuations become important near the phase transition. An important quantity that characterises the transverse fluctuations of the defect is the maximal height of the p vicious walkers in a fixed time (here time signifies the fixed longitudinal distance between the pair). Such extreme value questions have recently been studied extensively for a *single* Brownian bridge or an excursion (with certain constraints) in the

context of the maximal height of a fluctuating interface [12,13]. In this Letter, we obtain exactly the distribution of the maximal height for p nonintersecting Brownian bridges and excursions.

Motivated by the geometry of elementary excitations discussed above, we thus focus on “watermelons” pattern configurations [see Figs. 1(a) and 1(b)] where p nonintersecting Brownian walkers $x_1(\tau) < \dots < x_p(\tau)$, starting at 0 at time $\tau = 0$, arrive at the same position at $\tau = 1$. We consider both “ p watermelons with a wall” [Fig. 1(a)], where the walkers stay positive in the time interval $[0, 1]$ and “ p -watermelons without wall” [Fig. 1(b)] where the walkers are free to cross the origin in between. Our main focus is on H_p , the maximal height of the top walker in $[0, 1]$, $H_p = \text{Max}_\tau[x_p(\tau), 0 \leq \tau \leq 1]$.

In particular, we are interested in the cumulative distribution $F_p(M) = \text{Prob}[H_p \leq M]$ and in the moments $\langle H_p^s \rangle$. For $p = 1$, there exist well known results [14], e.g., $\langle H_1 \rangle = \sqrt{\pi}/2$ for an excursion, or $\langle H_1 \rangle = \sqrt{\pi}/8$ for a bridge. Recently, Bonichon and Mosbah (BM) [15], using an algorithm based on exact enumerative formulas [16], conjectured, from numerical simulations, that for $p > 1$, $\langle H_p \rangle_{\text{num}} \approx \sqrt{1.67p - 0.06}$ for watermelons with a wall

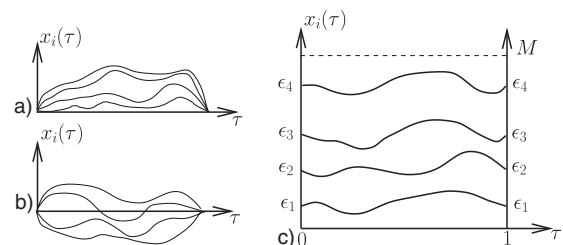


FIG. 1. (a) Four watermelons with a wall. (b) Four watermelons without a wall. (c) Illustration of the method to compute $F_4(M)$ using path-integral techniques where appropriate cutoffs ϵ_i 's have been introduced.

and $\langle H_p \rangle_{\text{num}} \simeq \sqrt{0.82p - 0.46}$ for watermelons without wall. These results stimulated several recent works [17–21] aiming at an analytical derivation of these estimates.

On the other hand, exploiting the recent connection between watermelons and the Airy processes [9,22], setting $\tilde{x}_p(\tau) = x_p(\tau)/\sqrt{\tau(1-\tau)}$, one expects that, in the limit $p \rightarrow \infty$, $\tilde{x}_p(\tau) = A\sqrt{p} + p^{-1/6}\xi$ where $A = 2^{3/2}$ (excursions) and $A = 2$ (bridges), where ξ is the Airy₂ process [9,22] of a suitably rescaled time parameter. Thus in the large p limit, the top curve approaches a limit shape, $x_p(\tau) \rightarrow A\sqrt{p}\sqrt{\tau(1-\tau)}$. Since the maximum of the top curve occurs at the midpoint $\tau = 1/2$, one expects that for $p \gg 1$, $\langle H_p \rangle \sim \langle x_p(\tau = \frac{1}{2}) \rangle \sim \sqrt{2p}$ for excursions and, similarly, $\langle H_p \rangle \sim \sqrt{p}$ for bridges. These exact asymptotic estimates differ considerably from the numerical estimates of BM suggesting that the latter only describe the pre-asymptotic behavior of $\langle H_p \rangle$. However, it calls for an explanation why this preasymptotic behavior as measured by BM should be about $\sqrt{1.67p}$ and $\sqrt{0.82p}$.

In this Letter, we present a method based on path integrals associated to corresponding free Fermions models to compute exactly $F_p(M)$. Our exact formula is useful for a number of reasons. It provides the exact asymptotic tails of the distribution of H_p which were not known before. For the average height, our formula explains the aforementioned discrepancy between the estimates of BM and the exact asymptotic behaviors of $\langle H_p \rangle$. We show that for moderate values of p (preasymptotic behavior), one obtains $\langle H_p \rangle \propto \pi\sqrt{p/6} = \sqrt{1.64493 \cdots p}$ for excursions and $\langle H_p \rangle \propto \pi\sqrt{p/12} = \sqrt{0.822467 \cdots p}$ for bridges, in nice agreement with BM's estimates. Finally, we show how our method allows for a physical derivation of the connection between p -watermelons' configurations and RMT.

Method.—To calculate the cumulative distribution $F_p(M)$, we use a path-integral method which needs to be suitably adapted to this problem. Indeed one notices that the p -watermelons' configurations described above [see, e.g., Figs. 1(a) and 1(b)] are ill defined for systems in continuous space and time. For such Brownian walks, it is well known that if two walkers cross each other once, they will recross each other infinitely many times immediately after the first crossing. Therefore, it is impossible to enforce the constraint $x_i(0) = x_{i+1}(0) = 0$ and simultaneously forcing $x_i(\tau) < x_{i+1}(\tau)$ immediately after. The cleanest way to circumvent this problem is to consider discrete time random walks moving on a discrete one-dimensional lattice (so called Dyck path): this is the method used in Refs. [16,17,19,20]. By taking the diffusion continuum limit, one would then arrive at nonintersecting Brownian motions [23]. This method is however mathematically cumbersome. Alternatively, following Refs. [12,24], we can go around this problem by assuming that the starting and finishing positions of the p walkers are $0 < \epsilon_1 < \dots < \epsilon_p$ [see Fig. 1(c)]. Only at the end we take

the limit $\epsilon_i \rightarrow 0$ and show that it is well defined. In addition, in order to compute $F_p(M)$, we put an absorbing hard wall at M such that

$$F_p(M) = \lim_{\epsilon_i \rightarrow 0} \left[\frac{N(\boldsymbol{\epsilon}, M)}{N(\boldsymbol{\epsilon}, M \rightarrow \infty)} \right], \quad (1)$$

where $\boldsymbol{\epsilon} \equiv \epsilon_1, \dots, \epsilon_p$ and $N(\boldsymbol{\epsilon}, M)$ is the probability that the p Brownian paths starting at $0 < \epsilon_1 < \dots < \epsilon_p$ at $\tau = 0$ come back to the same points at $\tau = 1$ without crossing each other and staying within the interval $[\mu, M]$, with $\mu = 0$ for excursions and $\mu \rightarrow -\infty$ for bridges. This procedure is depicted in Fig. 1(c).

The probability measure associated to p unconstrained Brownian paths $x_1(\tau), \dots, x_p(\tau)$ over the time interval $[0, 1]$ is proportional to $\exp[-\frac{1}{2} \sum_{i=1}^p \int_0^1 (\frac{dx_i}{d\tau})^2 d\tau]$. Here, we have to incorporate the constraint that they stay in the interval $[\mu, M]$. Therefore one can use path-integral techniques to write $N(\boldsymbol{\epsilon}, M)$ in Eq. (1) as the propagator

$$N(\boldsymbol{\epsilon}, M) = \langle \boldsymbol{\epsilon} | e^{-\hat{H}_M} | \boldsymbol{\epsilon} \rangle, \quad (2)$$

with $\hat{H}_M = \sum_{i=1}^p [-\frac{1}{2} \frac{\partial^2}{\partial x_i^2} + V(x_i)]$, where $V(x)$ is a confining potential with $V(x) = 0$ if $x \in [\mu, M]$ and $V(x) = \infty$ outside this interval. Denoting by E the eigenvalues of \hat{H}_M and $|E\rangle$ the corresponding eigenvectors one has

$$N(\boldsymbol{\epsilon}, M) = \sum_E |\Psi_E(\boldsymbol{\epsilon})|^2 e^{-E}, \quad (3)$$

where we introduced the notation $\langle \mathbf{x} | E \rangle = \Psi_E(\mathbf{x})$. Importantly, to take into account the fact that we are considering here nonintersecting Brownian paths, the many-body wave function $\Psi_E(\mathbf{x}) \equiv \Psi_E(x_1, \dots, x_p)$ must be Fermionic, i.e., it vanishes if any of the two coordinates are equal. This many-body antisymmetric wave function is thus constructed from the one-body eigenfunctions of \hat{H}_M by forming the associated Slater determinant.

Watermelons with a wall.—In that case $\mu = 0$ and the one-body eigenfunctions are given by $\phi_n(x) = \sqrt{\frac{2}{M}} \sin \frac{n\pi x}{M}$ with discrete eigenvalues $\frac{n^2 \pi^2}{2M^2}$, $n \in \mathbb{N}^*$. Therefore one has

$$\Psi_E(\boldsymbol{\epsilon}) = \frac{1}{\sqrt{p!}} \det_{1 \leq i, j \leq p} \phi_{n_i}(\epsilon_j), \quad E = \frac{\pi^2}{2M^2} \mathbf{n}^2, \quad (4)$$

where we use the notation $\mathbf{n}^2 = \sum_{i=1}^p n_i^2$, $n_i \in \mathbb{N}^*$. From this expression (4), one checks that, in the limit $\epsilon_1 \rightarrow 0, \dots, \epsilon_p \rightarrow 0$, powers of ϵ_i 's cancel between the numerator and the denominator in Eq. (1), yielding

$$F_p(M) = \frac{A_p}{M^{2p^2+p}} \sum_{n_1, \dots, n_p} [\Xi(\mathbf{n})]^2 e^{-(\pi^2/2M^2)\mathbf{n}^2}, \quad (5)$$

$$\Xi(\mathbf{n}) = \prod_{1 \leq j < k \leq p} (n_j^2 - n_k^2) \prod_{i=1}^p n_i,$$

where A_p , a constant independent of M , is determined by requiring that $\lim_{M \rightarrow \infty} F(M) = 1$. It can be evaluated

using a Selberg's integral [25] yielding $A_p = \pi^{2p^2+p}/[2^{p^2-p/2} \prod_{j=0}^{p-1} \Gamma(2+j)\Gamma(\frac{3}{2}+j)]$. For $p = 1$, our expression gives back the well known result for a Brownian excursion [26]. For $p = 2$, we have checked, using the Poisson summation formula that our expressions in Eq. (5) yield back the result of Ref. [18]. For generic p , the probability distribution function (PDF) $F'_p(M)$ is bell shaped, exhibiting a single mode. At variance with previous studies [18,19], our expression (5) is easily amenable to an asymptotic analysis for small M . Indeed, when $M \rightarrow 0$, the leading contribution to the sum in (5) comes from $n_i = i$ and its $p!$ permutations, yielding for $M \rightarrow 0$

$$F_p(M) \sim \frac{\alpha_p}{M^{2p^2+p}} e^{-(\pi^2/12M^2)p(p+1)(2p+1)}, \quad (6)$$

where α_p can be explicitly computed, yielding for instance $\alpha_2 = 12\pi^9$. For large M , one can use the Poisson summation formula to obtain $1 - F_p(M) \propto \exp(-2M^2)$.

From the distribution in Eq. (5), one can compute the moments of the distribution $\langle H_p^s \rangle$. For $p \geq 2$, one obtains that $\langle H_p \rangle$ can be expressed in terms of integrals involving the Jacobi theta function $\vartheta(u) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 u}$ and its derivatives, thus recovering, by a simpler physical derivation, the results of Refs. [17–20]. In particular, one has $\langle H_2 \rangle = 1.82262\dots$ [17]. For moderate values of p , one observes that the main contribution to the average $\langle H_p \rangle = \int_0^\infty M F'_p(M) dM$ comes from relatively small M where $F'_p(M)$ is dominated, as before in Eq. (6), by the terms where $n_i = i$ and its $p!$ permutations. It is easy to see that the PDF, restricted to this first term (6) exhibits a maximum for $M^* \sim \pi\sqrt{p/6}$. Therefore, one expects that $\langle H_p \rangle \sim M^* = \sqrt{1.64493 \dots p}$, in good agreement with the estimates of BM [15]. For larger values of p the average $\langle H_p \rangle$ picks up contributions from larger values of M where $F'_p(M)$ can not be approximated by a single term as in Eq. (6) and therefore the estimate of BM ceases to be correct. Instead, one has the exact asymptotic behavior $\langle H_p \rangle \sim \sqrt{2p}$ for $p \gg 1$, which can be obtained directly from our formula in Eq. (5) [27].

Watermelons without walls.—In the case of Brownian bridges, one can apply the same formalism as above (1)–(3) with $\mu \rightarrow -\infty$, i.e. $\hat{H}_M = \sum_{i=1}^p \frac{-1}{2} \frac{\partial^2}{\partial x_i^2}$. In that case, the one-body eigenfunctions are given by $\psi_k(x) = \sqrt{\frac{2}{\pi}} \sin[k(M-x)]$ with a continuous spectrum $E_k = k^2/2$, $k \in \mathbb{R}^+$. Therefore, $\Psi_E(\epsilon)$ entering the expression of $N(\epsilon, M)$ in Eq. (3) is formally given by Eq. (4) where ϕ_{n_i} is replaced by ψ_{k_i} and $E = \frac{k^2}{2}$. One obtains

$$F_p(M) = \frac{B_p}{M^{p^2}} \int_0^\infty dy_1 \dots \int_0^\infty dy_p e^{-(y^2/2M^2)} \Theta_p(\mathbf{y})^2, \quad (7)$$

$$\Theta_p(\mathbf{y}) = \det_{1 \leq i, j \leq p} y_i^{j-1} \cos\left(y_i + j \frac{\pi}{2}\right),$$

where $B_p = 2^{2p}/[(2\pi)^{p/2} \prod_{j=1}^p \Gamma(j+1)]$. This yields, for

instance, $F_2(M) = 1 - 4M^2 e^{-2M^2} - e^{-4M^2}$. From Eq. (7), one obtains the asymptotic behavior for $M \rightarrow 0$ as

$$F_p(M) \propto M^{p^2+p}, \quad (8)$$

whereas for large M one has $1 - F_p(M) \propto \exp(-2M^2)$. As in the case of watermelons with a wall, the PDF $F'_p(M)$ is also bell shaped with a single mode. Notice, however, that the presence of the wall has drastic effects on the small M behavior of $F_p(M)$ [see Eq. (6) and (8)] whereas, as expected, it has less influence for large M .

From $F_p(M)$ in (7), one computes the moments $\langle H_p^s \rangle$, yielding $\langle H_2 \rangle = \frac{1+\sqrt{2}}{4} \sqrt{\pi}$ or $\langle H_3 \rangle = \frac{45+36\sqrt{2}-8\sqrt{6}}{96} \sqrt{\pi}$, recovering (to leading order) recent results obtained by rather involved combinatorial techniques [20].

To make contact with BM's estimates, one first focuses on $p = 2$ and notices that $\Theta_2(y_1, y_2)$ in Eq. (7) exhibits saddles for $y_1 = \pm\pi/2$, $y_2 = \pm\pi$ and for symmetric points obtained by permutations: this is shown in Fig. 2 (a). In fact this property can be generalized to higher values of p and one can show that $\Theta_p(\mathbf{y})$ has saddles which are located around $y_1 = \pm\pi/2$, $y_2 = \pm\pi, \dots, y_p = \pm p\pi/2$ and the points obtained by permutations. Of course $\Theta_p(\mathbf{y})$ develops saddles for higher values of y^2 but their weights are exponentially suppressed in Eq. (7). For moderate values of p , one expects that $\langle H_p \rangle$ is dominated by these saddles $y_1 = \pm\pi/2$, $y_2 = \pm\pi, \dots, y_p = \pm p\pi/2$. Therefore performing a saddle point calculation, one has $F_p(M) \propto e^{-p^2 \chi(M/\sqrt{p})}$, with $\chi(y) = \log y + \pi^2/(24y^2)$, which has a minimum for $y^* = \pi/\sqrt{12}$. This yields $\langle H_p \rangle \sim \pi\sqrt{p/12} = \sqrt{0.822467 \dots p}$, in good agreement with the estimates of BM [15]. For larger values of p one expects that $\langle H_p \rangle$ picks up contributions from larger values of M where $F_p(M)$ can not be reduced to these first saddles. In Fig. 2(b), one shows a comparison between the exact value of $\langle H_p \rangle^2$ computed from Eq. (7) and the estimate of BM. This clearly shows that the estimate of BM corre-

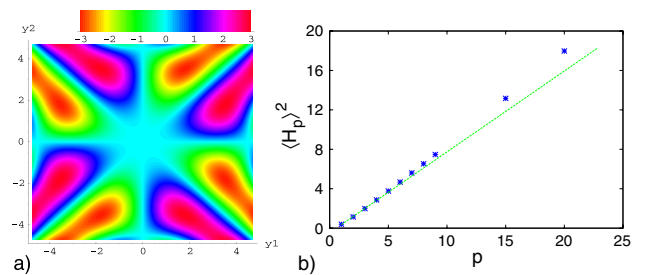


FIG. 2 (color online). (a) Contour plot of $\Theta_2(y_1, y_2) = y_2 \sin y_1 \cos y_2 - y_1 \cos y_1 \sin y_2$ given in (7). It exhibits saddles for $(y_1, y_2) = (\pm\pi/2, \pm\pi)$ and symmetric points obtained by permutations. (b) Plot of $\langle H_p \rangle^2$ as a function of p . The dotted line is the estimate from BM [15]. The quality of this estimate for $p \leq 10$ has its origin in the saddles of $\Theta_p(\mathbf{y})$ shown, for $p = 2$, on the left panel. For larger values of p one has instead $\langle H_p \rangle^2 \propto p$.

spond to the preasymptotic behavior. Instead, for large p , one expects here $\langle H_p \rangle \propto \sqrt{p}$.

Extension of the method.—The method presented here can be used to derive many other results. As an interesting example, showing explicitly the connection between watermelons and RMT, we compute the joint probability distribution $P_{\text{joint}}(x_1, \dots, x_p, \tau)$, first for p bridges. Following the same steps as above, Eq. (1)–(3), and using the Markov property of Brownian paths, one has

$$P_{\text{joint}}(\mathbf{x}, \tau) = \lim_{\epsilon_i \rightarrow 0} \frac{\langle \epsilon | e^{-\tau \hat{H}_0} | \mathbf{x} \rangle \langle \mathbf{x} | e^{-(1-\tau)\hat{H}_0} | \epsilon \rangle}{\langle \epsilon | e^{-\hat{H}_0} | \epsilon \rangle} \quad (9)$$

with $\hat{H}_0 = \sum_{i=1}^p \frac{-1}{2} \frac{\partial^2}{\partial x_i^2}$. One can show that powers of ϵ_i 's cancel between the numerator and the denominator in (9), yielding $P_{\text{joint}}(\mathbf{x}, \tau) \propto Q(\mathbf{x}, \tau)Q(\mathbf{x}, 1-\tau)$ with

$$Q(\mathbf{x}, \tau) = \int d\mathbf{k} \prod_{i < j} (k_i - k_j) e^{-(\tau \mathbf{k}^2/2)} \det_{1 \leq m, n \leq p} e^{i(x_m k_n)}, \quad (10)$$

where $\int d\mathbf{k} \equiv \int_{-\infty}^{\infty} dk_1 \cdots \int_{-\infty}^{\infty} dk_p$. After some algebra to evaluate the integrals in Eq. (10) one finally obtains, for p watermelons without wall

$$P_{\text{joint}}(\mathbf{x}, \tau) = Z_p^{-1} \sigma(\tau)^{-p^2} \prod_{i < j} (x_i - x_j)^2 e^{-(\mathbf{x}^2/2\sigma^2(\tau))}, \quad (11)$$

with $\sigma(\tau) = \sqrt{\tau(1-\tau)}$ and Z_p a normalization constant. This expression in Eq. (11) shows that this joint probability is exactly the one of the eigenvalues of the Gaussian unitary ensemble of random matrices (GUE) [7,8,10]. In particular, for $p \gg 1$, defining the rescaled variable $\eta = \sqrt{2} p^{1/6} (\frac{x_p(\tau)}{\sqrt{2\sigma(\tau)}} - \sqrt{2p})$, one obtains that the cumulative distribution of η is given by $\text{Proba}[\eta \leq x] = \mathcal{F}_2(x)$, the Tracy-Widom distribution for $\beta = 2$ [28].

For excursions, a similar calculation shows that

$$P_{\text{joint}}(\mathbf{x}, \tau) = Z_p'^{-1} \sigma(\tau)^{-p(2p+1)} [\Xi(\mathbf{x})]^2 e^{-(\mathbf{x}^2/2\sigma^2(\tau))}, \quad (12)$$

where $\Xi(\mathbf{x})$ is defined in (5) and Z_p' a normalization constant. Hence the joint distribution of $y_i = x_i^2/2\sigma^2(\tau)$ is formally identical to the distribution of the eigenvalues of Wishart matrices [25] with $M - N = \frac{1}{2}$, and $N = p$. In that case, from the results for the largest eigenvalue of Wishart matrices we conclude that for $p \gg 1$, the cumulative distribution of the rescaled variable $\zeta = 2^{2/3} p^{1/6} (\frac{x_p(\tau)}{\sqrt{2\sigma(\tau)}} - 2\sqrt{p})$ is again given by $\mathcal{F}_2(x)$ [29].

Conclusion.—To conclude, using methods of many-body physics, where appropriate cutoffs ϵ_i 's have been introduced [see Fig. 1(c)], we have obtained exact results for the distribution of the maximal height for p watermelons with a wall (5) and without walls (7), which is physically relevant to describe the geometrical properties of dislocations arising in $p \times 1$ commensurate adsorbed phases close to the C-IC transition. Our expressions ex-

plain the discrepancy between the estimates of BM [15] and the true asymptotic behaviors for the average $\langle H_p \rangle$. Besides, we obtained a quantitative description of the preasymptotic regime actually measured in the numerical experiments of BM. We hope that the path-integral method presented here, which is rather general, and the precise connection to RMT will allow further future studies.

We thank P. Ferrari for useful discussions.

-
- [1] P. G. de Gennes, J. Chem. Phys. **48**, 2257 (1968).
 - [2] M. E. Fisher, J. Stat. Phys. **34**, 667 (1984).
 - [3] J. W. Essam, A. J. Guttmann, Phys. Rev. E **52**, 5849 (1995).
 - [4] A. J. Bray and K. Winkler, J. Phys. A **37**, 5493 (2004).
 - [5] H. L. Richards and T. L. Einstein, Phys. Rev. E **72**, 016124 (2005).
 - [6] P. Ferrari and M. Prähofer, Markov Processes Relat. Fields **12**, 203 (2006).
 - [7] K. Johansson, Probab. Theory Relat. Fields **123**, 225 (2002).
 - [8] M. Katori and H. Tanemura, J. Math. Phys. (N.Y.) **45**, 3058 (2004).
 - [9] C. A. Tracy and H. Widom, Ann. Appl. Probab. **17**, 953 (2007).
 - [10] F. J. Dyson, J. Math. Phys. (N.Y.) **3**, 1191 (1962); **3**, 1199 (1962).
 - [11] D. A. Huse and M. E. Fisher, Phys. Rev. B **29**, 239 (1984).
 - [12] S. N. Majumdar and A. Comtet, Phys. Rev. Lett. **92**, 225501 (2004); J. Stat. Phys. **119**, 777 (2005).
 - [13] G. Schehr and S. N. Majumdar, Phys. Rev. E **73**, 056103 (2006); G. Györgyi *et al.*, Phys. Rev. E **75**, 021123 (2007); T. W. Burkhardt *et al.*, Phys. Rev. E **76**, 041119 (2007).
 - [14] A. N. Borodin and P. Salminen, *Handbook of Brownian Motion-Facts and Formulae* (Birkhäuser, Basel 2002).
 - [15] N. Bonichon and M. Mosbah, Theor. Comput. Sci. **307**, 241 (2003).
 - [16] C. Krattenthaler, A. J. Guttmann, and X. G. Viennot, J. Phys. A **33**, 8835 (2000).
 - [17] M. Fulmek, The Electronic J. Combinatorics **14**(1), R64/1-20 (2007).
 - [18] M. Katori, M. Izumi, and N. Kobayashi, J. Stat. Phys. **131**, 1067 (2008).
 - [19] T. Feierl, arXiv:0802.2691.
 - [20] T. Feierl, arXiv:0806.0037.
 - [21] S. R. Finch, "Stars and watermelons", unpublished, <http://algo.inria.fr/resolve/wm.pdf>.
 - [22] M. Prähofer and H. Spohn, J. Stat. Phys. **108**, 1071 (2002).
 - [23] F. Gillet, arXiv:math.PR/0307204.
 - [24] S. N. Majumdar, Curr. Sci. **89**, 2076 (2005).
 - [25] M. L. Mehta, *Random Matrices* (Academic Press, New York, 1991).
 - [26] K. L. Chung, Bull. Am. Math. Soc. **81**, 742 (1975).
 - [27] Details will be published elsewhere.
 - [28] C. A. Tracy and H. Widom, Commun. Math. Phys. **159**, 151 (1994); **177**, 727 (1996).
 - [29] K. Johansson, Commun. Math. Phys. **209**, 437 (2000); I. M. Johnstone, Ann. Stat. **29**, 295 (2001).

Convex Hull of N Planar Brownian Motions: Exact Results and an Application to Ecology

Julien Randon-Furling,¹ Satya N. Majumdar,¹ and Alain Comtet^{1,2}

¹*Laboratoire de Physique Théorique et Modèles Statistiques (UMR CNRS 8626) Université Paris-Sud, Bâtiment 100, 91405 Orsay Cedex, France*

²*Institut Henri Poincaré, Université Pierre et Marie Curie-Paris 6, Paris, France*
(Received 2 July 2009; published 29 September 2009)

We compute exactly the mean perimeter and area of the convex hull of N independent planar Brownian paths each of duration T , both for open and closed paths. We show that the mean perimeter $\langle L_N \rangle = \alpha_N \sqrt{T}$ and the mean area $\langle A_N \rangle = \beta_N T$ for all T . The prefactors α_N and β_N , computed exactly for all N , increase very slowly (logarithmically) with increasing N . This slow growth is a consequence of extreme value statistics and has interesting implications in an ecological context in estimating the home range of a herd of animals with a population size N .

DOI: 10.1103/PhysRevLett.103.140602

PACS numbers: 05.40.Fb, 02.50.-r, 87.23.Cc

Ecologists often need to estimate the home range of an animal or a group of animals, in particular, for habitat-conservation planning [1]. The home range of a group of animals simply means the two-dimensional space over which they typically move around in search of food. There exist various methods to estimate this home range, based on the monitoring of the positions of the animals over a certain period of time [2]. One method consists in drawing the minimum convex polygon enclosing all monitored positions, called the convex hull. While this may seem simpleminded, it remains, under certain circumstances, the best way to proceed [3]. The monitored positions, for one animal, will appear as the vertices of a path whose statistical properties will depend on the type of motion the animal is performing. In particular, during phases of food searching known as foraging, the monitored positions can be described as the vertices of a random walk in the plane [4,5]. For animals whose daily motion consists mainly in foraging, quantities of interest about their home range, such as its perimeter and area, can be estimated through the average perimeter and area of the convex hull of the corresponding random walk [Fig. 1(a)]. If the recorded positions are numerous (which might result from a very fine and/or long monitoring), the number of steps of the random walker becomes large and to a good approximation the trajectory of a discrete-time planar random walk (with finite variance of the step sizes) can be replaced by a continuous-time planar Brownian motion of a certain duration T .

The home range of a single animal can thus be characterized by the mean perimeter and area of the convex hull of a planar Brownian motion of duration T starting at origin O . Both “open” (where the end point of the path is free) and “closed” (that are constrained to return to the origin in time T) are of interest. The latter corresponds, for instance, to an animal returning every night to its nest after spending the day foraging in the surroundings. For an open path, the mean perimeter $\langle L_1 \rangle = \sqrt{8\pi T}$ and the mean area $\langle A_1 \rangle = \pi T/2$ are known in the mathematics

literature [6–8]. For a closed path, only the mean perimeter is known [9]: $\langle L_1 \rangle = \sqrt{\pi^3 T/2}$.

In any given habitat, however, an animal is hardly alone; they live in herds with typically a large population size N . To study their global home range via the convex hull model, one needs to study the convex hull of N planar Brownian motions. Assuming that the animals do not develop any substantial interaction between them during foraging, this leads us to the main question addressed in this Letter: What is the mean perimeter and area of the convex hull of N independent planar Brownian motions (both open and closed) each of duration T ? This question is of vital importance in ecological conservation: if the population size increases by, say tenfold, by how much should one increase the home range, i.e., the conservation area?

Using the standard scaling property of Brownian motion, length scale \sim (time scale)^{1/2}, it follows that the mean perimeter and area of the global convex hull of N independent Brownian paths will behave as $\langle L_N \rangle = \alpha_N \sqrt{T}$ and $\langle A_N \rangle = \beta_N T$ for all T . The main challenge is to estimate the N dependence of the prefactors α_N and β_N . The central result of this Letter is to provide exact formulas for α_N and β_N for all N , both for open and closed paths.

Most interestingly, we find that the mean perimeter and area of the convex hull increases very slowly with N for large N : $\alpha_N \simeq 2\pi\sqrt{2}\ln N$, $\beta_N \simeq 2\pi\ln N$ for open paths,

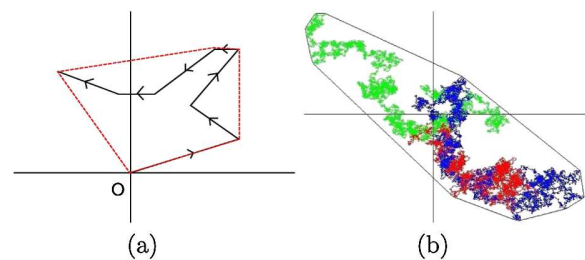


FIG. 1 (color online). Convex hull of (a) a seven-step random walk starting at O and (b) three closed Brownian paths.

and $\alpha_N(c) \simeq \pi\sqrt{2\ln N}$, $\beta_N(c) \simeq \frac{\pi}{2} \ln N$ for closed paths, c referring to closed paths. This leads to our main conclusion: the home range increases very slowly with increasing population size. Indeed, the $\ln N$ behavior of the prefactor $\beta_N(c)$ in the mean area indicates that, for instance, a tenfold increase in the number of animals in the herd will result, on average, in the addition of only about 3.6 T units of area to the home range of the herd—good news for conservation.

To proceed, let I denote any set of points on the plane with coordinates (x_i, y_i) . Let C denote the convex hull of I , i.e., the minimal convex polygon enclosing this set. Geometrically, the convex hull can be constructed very simply: consider any arbitrary direction from the origin O specified by the angle θ with respect to the x axis. Bring a straight line from infinity perpendicularly along direction θ and stop when it touches a point of the set I [e.g., in Fig. 2, this point has coordinates (x_{k^*}, y_{k^*})]. Let $M(\theta)$ denote the Euclidean distance of this perpendicular line from the origin when it stops, measuring the maximal extension of the set I along the direction θ . This support function can be simply written as

$$M(\theta) = \max_{i \in I} \{x_i \cos \theta + y_i \sin \theta\}. \quad (1)$$

Knowing $M(\theta)$, one can express the perimeter and the area of the convex hull C via Cauchy's formulas [10]:

$$L = \int_0^{2\pi} d\theta M(\theta) \quad (2)$$

$$A = \frac{1}{2} \int_0^{2\pi} d\theta [M^2(\theta) - (M'(\theta))^2], \quad (3)$$

where $M'(\theta) = dM/d\theta$. In this Letter we show how these two formulas can be used to compute the mean perimeter and area of the convex hull of N planar Brownian motions.

To illustrate the main idea let us start with a single open Brownian path; the generalizations to $N > 1$ and for closed paths will follow. In this case the set I consists of the vertices of a Brownian path of duration T starting at the

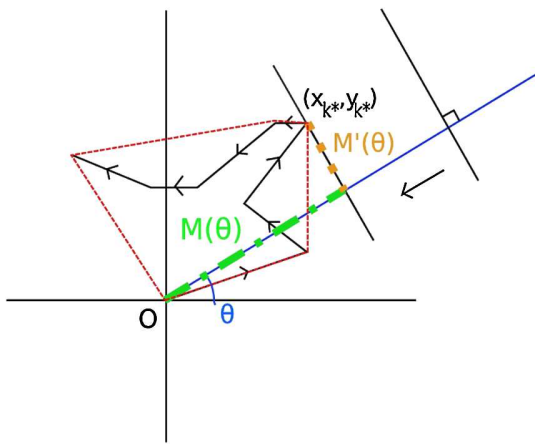


FIG. 2 (color online). Support function of a seven-step random walk.

origin O . Since it is a continuous path, we can conveniently label the coordinates by $\mathcal{B}(\tau) = (x(\tau), y(\tau))$ with $0 \leq \tau \leq T$. Here $x(\tau)$ and $y(\tau)$ are just two independent one-dimensional Brownian paths each of duration T and evolve via the Langevin equations: $\dot{x}(\tau) = \eta_x(\tau)$ and $\dot{y}(\tau) = \eta_y(\tau)$ where $\eta_x(\tau)$ and $\eta_y(\tau)$ are independent Gaussian white noises, each with zero mean and delta correlated, e.g., $\langle \eta_x(\tau) \eta_x(\tau') \rangle = \delta(\tau - \tau')$ and the same for η_y . Clearly then $\langle x^2(\tau) \rangle = \tau$ and $\langle y^2(\tau) \rangle = \tau$.

Let us now consider a fixed direction θ . Then, $z_\theta(\tau) = x(\tau) \cos \theta + y(\tau) \sin \theta$ and $h_\theta(\tau) = -x(\tau) \sin \theta + y(\tau) \cos \theta$ are also two independent one-dimensional Brownian motions (each of duration T) parametrized by θ . Then $M(\theta)$ in Eq. (1) is simply the maximum of the one-dimensional Brownian motion $z_\theta(\tau)$ over the time interval $\tau \in [0, T]$, i.e., $M(\theta) = \max[z_\theta(\tau)]$. Let τ^* be the time at which this maximum is achieved. Then, $M(\theta) = z_\theta(\tau^*) = x(\tau^*) \cos \theta + y(\tau^*) \sin \theta$. Taking derivative with respect to θ gives $M'(\theta) = -x(\tau^*) \sin \theta + y(\tau^*) \cos \theta = h_\theta(\tau^*)$. Thus, while $M(\theta)$ is the maximum value of the first Brownian motion $z_\theta(\tau)$, $M'(\theta)$ is the value of the second independent Brownian motion $h_\theta(\tau)$ but at the time $\tau = \tau^*$ at which the first one achieves its maximum [see Figs. 3(a) and 3(b)]. Note, in particular, that for $\theta = 0$, $z_0(\tau) = x(\tau)$ and $h_0(\tau) = y(\tau)$. Thus $M(0)$ is the maximum of $x(\tau)$ in $\tau \in [0, T]$ while $M'(0) = y(\tau^*)$ is the value of y at the time τ^* when x achieves its maximum.

Taking the expectation value and using isotropy of the mean over the choice of directions, it follows that

$$\langle L \rangle = 2\pi \langle M(0) \rangle \quad (4)$$

$$\langle A \rangle = \pi (\langle [M(0)]^2 \rangle - \langle [M'(0)]^2 \rangle). \quad (5)$$

For a one-dimensional Brownian motion $x(\tau)$ over $[0, T]$, the distribution of its maximum is well known [11]. The cumulative distribution, $Q_1(m, T) = \text{Prob}[M(0) \leq m] = \text{erf}(m/\sqrt{2T})$, where $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du$. The first two moments can be easily computed: $\langle M(0) \rangle = \sqrt{2T/\pi}$ and $\langle [M(0)]^2 \rangle = T$. Equation (4) then gives the mean perimeter, $\langle L_1 \rangle = \alpha_1 \sqrt{T}$ with $\alpha_1 = \sqrt{8\pi}$. The calculation of the mean area is slightly trickier since we need $\langle [M'(0)]^2 \rangle$. We first note that for any fixed time τ^* , $E[y^2(\tau^*)] = \tau^*$ (since it is a free Brownian motion) where the expectation is taken over all realizations of the process y at fixed τ^* . But τ^*

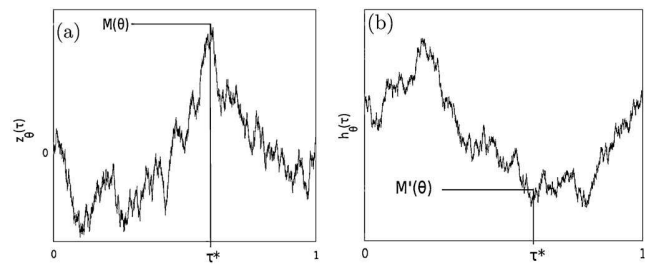


FIG. 3. (a) Location τ^* of the maximum $M(\theta)$ of $z_\theta(\tau)$ and (b) corresponding value of $M'(\theta) = h_\theta(\tau^*)$.

itself is a random variable, being the time at which the first process x achieves its maximum. Its distribution is also well known and given by the celebrated arcsine law of Lévy [12]: $P(\tau^*, T) = [\tau^*(T - \tau^*)]^{-1/2}/\pi$. Thus, averaging further over τ^* taken from this distribution, we finally get $\langle [M'(0)]^2 \rangle = \langle \tau^* \rangle = T/2$. Equation (5) then gives the exact mean area $\langle A_1 \rangle = \beta_1 T$ with $\beta_1 = \pi/2$.

For a single closed path, the analysis is similar, except that $x(\tau)$ and $y(\tau)$ are now two independent one-dimensional Brownian bridges of duration T ; i.e., both start at the origin but are constrained to return to the origin at time T : $x(0) = x(T) = 0$ and $y(0) = y(T) = 0$. The distribution of the maximum of a Brownian bridge is also well known [11] and its first two moments are given by $\langle M(0) \rangle = \sqrt{\pi T}/8$ and $\langle [M(0)]^2 \rangle = T/2$. Thus, the mean perimeter of the convex hull is given from Eq. (4): $\langle L_1 \rangle = \alpha_1(c)\sqrt{T}$ with $\alpha_1(c) = \sqrt{\pi^3}/2$. To compute the mean area, we note that for a Brownian bridge $y(\tau)$, at fixed time τ^* , $E[y^2(\tau^*)] = \tau^*(T - \tau^*)/T$. Moreover, the distribution of τ^* , the time at which a bridge achieves its maximum, is known to be uniform [11]. Thus, taking average over τ^* with uniform distribution, $P(\tau^*, T) = 1/T$, we get $\langle [M'(0)]^2 \rangle = \langle y^2(\tau^*) \rangle = T/6$. Equation (5) then gives the exact mean area of the convex hull of a planar closed Brownian path: $\langle A_1 \rangle = \beta_1(c)T$ with $\beta_1(c) = \pi/3$. To our knowledge, this result, even for a single closed path, appears to be new.

This method can then be generalized to N independent planar Brownian paths, open or closed. We now have two sets of N Brownian paths: $x_j(\tau)$ and $y_j(\tau)$ ($j = 1, 2, \dots, N$). All paths are independent of each other. Since isotropy holds, we can still use Eqs. (4) and (5), except that $M(0)$ now denotes the global maximum of a set of N independent one-dimensional Brownian paths (or bridges for closed paths) $x_j(\tau)$ ($j = 1, 2, \dots, N$), each of duration T ,

$$M(0) = \max_{1 \leq j \leq N} \max_{0 \leq \tau \leq T} [x_1(\tau), x_2(\tau), \dots, x_N(\tau)]. \quad (6)$$

Let j_* and τ^* denote the label of the path and the time at which this global maximum is achieved. Then, using argument similar to the $N = 1$ case, it is easy to see that $M'(0) = y_{j_*}(\tau^*)$, i.e., the position of the j_* -th y path at the time when the x paths achieve their global maximum.

To compute the first two moments of $M(0)$, we first compute the distribution $P_N[M(0), T]$ of the global maximum of N independent Brownian paths (or bridges) $x_j(\tau)$. This is a standard extreme value calculation. Consider first N open Brownian paths. It is easier to compute the cumulative probability, $Q_N(m, T) = \text{Prob}[M(0) \leq m]$. Since the Brownian paths are independent, it follows that $Q_N(m, T) = [Q_1(m, T)]^N$, where $Q_1(m, T) = \text{erf}(m/\sqrt{2T})$ for a single path mentioned before. Knowing this cumulative distribution $Q_N(M(0), T)$, the first two moments $\langle M(0) \rangle$ and $\langle [M(0)]^2 \rangle$ can be computed for all N . Using the result for $\langle M(0) \rangle$ in Eq. (4) gives us the mean perimeter, $\langle L_N \rangle = \alpha_N \sqrt{T}$ with

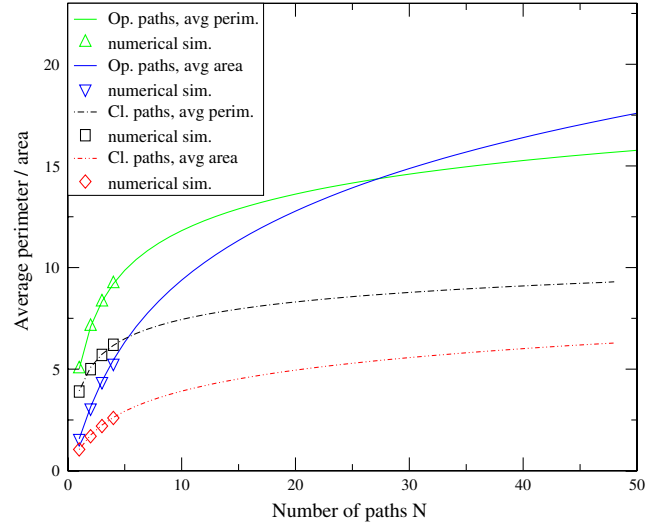


FIG. 4 (color online). Setting $T = 1$, the analytical results for average perimeter α_N [Eq. (7)] and area β_N [Eq. (9)] of N open (Op.) Brownian paths, and similarly the average perimeter $\alpha_N(c)$ [Eq. (10)] and area $\beta_N(c)$ [Eq. (11)] of N closed (Cl.) Brownian paths, plotted against N . The symbols denote results from numerical simulations (up to $N = 4$).

$$\alpha_N = 4N\sqrt{2\pi} \int_0^\infty du u e^{-u^2} [\text{erf}(u)]^{N-1}. \quad (7)$$

The first few values are $\alpha_1 = \sqrt{8\pi} = 5.013\dots$, $\alpha_2 = 4\sqrt{\pi} = 7.089\dots$, $\alpha_3 = 24 \tan^{-1}(1/\sqrt{2})/\sqrt{\pi} = 8.333\dots$, etc. (see Fig. 4 for a plot of α_N versus N). For large N , one can analyze the integral in Eq. (7) by the saddle point method giving [13] $\alpha_N \simeq 2\pi\sqrt{2} \ln N$. This logarithmic dependence on N is thus a direct consequence of extreme value statistics [14], and the calculation of the mean perimeter of the convex hull of N paths is a nice application of extreme value statistics.

To compute the mean area, we need to calculate $\langle [M'(0)]^2 \rangle$ in Eq. (5). We proceed as in the $N = 1$ case. For a fixed label j and fixed time τ , the expectation $E[y_j^2(\tau)] = \tau$, which follows from the fact that $y_j(\tau)$ is simply a Brownian motion. Thus, $E[y_{j_*}^2(\tau^*)] = \tau^*$. Next, we need to average over τ^* , which is the time at which the global maximum in Eq. (6) happens. The distribution $P_N(\tau^*, T)$ of the time of global maximum of N independent Brownian motions, to our knowledge, is not known in the probability literature. We were able to compute this exactly for all N [13]. Skipping details, we find that $P_N(\tau^*, T) = \frac{1}{T} F_N(\tau^*/T)$ where

$$F_N(z) = \frac{a_N}{\sqrt{z(1-z)}} \int_0^\infty dx x e^{-x^2} [\text{erf}(x\sqrt{z})]^{N-1} \quad (8)$$

and a_N is a normalization constant fixed by $\int_0^1 F_N(z) dz = 1$. It is easy to check that for $N = 1$, it reproduces the arcsine law mentioned before. Averaging over τ^* drawn from this distribution, we can then compute $\langle [M'(0)]^2 \rangle = \int_0^1 \tau^* P_N(\tau^*, T) d\tau^*$. Substituting this in Eq. (5) gives the

exact mean area for all N , $\langle A_N \rangle = \beta_N T$ with

$$\beta_N = 4N\sqrt{\pi} \int_0^\infty du u [\operatorname{erf}(u)]^{N-1} [ue^{-u^2} - g(u)], \quad (9)$$

where $g(u) = (1/2\sqrt{\pi}) \int_0^1 [e^{-u^2/t} dt / \sqrt{t(1-t)}]$. For example, the first few values are given by $\beta_1 = \pi/2 = 1.570\dots$, $\beta_2 = \pi = 3.141\dots$, $\beta_3 = \pi + 3 - \sqrt{3} = 4.409\dots$, etc. (Figure 4 shows a plot of β_N versus N .) The large N analysis gives [13] $\beta_N \simeq 2\pi \ln N$. Thus for large N , the shape of the convex hull approaches a circle of radius $\sqrt{2 \ln N}$ which, incidentally, coincides with the set of distinct sites visited by N Brownian motions [15].

For N closed Brownian planar paths one proceeds in a similar way. Without repeating the analysis, we just mention our main results [13]. The mean perimeter and area are given by $\langle L_N \rangle = \alpha_N(c)\sqrt{T}$ and $\langle A_N \rangle = \beta_N(c)T$ where, for all N ,

$$\alpha_N(c) = \frac{\pi^{3/2}}{\sqrt{2}} \sum_{k=1}^N \binom{N}{k} \frac{(-1)^{k+1}}{\sqrt{k}} \quad (10)$$

$$\beta_N(c) = \frac{\pi}{2} \left[\sum_{k=1}^N \frac{1}{k} - \frac{N}{3} + \frac{1}{2} \sum_{k=2}^N (-1)^k f(k) \right] \quad (11)$$

and

$$f(k) = \binom{N}{k} (k-1)^{-3/2} [k \tan^{-1}(\sqrt{k-1}) - \sqrt{k-1}].$$

The first few values are $\alpha_1(c) = \sqrt{\pi^3/2} = 3.937\dots$, $\alpha_2(c) = \sqrt{\pi^3}(\sqrt{2}-1/2) = 5.090\dots$, $\alpha_3(c) = \sqrt{\pi^3}(3/\sqrt{2}-3/2+1/\sqrt{6}) = 5.732\dots$, and $\beta_1(c) = \pi/3 = 1.047\dots$, $\beta_2(c) = \pi(4+3\pi)/24 = 1.757\dots$, $\beta_3(c) = 2.250\dots$, etc. [see Fig. 4 for a plot of $\alpha_N(c)$ and $\beta_N(c)$ versus N]. Large N analysis shows that [13] $\alpha_N(c) \simeq \pi\sqrt{2 \ln N}$ and $\beta_N(c) \simeq \frac{\pi}{2} \ln N$, smaller, respectively, by a factor 1/2 and 1/4 than the corresponding results for open paths.

We have also computed the mean perimeter and area of the convex hull of $N = 1, 2, 3, 4$ Brownian paths (both open and closed) via numerical simulations. For each N , we constructed N independent normal Gaussian random walks of 10^4 steps each with a time step $\Delta\tau = 10^{-4}$. For each realization of the N walks, we constructed the convex hull using the Graham scan algorithm [16] and computed its perimeter and area and then averaged over 10^3 samples. We find excellent agreement with our analytical predictions (Fig. 4).

The method presented here is general and can, in principle, be applied to compute the mean perimeter and area of the convex hull of any set of random points. In fact, when the set consists of just random points, each drawn independently from a given distribution, the statistics of the perimeter and area of their convex hull has been studied before [17]. Our method, using extreme value statistics, can easily recover these results [13], but can go further, e.g., when the points are correlated as in the case of Brownian paths.

This work leads to several interesting open questions. It would be interesting to extend the results presented here for normal diffusion to the case where the trajectories of animals such as birds, deer, or spider monkeys undergo anomalous diffusion, e.g., Lévy flights [18–20]. Another interesting question concerns the effect of interactions or collective behavior of the animals on the statistics of their convex hulls. For animals like birds that move in three-dimensional space, it would be interesting to study the statistics of the convex polytope of their trajectories, such as the mean surface area and the volume of such a polytope [13]. Finally, the distribution of the time of maximum τ^* , a crucial ingredient in our method, has recently been computed exactly for a variety of constrained one-dimensional Brownian motions [21]. These results may be useful to study the statistics of convex hulls of constrained planar Brownian paths (see, e.g., [22]).

We thank D. Dhar and H. Larralde for useful discussions.

-
- [1] D. Murphy and B. Noon, *Ecol. Appl.* **2**, 3 (1992).
 - [2] B.J. Worton, *Biometrics* **51**, 1206 (1995).
 - [3] S.A. Boyle *et al.*, *Folia Primatol.* **80**, 33 (2009).
 - [4] F. Bartumeus *et al.*, *Ecology* **86**, 3078 (2005).
 - [5] H.C. Berg, *Random Walks in Biology* (Princeton University Press, New York, 1983); L. Edelman-Keshet, *Mathematical Models in Biology* (McGraw Hill, Boston, 1988).
 - [6] L. Takács, *Am. Math. Mon.* **87**, 142 (1980).
 - [7] M. El Bachir, Ph.D. thesis, Université Paul Sabatier, Toulouse, France, 1983.
 - [8] G. Letac, *J. Theor. Probab.* **6**, 385 (1993).
 - [9] A. Goldman, *Probab. Theory Relat. Fields* **105**, 57 (1996).
 - [10] A. Cauchy, *Mem. Acad. Sci. Inst. Fr.* **22**, 3 (1850); L.A. Santaló, *Integral Geometry and Geometric Probability* (Addison-Wesley, Reading, MA, 1976).
 - [11] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1968).
 - [12] P. Lévy, *Compos. Math.* **7**, 283 (1939).
 - [13] Details will be published elsewhere.
 - [14] E.J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
 - [15] H. Larralde *et al.*, *Phys. Rev. A* **45**, 7128 (1992).
 - [16] R.L. Graham, *Inf. Proc. Lett.* **1**, 132 (1972).
 - [17] A. Rényi and R. Sulanke, *Z. Wahrsch. Verw. Geb.* **2**, 75 (1963).
 - [18] G.M. Viswanathan *et al.*, *Nature (London)* **381**, 413 (1996); see also A.M. Edwards *et al.*, *Nature (London)* **449**, 1044 (2007).
 - [19] J. Klafter and I.M. Sokolov, *Phys. World* **18**, 29 (2005).
 - [20] G. Ramos-Fernández *et al.*, *Behav. Ecol. Sociobiol.* **55**, 223 (2004).
 - [21] J. Randon-Furling and S.N. Majumdar, *J. Stat. Mech.* (2007) P10008; S.N. Majumdar *et al.*, *J. Phys. A* **41**, 365005 (2008).
 - [22] P. Biane and G. Letac, arXiv:0905.2256.

Bibliographie

- [1] Aldama A. A. and Ramirez A. I. Dam design flood estimation based on bivariate extreme-value distributions. In *The extremes of the extremes : extraordinary floods*, pages 257–262, Wallingford, 2002. IAHS Press. 1
- [2] A. Bengtsson and C. Nilsson. Extreme value modelling of storm damage in Swedish forests. *Nat. Hazards Earth Syst. Sci.*, 7 :515–521, 2007. 1
- [3] E. Alvarado, D. V. Sandberg, and S. G. Pickford. Modeling large forest fires as extreme events. *Northwest Science*, pages 66–75, 1998. 1
- [4] B. H. Lavenda and E. Cipollone. Extreme value statistics and thermodynamics of earthquakes : large earthquakes. *Annali di Geofisica*, 43(3) :469–484, 2001. 1
- [5] J. N. Al-Abbasi and K. J. Fahmi. Estimating maximum magnitude earthquakes in Iraq using extreme value statistics. *Geophysical journal of the Royal Astronomical Society*, 82(3) :535–548, 1985. 1
- [6] J. F. Eichner, J. W. Kantelhardt, A. Bunde, and S. Havlin. Extreme value statistics in records with long-term persistence. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(1) :016130, 2006. 1
- [7] Einmahl J. H. J. and Magnus J. R. Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, 103(484) :1382–1391, 2008. 1
- [8] J.-P. Bouchaud and M. Potters. *Theory of financial risk and deriva-*

tive pricing : from statistical physics to risk management. Cambridge University Press, Cambridge, 2nd edition, 2003. 1, 2, 34

- [9] H.E. Daniels. The statistical theory of the strength of bundles and threads. *Proc. Royal Soc.*, A183 :405–435, 1945. 1
- [10] Z.P. Bazant, S.D. Pang, M. Vorechovsky, D. Novak, and R. Pukl. Statistical size effect in quasibrittle materials : Computation and extreme value theory. In V. C. Li, K. J. Willam, C. K. Y. Leung, and S. L. Billington, editors, *5th Int. Conference FraMCoS - Fracture Mechanics of Concrete and Concrete Structures*, volume 1, pages 189–196. 1
- [11] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Cambridge Philos. Soc.*, 24 :180–190, 1928. 1, 108
- [12] M. Fréchet. Sur la loi de probabilité de l'écart maximum. *Ann. de la Soc. Polonaise de Math*, 1927, 6 :93, 1927. 1, 108
- [13] E.J. Gumbel. Les valeurs extrêmes des distributions statistiques. *Annales de l'I.H.P.*, 5(2) :115–158, 1935. 1, 103, 108
- [14] B. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.*, 44(3) :423–453, 1943. 1, 103, 108
- [15] J.-P. Bouchaud and M. Mézard. Universality classes for extreme value statistics. *J. Phys. A.*, 30 :7997–8015, 1997. 1, 108
- [16] S. N. Majumdar and P. L. Krapivsky. Extreme value statistics and traveling fronts : various applications. *Physica A : Statistical Mechanics and its Applications*, 318(1-2) :161–170, 2003. 1, 108
- [17] G. Biroli, J.-P. Bouchaud, and M. Potters. Extreme value problems in random matrix theory and other disordered systems. *J. Stat. Mech.*, page P07019, 2007. 1
- [18] B. Duplantier. Le mouvement brownien, "divers et ondoyant". In *Séminaire Poincaré*, volume 1, pages 155–212, 2005. 1, 2, 156

- [19] E. Frey and K. Kroy. Brownian motion : a paradigm of soft matter and biological physics. *Annalen der Physik*, 14(1–3) :20–50, 2005. 1, 2
- [20] S. N. Majumdar. Brownian functionals in Physics and Computer Science. *Current Science*, 89 :2075, 2005. 1, 2, 20, 55
- [21] D. Chowdhury. 100 years of Einstein’s theory of Brownian motion : From pollen grains to protein trains. *Resonance*, 10(11) :42–54, 2005. 1, 2
- [22] J.-P. Kahane. Le mouvement brownien et son histoire, réponses à quelques questions. [http ://images.math.cnrs.fr/Le-mouvement-brownien-et-son.html](http://images.math.cnrs.fr/Le-mouvement-brownien-et-son.html), 2006. 1, 156
- [23] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. d. Physik*, 17 :549–560, 1905. 1, 155
- [24] R. Brown. A Brief Account of Microscopical Observations Made in the Months of June, July and August, 1827, on the Particles Contained in the Pollen of Plants ; and on the General Existence of Active Molecules in Organic and Inorganic Bodies. *Edinburgh Journal of Science*, 9 :336–346, 1828. 2, 155
- [25] J. Perrin. Mouvement brownien et réalité moléculaire. *Ann. Chim. Phys.*, 18 :5–114, 1909. 2, 156
- [26] Freeman J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6) :1191–1198, 1962. 2
- [27] P. Reimann. Brownian motors : noisy transport far from equilibrium. *Physics reports*, 361 :57–265, 2002. 2
- [28] S. Chandrasekhar. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15(1) :1–89, 1943. 2

- [29] P. Chatterjee, L. Hernquist, and A. Loeb. Brownian motion in gravitationally interacting systems. *Phys. Rev. Lett.*, 88(12) :121103, 2002. 2
- [30] P. G. Saffman and M. Delbrück. Brownian motion in biological membranes. *PNAS*, 72(8) :3111–3113, 1975. 2
- [31] P. Hänggi. Brownian machinery in physics and biology. In G. Bosman, editor, *Noise in physical systems and $\frac{1}{f}$ fluctuations*, pages 397–399. World Scientific, 2001. 2
- [32] R. D. Astumian. Thermodynamics and Kinetics of a Brownian Motor. *Science*, 276(5314) :917–922, 1997. 2
- [33] G Louchard. Brownian motion and algorithm complexity. *BIT*, 26(1) :17–34, 1986. 2
- [34] R. T. Smythe and J. Wellner. Stochastic analysis of shell sort. *Algorithmica*, 31(3) :442–457, 2001. 2
- [35] M. F. M. Osborne. Brownian motion in the stock market. *Operations Research*, 7(2) :145–173, 1959. 2
- [36] S. N. Majumdar and J.-P. Bouchaud. Optimal time to sell a stock in the Black–Scholes model : comment on ”Thou shalt buy and hold”, by A. Shiryaev, Z. Xu and X.Y. Zhou. *Quantitative Finance*, 8(8) :753–760, 2008. 2, 35, 42, 150
- [37] S. N. Majumdar and A. Comtet. Exact maximal height distribution of fluctuating interfaces. *Phys. Rev. Letters*, 92 :225501, 2004. 2, 11, 20, 37, 55
- [38] S. N. Majumdar and A. Comtet. Airy distribution function : from the area under a Brownian excursion to the maximal height of fluctuating interfaces. *J. Stat. Phys.*, 119 :777–826, 2005. 2, 55
- [39] M. J. Kearney. On a random area variable arising in discrete-time queues and compact directed percolation. *J. Phys. A : Math. Gen.*, 37 :8421–8431, 2004. 2, 34

- [40] P. Lévy. Sur certains processus stochastiques homogènes. *Comp. Math.*, 7 :283–339, 1939. 5, 16, 104, 157
- [41] W. Feller. *An Introduction to Probability Theory and its Applications*. Wiley, New York, 1968. 5
- [42] M. E. Fisher. Walks, walls, wetting and melting. *J. Stat. Phys.*, 34(667–729), 1984. 5, 6, 54
- [43] P.-G. de Gennes. Soluble models for fibrous structures with steric constraints. *J. Chem. Phys.*, 48 :2257–2259, 1968. 5, 54
- [44] W. Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.*, 7 :143–149, 1979. 20, 174
- [45] K. L. Chung. Maxima in Brownian excursions. *Bull. Amer. Math. Soc.*, 81 :742–745, 1975. 24
- [46] K. L. Chung. Excursions in Brownian motion. *Ark. Mat.*, 14 :155–177, 1976. 24
- [47] D. P. Kennedy. The distribution of the maximum Brownian excursion. *J. Appl. Prob.*, 13 :371–376, 1976. 24
- [48] S.N. Majumdar, J. Randon-Furling, M. Kearney, and M. Yor. On the time to reach maximum for a variety of constrained Brownian motions. *J. Phys. A : Math. Theor.*, 41(365005), 2008. 24, 27
- [49] R. T. Durrett, D. L. Iglehart, and D. R. Miller. Weak convergence to Brownian meander and Brownian excursion. *Ann. Probability*, 5(1) :117–129, 1977. 27
- [50] R. T. Durrett and D. L. Iglehart. Functionals of Brownian meander and Brownian excursion. *Ann. Prob.*, 5 :130–135, 1977. 27
- [51] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, 1964. 30, 39, 123

- [52] S. Redner. *A guide to first-passage processes*. Cambridge University Press, Cambridge, 2001. 33
- [53] M. J. Kearney and S. N. Majumdar. On the area under a continuous time Brownian motion till its first-passage time. *J. Phys. A : Math. Gen.*, 38 :4097–4104, 2005. 33, 36, 39
- [54] J. Randon-Furling and S. N. Majumdar. Distribution of the time at which the deviation of a Brownian motion is maximum before its first-passage time. *J. Stat. Mech.*, 2007. 33, 42
- [55] S. Asmussen. *Applied probability and queues*. Springer, New York, 2nd edition, 2003. 34
- [56] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, 2000. 41, 86, 89, 92
- [57] H. Bateman. *Tables of Integral Transforms*. McGraw-Hill, New York, 1954. 41
- [58] T. A Witten. Polymer solutions : a geometric introduction. *Rev. Mod. Phys.*, 70(4) :1531–1544, 1998. 54
- [59] J. W. Essam and A. J. Guttmann. Vicious walkers and directed polymer networks in general dimensions. *Phys. Rev. E*, 52 :5849, 1995. 54
- [60] A. J. Bray and K. Winkler. Vicious walkers in a potential. *J. Phys. A : Math. Gen.*, 37 :5493–5501, 2004. 54
- [61] H. L. Richards and T. L. Einstein. Beyond the Wigner distribution : Schrödinger equations and terrace width distributions. *Phys. Rev. E*, 72 :016124, 2005. 54
- [62] P. Ferrari and M. Praehofer. One-dimensional stochastic growth and Gaussian ensembles of random matrices. *Markov Proc. Relat. Fields*, 12 :203–236, 2005. 54

- [63] G. Schehr, S. N. Majumdar, A. Comtet, and J. Randon-Furling. Exact distribution of the maximal height of p vicious walkers. *Phys. Rev. Lett.*, 101 :150601, 2008. 55, 56
- [64] C. Nadal and S. N. Majumdar. Non-intersecting Brownian interfaces and Wishart random matrices. *Phys. Rev. E*, 79 :061117, 2009. 55
- [65] T. D. Rossing, editor. *Handbook of acoustics*. Springer, 2007. 59
- [66] R. L. Dobrushin, R. Kotecký, and S. Shlosman. *Wulff construction : a global shape from local interaction*. Translations of Mathematical Monographs. American Mathematical Society, Providence, Rhode Island, 1992. 59
- [67] J. Haushofer, C. I. Bake, M. S. Livingstone, and N. Kanwisher. Privileged coding of convex shapes in human object-selective cortex. *Journal of Neurophysiology*, 100 :753–762, 2008. 59
- [68] S.G. Akl and G.T. Toussaint. Efficient convex hull algorithms for pattern recognition applications. In *International Conference on Pattern Recognition*, pages 483–487, 1978. 59
- [69] R. Meier, F. Ackermann, G. Herrmann, S. Posch, and G. Sagerer. Segmentation of molecular surfaces based on their convex hull. In *International Conference on Image Processing (ICIP'95)*, volume 3, page 3552, 1995. 59
- [70] R.L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Information Processing Letters*, 1 :132,133, 1972. 59, 141, 174
- [71] R. A. Jarvis. On the identification of the convex hull of a finite set of points in the plane. *Information Processing Letters*, 2 :18–21, 1973. 59
- [72] W. Eddy. A new convex hull algorithm for planar sets. *ACM Trans. Math. Software*, 3(4) :398–403, 1977. 59

- [73] F. P. Preparata and S. J. Hong. Convex hulls of finite sets of points in two and three dimensions. *Communications of the ACM*, 20(2) :87–93, 1977. 59
- [74] L. Devroye. How to reduce the average complexity of convex hull finding algorithms. *Comp. Math. Appl.*, 7 :299–308, 1981. 59, 141
- [75] D. G. Kirkpatrick and R. Seidel. The ultimate planar convex hull algorithm? *SIAM Journal on Computing*, 15(1) :287–299, 1986. 59
- [76] B. Bhattacharya and S. Sen. On a simple, practical, optimal, output-sensitive randomized planar convex hull algorithm. *Journal of Algorithms*, 25 :173–193, 1997. 59
- [77] R. Wenger. Randomized quick hull. *Algorithmica*, 17 :322–329, 1997. 59
- [78] R. Seidel. Convex hull computations. In J. E. Goodman and J. O’Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 361–375. CRC Press, 1997. 59, 141
- [79] G.T. Toussaint. A historical note on convex hull finding algorithms. *Pattern Recognition Letters*, 3 :21–28, 1985. 59
- [80] F. Yaacoub, Y. Hamam, A. Abche, and C. Fares. Convex hull in medical simulations : a new hybrid approach. In *32nd Annual Conference of IEEE Industrial Electronics Society, IECON’06*, pages 3308–3313, 2006. 59
- [81] Nikolay M. Sirakov. A new active convex hull model for image regions. *J. Math. Imaging Vis.*, 26(3) :309–325, 2006. 59
- [82] B. J. Worton. A convex hull-based estimator of home-range size. *Biometrics*, 51(4) :1206–1215, 1995. 59, 126
- [83] J. Randon-Furling, S. N. Majumdar, and A. Comtet. Convex Hull of n Planar Brownian Motions : Exact results and an application to Ecology. *Phys. Rev. Lett.*, 103 :140602, 2009. 60

- [84] P. Lévy. *Processus stochastiques et mouvement brownien*. Gauthiers-Villars, Paris, 1948. 61, 69, 125, 157
- [85] J. Geffroy. Contribution à la théorie des valeurs extrêmes. II. *Publ. Inst. Stat. Univ. Paris*, 7(8) :3–65, 1959. 62, 120
- [86] J. Geffroy. Localisation asymptotique du polyèdre d'appui d'un échantillon Laplacien à k dimensions. *Publ. Inst. Stat. Univ. Paris*, 10 :213–228, 1961. 62
- [87] F. Spitzer and H. Widom. The circumference of a convex polygon. *Proc. Am. Math. Soc.*, 12 :506–509, 1961. 62, 63, 76
- [88] G. Baxter. A combinatorial lemma for complex numbers. *Ann. Math. Stat.*, 32(3) :901, 1961. 63, 76
- [89] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrsch.*, 2 :75–84, 1963. 64
- [90] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten II. *Z. Wahrsch.*, 3 :138–147, 1964. 64, 76, 113, 114, 115, 124
- [91] H. Raynaud. Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans \mathbb{R}^n . *C. R. Acad. Sciences*, 261 :627–629, 1965. 66, 67
- [92] H. Carnal. Die konvexe Hülle von n rotationssymmetrisch verteilten Punkten. *Z. Wahrsch.*, 15 :168–176, 1970. 67, 69, 76, 118
- [93] H. Raynaud. Sur l'enveloppe convexe des nuages de points aléatoires dans \mathbb{R}^n . *J. Appl. Prob.*, 7(1) :35–48, 1970. 67
- [94] W. Eddy. The distribution of the convex hull of a Gaussian sample. *J. Appl. Prob.*, 17 :686–695, 1980. 67, 68, 69, 70, 95
- [95] B. Brown and S. I. Resnick. Extreme values of independent stochastic processes. *J. Appl. Prob.*, 14(4) :732–739, 1977. 67, 151

- [96] L. Takács. Expected perimeter length. *Amer. Math. Month.*, 87 :142, 1980. 68, 76, 77, 95, 125, 127, 132
- [97] G. Letac. *American Mathematical Monthly*, 1978. 68
- [98] W. Eddy and J.D. Gale. The convex hull of a spherically symmetric sample. *Adv. Appl. Prob.*, 13 :751–763, 1981. 68, 70
- [99] M. El Bachir. *L'enveloppe convexe du mouvement brownien*. PhD thesis, Université Paul Sabatier, Toulouse, France, 1983. 69, 70, 76, 77, 95, 125, 127, 132
- [100] C. Buchta. Zufallspolygone in konvexen Vielecken. *J. reine u. angew. Math.*, 347 :212–220, 1983. 70, 76
- [101] F. Affentranger. The expected volume of a random polytope in a ball. *Journal of Microscopy*, 151 :277–287, 1988. 70, 76
- [102] C. Buchta. Zufällige Polyeder. *Lecture Notes in Mathematics 114*, Springer, 1985. 70
- [103] R. Schneider. Random approximation of convex sets. *Journal of Microscopy*, 151 :211, 1988. 70
- [104] W. Weil and J. A. Wieacker. *Handbook of Convex Geometry*, volume B, pages 1391–1438. North Holland, 1993. 70
- [105] H. Brozius and J. de Haan. On limiting laws for the convex hull of a sample. *J. Appl. Prob.*, 24 :852, 1987. 70, 75
- [106] H. Brozius. Convergence in mean of some characteristics of the convex hull. *Adv. Appl. Prob.*, 21 :526–542, 1989. 70
- [107] R. Davis, E. Mulrow, and S. Resnick. The convex hull of a random sample in \mathbb{R}^d . *Comm. Statist. Stoch. Models*, 3(1) :1–27, 1987. 70
- [108] D. Aldous, B. Fristedt, P. S. Griffin, and W. E. Pruitt. The number of extreme points in the convex hull of a random sample. *J. Appl. Prob.*, 28 :287–304, 1991. 70

- [109] M. Cranston, P. Hsu, and P. March. Smoothness of the convex hull of planar Brownian motion. *Annals Prob*, 17(1) :144, 1989. 70
- [110] M. Shimura. A limit theorem for conditional random walk. *Nagoya Math. J.*, 95 :105–116, 1984. 70
- [111] M. Shimura. Excursions in a cone for two-dimensional Brownian motion. *J. Math. Kyoto Univ.*, 13 :433–443, 1985. 70
- [112] K. Burdzy. Brownian motion in cones. *Ann. Prob.*, 13 :1006–1010, 1985. 70
- [113] J.-F. Le Gall. Mouvement brownien, cônes et processus stables. *Prob. Theory Relat. Fields*, 76 :587–627, 1987. 71
- [114] P. Lévy. Le caractère universel de la courbe du mouvement brownien et la loi du logarithme itéré. *Circ. Mat. Palermo Ser.*, 2.4 :337–366, 1955. 71
- [115] S. N. Evans. On the Hausdorff dimension of Brownian cone points. *Math. Proc. Camb. Philos. Soc.*, 98 :343–353, 1985. 71
- [116] K. Burdzy and J. San Martin. Curvature of the convex hull of planar Brownian motion near its minimum point. *Stoch. Process. Appl.*, 33(1) :89–103, 1989. 71
- [117] D. Khoshnevisan. Moment Inequalities for Functionals of the Brownian Convex Hull. *Annals Prob.*, 20(2) :627, 1992. 71
- [118] G. Letac. An explicit calculation of the mean of the perimeter of the convex hull of a plane random walk. *J. Theor. Prob*, 6(2) :385, 1993. 71, 76
- [119] T. Snyder and J. Steele. Convex hulls of random walks. *Proc. Am. Math. Soc.*, 117(4) :1165, 1993. 71, 76
- [120] A. Goldman. Le spectre de certaines mosaïques poissoniennes du plan et l’enveloppe convexe du pont brownien. *Prob. Theor. Relat. Fields*, 105 :57–83, 1996. 73, 76, 125, 127

- [121] A. Goldman. Sur une conjecture de D.G. Kendall concernant la cellule de Crofton du plan et sur sa contrepartie brownienne. *C.R. Acad. Sc. Paris*, 326 :233–237, 1998. 74, 125
- [122] P. Groeneboom. Limit theorems for convex hulls. *Prob. Theory Relat. Fields*, 79 :327–368, 1988. 75
- [123] S. Finch and I. Hueter. Random convex hulls : a variance revisited. *Adv. Appl. Prob.*, 36(4) :981–986, 1994. 75
- [124] T. Hsing. On the asymptotic distribution of the area outside a random convex hull in a disk. *Ann. Appl. Prob.*, 4 :478–493, 1994. 75
- [125] J. Cabo and P. Groeneboom. Limit theorems for functionals of convex hulls. *Prob. Theory Relat. Fields*, 100 :31–55, 1994. 75
- [126] V. Vu. Central limit theorems for random polytopes in a smooth convex set. *Adv. Math*, 207 :221–243, 2006. 75
- [127] P. Calka and T. Schreiber. Large deviation probabilities for the number of vertices of random polytopes in the ball. *Adv. Appl. Prob.*, 38 :47–58, 2006. 75
- [128] M. Reitzner. Random polytopes and the Efron-Stein jackknife inequality. *Ann. Prob.*, 31 :2136–2166, 2003. 75
- [129] M. Reitzner. The combinatorial structure of random polytopes. *Adv. Math.*, 191 :178–208, 2005. 75
- [130] M. Reitzner. Central limit theorems for random polytopes. *Prob. theory Relat. Fields*, 133 :483–507, 2005. 75
- [131] I. Bárány. *Stochastic Geometry*. C.I.M.E., Lecture Notes Math. 1892. Springer, 2006. 75
- [132] I. Bárány and M. Reitzner. Random polytopes. *Preprint*, 2008. www.renyi.hu/~barany/cikkek/clt-pol2.pdf. 75
- [133] I. Bárány and V. Vu. Central limit theorems for Gaussian polytopes. *Ann. Prob.*, 35(4) :1593–1621, 2007. 75

- [134] P. Biane and G. Letac. The mean perimeter of some random plane convex sets generated by a Brownian motion. *J. Theor. Prob.*, 2010. arXiv :0905.2256. 75, 77
- [135] B. Efron. The convex hull of a random set of points. *Biometrika*, 52(3 and 4) :331, 1965. 76, 108
- [136] H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss. Number of distinct sites visited by n random walkers. *Phys. Rev. A*, 45(10) :7128–7139, 1992. 77, 143
- [137] L. Acedo and S. B. Yuste. Multiparticle random walks. *Recent Res. Dev. Stat. Phys.*, 2 :83–106, 2002. 77, 143, 144, 145
- [138] P. E. Rouse. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *J. Chem. Phys.* 21, 1272 (1953, 21 :1272–1280, 1953. 79
- [139] Steven Finch. Variants of brownian motion, 2004. <http://pauillac.inria.fr/algo/resolve/br.pdf>. 81
- [140] A. Comtet and J. Desbois. Brownian motion in wedges, last passage time and the second arc-sine law. *J. Phys. A : Math. Gen.*, 36(17) :L255–L261, 2003. 84
- [141] H. S. Carslaw and J. C. Jaeger. *Conduction of heat in solids*. Oxford University Press, Oxford, 1959. 86
- [142] J. S. Dowker. Heat kernel expansion on a generalized cone. *J. Math. Phys.*, 30 :770–773, 1989. 86
- [143] A. Cauchy. La rectification des courbes. Mémoire de l'Académie des Sciences, 1832. 96, 158, 177
- [144] M.W. Crofton. On the theory of local probability, applied to straight lines at random in a plane. *Trans. Roy. Soc.*, 158 :181–199, 1868. 96
- [145] E. Barbier. Note sur le problème de l'aiguille et le jeu du joint couvert. *J. Math. Pures Appl.*, 5 :273–286, 1860. 96, 102, 143, 158

- [146] L. Santaló. *Integral Geometry and Geometric Probability*. Encyclopedia of Mathematics and Its Applications. Addison-Wesley, 1976. 96, 101
- [147] F. Valentine. *Convex sets*. McGraw-Hill, 1964. 96, 101
- [148] S. Ayari and S. Dubuc. La formule de Cauchy sur la longueur d'une courbe. *Canad. Math. Bull.*, 40 :3–9, 1997. 96, 101
- [149] E.J. Gumbel. *Statistics of Extremes*. Columbia University Press, New York, 1958. 103, 108
- [150] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer Verlag, London, 2001. 103
- [151] H. E. Daniels and T. H. R. Skyrme. The maximum of a random walk whose mean path has a maximum. *Adv. Appl. Prob.*, 17(1) :85–99, 1985. 104
- [152] A. M. Odlyzko. Search for the maximum of a random walk. *Random Structures Algorithms*, 6(275–295), 1995. 104
- [153] P. Chassaing, J.-F. Marckert, and M. Yor. A stochastically quasi-optimal search algorithm for the maximum of the simple random walk. *Ann. Appl. Prob.*, 13(4) :1264–1295, 2003. 104
- [154] W. Weibull. A statistical distribution function of wide applicability. *J. Appl. Mech.*, 18(3) :293–297, 1951. 108
- [155] H. J. Hilhorst, P. Calka, and G. Schehr. Sylvester's question and the random acceleration process. *J. Stat. Mech. : Theory and Exp.*, 2008. P10010. 115
- [156] D. Murphy and B. Noon. Integrating Scientific Methods with Habitat Conservation Planning : Reserve Design for Northern Spotted Owls. *Ecological Application*, 2 :3–17, 1992. 125
- [157] S.A. Boyle, W.C Lourenço, L.R. da Silva, and A.T. Smith. Home-range estimates vary with sample size and methods. *Folia Primatologica*, 80 :33–42, 2009. 126

- [158] F. Bartumeus, M.G.E. da Luz, G.M. Viswanathan, and J. Catalan. Animal search strategies : A quantitative random-walk analysis. *Ecology*, 2005. 126
- [159] H.C. Berg. *Random Walks in Biology*. Princeton University Press, New York, 1983. 126
- [160] L. Edelstein-Keshet. *Mathematical Models in Biology*. Random House, New York, 1988. 126
- [161] B. Teissier. Volumes des corps convexes, géométrie et algèbre. In *Leçons de mathématiques d'aujourd'hui*, volume 3 of *Le Sel et le Fer*. Cassini, Paris, 2007. 142
- [162] H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss. Territory Covered by N Diffusing Particles. *Nature*, 355 :423–426, 1992. 143
- [163] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6) :1226–1229, 1995. 152
- [164] A. Cavagna, A. Cimorelli, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, R. Santagati, and F. Stefanini. New statistical tools for analyzing the structure of animal groups. *Mathematical Biosciences*, 214 :32–37, 2008. 152
- [165] P. Romanczuk, I. D. Couzin, and L. Schimansky-Geier. Collective motion due to individual escape and pursuit response. *Phys. Rev. Lett.*, 102 :010602, 2009. 152
- [166] B. J. Ford. Brownian movement in clarkia pollen : A reprise of the first observations. *The Microscope*, 40(4) :235–241, 1992. 155
- [167] A. Avogadro. Essai d'une manière de déterminer les masses relatives des molécules élémentaires des corps, et les proportions selon lesquelles elles entrent dans ces combinaisons. *Journal de physique, de chimie et d'histoire naturelle*, 73 :58–76, 1811. 156

- [168] J. Perrin. *Les Atomes*. Félix Alcan, Paris, 1913. 156
- [169] A. Einstein. Zur Theorie der Brownschen Bewegung. *Ann. d. Physik*, 19 :371–381, 1906. 156
- [170] C. C. W. Taylor. *The Atomists : Leucippus and Democritus*. University of Toronto Press, Toronto, 1999. 156
- [171] W. Sutherland. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin. *Phil. Mag.*, 9 :781, 1905. 156
- [172] M. v. Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. d. Phys.*, 21 :756–780, 1906. 156
- [173] M. v. Smoluchowski. Sur le chemin moyen parcouru par les molécules d'un gaz et sur son rapport avec la théorie de la diffusion. *Bull. Int. Acad. Sc. Crac.*, pages 202–213, mars 1906. 156
- [174] M. v. Smoluchowski. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Phys. Zeit.*, 17 :557–571 and 585–599, 1916. 156
- [175] P. Langevin. Sur la théorie du mouvement brownien. *C. R. Ac. Sci. Paris*, 146 :530–532, 1908. 157
- [176] L. Bachelier. La théorie de la spéculation. *Ann. Sci. Éc. Norm. Sup.*, 17 :21–86, 1900. 157
- [177] R. Feynman. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Mod. Phys.* 20, 367 - 387 (1948), 20 :367–387, 1948. 157
- [178] G. E. Uhlenbeck and L. S. Ornstein. On the theory of Brownian motion. *Phys. Rev.*, 36 :823–841, 1930. 157
- [179] M. C. Wang and G. E. Uhlenbeck. On the theory of Brownian motion II. *Rev. Mod. Phys.*, 17 :323–342, 1945. 157

- [180] G. Matheron. *Random convex sets and integral geometry*. Wiley, New York, 1975. 158
- [181] J. J. Sylvester. Problem 1491. *The Educational Times*, April 1864. 158
- [182] E. Seneta, K. H. Parshall, and F. Jongmans. Nineteenth-century developments in geometric probability : J. J. Sylvester, M. W. Crofton, J.-E. Barbier, and J. Bertrand. *Archive for History of Exact Sciences*, 55(6) :501–524, 2001. 158
- [183] G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5 :131–295, 1953–54. 158
- [184] D. G. Kendall. Foundations of a theory of random sets. In E. F. Harding and D. G. Kendall, editors, *Stochastic Geometry*, pages 322–376. Wiley, New York, 1973. 158
- [185] H. D. Fara and A. E. Scheidegger. Statistical geometry of porous media. *J. Geophys. Res.*, 66(10) :3279–3284, 1961. 159
- [186] G. Matheron. *Éléments pour une théorie des milieux poreux*. Masson, Paris, 1967. 159
- [187] R. J. Gardner. Geometric tomography. *Notices of the AMS*, 42(4) :422–429, 1995. 159
- [188] W. Weil. Stereology : a survey for geometers. In P. M. Gruber and J. M. Wills, editors, *Convexity and its applications*, pages 360–421. Birkhäuser, Basel, 1983. 159
- [189] A. Al-Jarrah, K. M. Dempsey, and M. L. Glasser. Generalized series of Bessel functions. *J. Comput. Appl. Math.*, 143(1) :1–8, 2002. 161
- [190] W. T. Vetterling, S. A. Teukolsky, W. H. Press, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992. 173