

Mémoire de magistère MMFAI

Rémi Peyre

12 octobre 2007

Avant-Propos

Le présent mémoire récapitule l'ensemble des travaux que j'ai effectués au cours de ma scolarité au magistère MMFAI de l'École Normale Supérieure. Il est constitué de onze documents essentiellement indépendants dont voici la description :

1. *Curriculum vitæ*. Ce texte récapitule l'ensemble de mon parcours au MMFAI.
2. *Marche aléatoire en milieu aléatoire*. Il s'agit de mon exposé de maîtrise, réalisé en collaboration avec Hendrik Weber, sous la direction de Thierry Bodineau.
3. *Une inégalité de Harnack*. Il s'agit de mon mémoire de DEA, réalisé sous la direction de Cédric Villani.
4. *Ramification des graphes*. Ce document est constitué par les notes d'un exposé d'une heure dans le cadre du groupe de travail « Probabilités » organisé à l'ENS en 2005.
5. *Approche probabiliste du problème d'équilibre*. Ce document est constitué par les notes d'un exposé d'une heure et demie dans le cadre d'un groupe de travail organisé par Yves Le Jan pour le DEA de probabilités d'Orsay en 2006.
6. *Autour du mouvement brownien*. Ce document est constitué par les notes d'une série de trois exposés de deux heures dans le cadre d'un groupe de travail étudiant à l'ENS Lyon en 2007. Il s'agit d'exposés introductifs destinés à des non-spécialistes.
7. *A probabilistic approach to Carne's bound*. C'est un article de recherche qui, après quelques dernières corrections, sera soumis à la revue *Potential Analysis* imminemment.
8. *Un nouveau type de décorrelations exponentielles pour les mesures de Gibbs*. Ce texte est constitué de notes personnelles développant quelques résultats nouveaux sur les mesures de Gibbs à haute température ; il contient sans doute encore de nombreuses coquilles.
9. *Moyennes déviations des mesures empiriques*. Ce texte est constitué de notes personnelles développant quelques résultats nouveaux sur la convergence des mesures empiriques ; il contient sans doute encore de nombreuses coquilles. Il sert de travail préparatoire au document suivant.
10. *Moyennes déviations pour des modèles de champs moyen à chocs*. Ce texte est constitué de notes personnelles développant quelques résultats nouveaux sur les limites de champ moyen ; il contient sans doute encore de nombreuses coquilles.

11. *Grand nombre de particules en interaction : la loi de Fourier et l'équation de Boltzmann en champ moyen.* Ce texte est la présentation de mon domaine de recherches : j'y aborde les thèmes d'intérêt de ma thèse, les questions que j'y ai résolues, les problèmes qui restent à surmonter et les méthodes que je compte employer à cette fin.

Table des matières

I Curriculum vitæ	6
II Marche aléatoire en milieu aléatoire	12
III Une inégalité de Harnack	52
IV Ramification des graphes	82
V Approche probabiliste du problème d'équilibre	94
VI Autour du mouvement brownien	118
VII A probabilistic approach to Carne's bound	156
VIII Un nouveau type de décorrelations exponentielles pour les mesures de Gibbs	184
IX Moyennes déviations des mesures empiriques	196
X Moyennes déviations pour des modèles de champs moyen à chocs	206
XI Grand nombre de particules en interaction : la loi de Fourier et l'équation de Boltzmann en champ moyen	222

Document I
Curriculum vitæ

Curriculum vitæ

Je suis entré à l'ÉNS en 2004 par le concours PC, hésitant entre continuer la physique ou me consacrer au mathématiques. Les bons contacts que j'ai eu avec le DMA, tant au niveau des élèves que des professeurs, m'ont rapidement amené à opter pour les mathématiques, et j'ai commencé le magistère MMFAI, cursus mathématiques — car, dans la mesure où le cursus mixte maths-physique servait essentiellement « à différer l'heure du choix », je n'y voyais pas d'intérêt, le mien étant définitif.

À l'origine, je comptais, imitant la plupart de mes camarades, passer ma licence et ma maîtrise en première année, puis continuer par l'agrégation tout en commençant un DEA. N'ayant pas encore d'*a priori* sur la branche des mathématiques que je voulais étudier, je me suis inscrit à tous les cours de mathématiques du premier semestre, à savoir *Algèbre I* (D. Harari), *Topologie & Calcul différentiel* (B. Perthame), *Intégration & Probabilités* (G. David) et *Logique* (P. Dehornoy), ainsi qu'au groupe de lecture « Graphes de Ramanujan ». Mais j'ai dû quitter Paris peu avant la fin du semestre à cause d'un arrêt de travail (dépression), ce qui fait que je n'ai pas passé les examens du premier semestre, et manqué le début du second.

À mon retour, il s'agissait de finir l'année le moins mal possible. À cause du retard pris, j'ai choisi de suivre les cours aux contenus les moins denses, à savoir *Processus aléatoires* (J-F. Le Gall), *Analyse fonctionnelle* (É. Séré) et *Analyse complexe & Théorie spectrale* (G. Skandalis). Entretemps, il m'avait fallu choisir un sujet de mémoire de maîtrise : je me suis associé avec H. Weber pour travailler sur le sujet « Marche aléatoire en milieu aléatoire » sous la direction de T. Bodineau, car le thème me plaisait par son aspect à la fois naturel, divertissant et riche en intuitions physiques.

C'est ainsi que j'ai validé ma licence avec la note de 17 (mention TB), en validant au passage les cours de G. David et P. Dehornoy à la session de juin, étant convenu avec la direction du MMFAI que, des cours étiquetés « maîtrise » ayant été pris en compte dans ma licence, l'inverse serait fait l'année suivante.

N'ayant donc pas obtenu de maîtrise, je n'avais pas la possibilité de passer l'agrégation et je me suis orienté vers un DEA, avec une inscription pédagogique pour contourner les problèmes administratifs. Comme le sujet de l'exposé de maîtrise m'avait particulièrement plu, j'avais résolu d'étudier les

probabilités, et, suite à la présentation des DEA de fin d'année, j'ai opté pour le cursus « Probabilités & Statistiques » de Paris XI (Orsay) qui me seyait particulièrement. Là, j'ai suivi au premier semestre les cours de W. Werner (*Percolation & Modèle d'Ising*), O. Raimond (*Calcul Stochastique*), T. Duquesne (*Processus Markoviens*) et P. Massart (*Grandes déviations & Inégalités de concentration*). En même temps, je participais au groupe de travail « Probabilités sur les arbres et les réseaux » à l'ÉNS, et j'y assistais également au cours de deuxième année *Systèmes dynamiques* (F. Le Roux), ainsi aux TDs des cours d'algèbre et de topologie en vue de préparer les examens. J'ai passé les sept examens à la fin du semestre.

Au second semestre, le DEA d'Orsay, trop petit pour dispenser assez de cours, autorise ses élèves à valider les cours d'autres universités. En l'occurrence, j'allai à Paris VII et y suivis les cours qui m'intéressaient le plus par le côté « naturel » et « pratique » des questions qu'ils étudiaient, à savoir *Ondes en milieu aléatoire* (J. Garnier), *Arbres et cartes aléatoires* (Ph. Chassaing) et *Limite hydrodynamique* (T. Bodineau). J'ai suivi les trois cours, mais n'ai validé que les deux premiers car il n'était pas nécessaire d'en faire plus. Il me fallait également préparer un mémoire de DEA : dans un premier temps, j'ai contacté R. Cerf à Orsay, qui souhaitait me faire travailler sur les principes de grandes déviations dans les modèles de physique statistique, mais le style très effacé de R. Cerf me gênait pour bien cerner le sujet et j'ai préféré renoncer. Je me suis alors tourné vers Cédric Villani (ÉNS Lyon) qui a accepté d'encadrer mon mémoire à distance en me donnant un sujet sur les diffusions dégénérées. En fait, le sujet a bifurqué vers une méthode probabiliste de régularité pour les solutions stationnaires de diffusions avec un champ non-gradient, que j'ai soutenu en octobre 2006. Pour tous ces choix de cours, directeurs de DEA, changement d'orientation, j'ai été grandement aidé par mes entretiens avec W. Werner, comme directeur de la formation d'Orsay dans un premier temps, puis simplement comme conseiller. Par ailleurs, toujours au second semestre de l'année 2005-2006, je suivais à l'ÉNS les cours de maîtrise que je n'avais pas fait l'année précédente, soit *Algèbre II* (M. Rosso) et *Variétés Différentielles* (F. Paulin). J'ai ainsi validé ma maîtrise (18, mention TB), et mon DEA *de facto* (mention TB), à la fin de ma deuxième année de la MMFAI.

Le fin de la deuxième année fut une période de réflexions, de doutes aussi, sur la suite à donner à mon parcours, et une fois encore les conseils de W. Werner m'ont bien aidé à y voir plus clair. La préparation de l'agrégation, initialement prévue, m'apparaissait maintenant comme rébarbative et, dans la mesure où j'étais certain de ne pas vouloir m'orienter vers l'enseignement en lycée/CPGE, j'ai finalement décidé de renoncer à passer l'agrégation pour me lancer sur une thèse. Dans la mesure où mes contacts avec C. Villani au cours de mon mémoire de DEA avaient été bons, et où j'appréciais sa façon de présenter les problèmes, j'ai commencé une thèse avec lui, déménageant à Lyon pour la circonstance.

Le sujet initial était la façon dont la conduction de la chaleur apparaît

à l'échelle macroscopique à partir des lois microscopiques du mouvement. Il s'agit d'un sujet vaste et difficile, à cheval sur les travaux des physiciens et des mathématiciens, et j'ai passé beaucoup de temps à en appréhender toutes les facettes. À cela s'ajoutant le changement de rythme, l'année a au final été très peu efficace et s'est même terminée sur un épisode de psychasthénie et un nouvel arrêt de travail. Il y a tout de même eu, de manière complètement indépendante, un travail de recherche sur les marches aléatoires réversibles, aujourd'hui sur le point d'être soumis à publication.

À mon retour, nous avons réorienté la thèse vers des problèmes de limite de champ moyen pour les modèles de particules à chocs. Dans un travail avec un précédent thésard (F. Bolley), mon directeur s'était en effet intéressé à l'obtention d'estimées non asymptotiques pour les limites de champ moyen ; malheureusement leur méthode ne supportait pas les évolutions par chocs, pourtant essentielles dans l'étude de l'équation de Boltzmann. J'ai rapidement eu quelques idées qui semblent prometteuses, et mon travail devrait se recentrer sur ces questions.

À l'heure actuelle donc, je suis enfin inscrit en thèse de manière officielle à l'ÉNS Lyon et continue mon travail sur les limites de champ moyen quantitatives sous la direction de C. Villani.

Document II

Marche aléatoire en milieu aléatoire

Marche aléatoire en milieu aléatoire

Rémi Peyre – Hendrik Weber

sous la direction de Thierry Bodineau

17 juin 2005

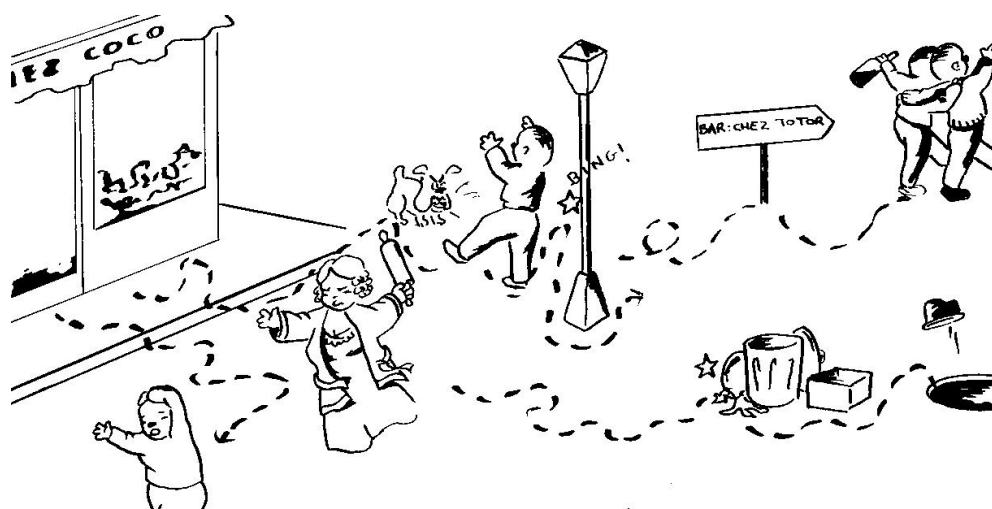


Table des matières

1	Généralités	16
1.1	Présentation	16
1.2	Cadre de notre étude	17
1.3	Notion de système ergodique	18
2	Récurrence et transience	22
2.1	Énoncé du théorème	22
2.2	Démonstration	22
3	Vitesse de dérive	26
3.1	Énoncé du théorème	26
3.2	Démonstration	27
4	Chaînes auxiliaires	32
4.1	Chaînes auxiliaires	32
4.2	Preuve alternative du théorème 8	38
5	Comparaison avec les marches aléatoires simples	40
5.1	Différence conceptuelle	40
5.2	Comparaison des théorèmes	42
5.3	Une interprétation : notion de puits	42
5.3.1	Profil de potentiel	42
5.3.2	Effet des puits	43
5.3.3	La mesure \bar{Q}	44
5.3.4	Cas transitoire à vitesse nulle	44
5.3.5	Approche heuristique du théorème 8	45
5.4	Exemples numériques	46
5.4.1	Vitesse dans un milieu simple	46
5.4.2	Le cas à vitesse nulle	47
5.4.3	Le cas récurrent	48

Chapitre 1

Généralités

1.1 Présentation

Qu'est-ce qu'une marche aléatoire ? Physiquement, cela correspond au déplacement aléatoire d'un mobile dans un certain milieu, par exemple un grain de pollen dans une goutte d'eau. À un instant donné, le mobile a une certaine probabilité de se déplacer dans chaque direction, et on s'intéresse aux propriétés de son mouvement global. Parfois, le déplacement du mobile entre deux instants dépend du point où il se trouve : ainsi, au cours d'un processus de diffusion, la diffusion d'un électron dans un cristal semiconducteur par exemple, l'électron a plus de chances de rester immobile entre deux instants lorsqu'il se trouve près d'un défaut électrophile du cristal que dans le cas contraire. On a alors une marche aléatoire dans un milieu irrégulier. Mais pour le physicien, il est absurde de chercher à décrire parfaitement le milieu : imagine-t-on ce que donnerait la description complète du café contenu dans un percolateur pour y étudier le mouvement des gouttes d'eau ! En fait, on se contente de donner les caractéristiques statistiques du milieu, et tout se passe comme si le milieu était choisi au hasard parmi un ensemble de milieux possédant les mêmes propriétés statistiques. Ce qui nous amène tout naturellement au concept de marche aléatoire en milieu aléatoire (MAMA).

Pour le mathématicien, une MAMA se construit ainsi : on choisit d'abord un cadre général pour la marche, un graphe qui représente l'espace dans lequel le mobile va évoluer (pour modéliser l'évolution d'un mobile dans l'espace à trois dimensions, par exemple, ce réseau pourra être \mathbb{Z}^3 , où les arêtes seront les (x, y) tels que $\|x - y\|_1 \leq 1$). Ensuite, on va définir un *milieu* en donnant, pour chaque sommet du graphe, une certaine loi définissant la probabilité de passer sur chacun des sommets qui lui sont connectés. Mais ce milieu n'est pas fixé une fois pour toutes : en fait, il obéit à une certaine distribution aléatoire P que le mathématicien choisit. Lors du déclenchement d'une MAMA, on commence par tirer au sort le milieu conformément à la mesure P , puis on laisse évoluer le mobile dans le milieu ainsi fixé¹, selon une chaîne de Markov.

¹Attention ! Une fois que le milieu a été choisi, il reste le même tout au long du déplacement du mobile. Si on retirait au sort le milieu à chaque étape, on se retrouverait avec une

1.2 Cadre de notre étude

On peut définir une MAMA de manière très générale, de façon à se rapprocher le plus possible des cas rencontrés en physique. Mais le traitement mathématique des marches devient alors difficile. Citons deux exemples :

- Quand on raisonne à deux ou trois dimensions, le comportement du mobile à l'infini est très compliqué à prévoir à cause de la possibilité qu'il a de contourner certains points ou groupes de points qui auraient tendance à lui imposer un certain comportement (par exemple, à l'immobiliser). Il existe ainsi des cas où le mobile a une probabilité non nulle de partir à l'infini dans deux directions opposées.
- Un autre problème peut être dû au couplage entre points voisins : on peut, par exemple, imaginer un milieu à une dimension où les points qui ont tendance à envoyer le mobile vers la gauche ont une probabilité plus grande que les autres d'être entourés de points ayant tendance à envoyer le mobile vers la droite. Il est alors possible d'utiliser la théorie des systèmes ergodiques pour traiter les cas où le couplage est nul à l'infini, mais c'est au prix d'une certaine complexité mathématique.

Finalement, nous nous restreindrons au cas particulier suivant :

Le mobile se déplace dans le graphe \mathbb{Z} , où chaque point est connecté à ses deux voisins ainsi qu'à lui-même. Un milieu ω est défini en donnant pour chaque point z des probabilités respectives ω_z^- , ω_z^0 et ω_z^+ d'envoyer le mobile à gauche, de le laisser immobile ou de l'envoyer à droite. On doit évidemment avoir $\omega_z^- + \omega_z^0 + \omega_z^+ = 1$, et le triplet $(\omega_z^-, \omega_z^0, \omega_z^+)$ est appelé *loi de transition en z* . On utilisera aussi la notation $\omega_z(y)$ pour désigner la probabilité qu'a le mobile de sauter de z vers y^2 . Si on note $\mathcal{M}_1(N_z)$ l'ensemble des lois de transition possibles en z , un milieu peut donc être considéré comme un élément de $\Omega \stackrel{\text{déf}}{=} \prod_{z \in \mathbb{Z}} \mathcal{M}_1(N_z)$.

La topologie vague sur $\mathcal{M}_1(N_z)$ est la même qu'en tant que sous-ensemble de \mathbb{R}^3 . On munit Ω de la topologie produit, c'est-à-dire la topologie dont les ensembles de la forme

$$\prod_{z < z^-} \mathcal{M}_1(N_z) \times \prod_{z = z^-}^{z^+} \mathcal{O}_z \times \prod_{z > z^+} \mathcal{M}_1(N_z)$$

avec \mathcal{O}_z ouvert de $\mathcal{M}_1(N_z)$ forment une base d'ouverts. Comme chaque $\mathcal{M}_1(N_z)$ est compact Ω l'est aussi. En particulier Ω est un espace polonais. On munit Ω de sa tribu borélienne \mathcal{F} .

Reste à définir une probabilité P sur Ω pour représenter la façon dont le milieu est tiré au sort. Dans tout cet exposé, nous ne considérons sur Ω que des probabilités telles que

marche aléatoire simple, cf remarque 10.

²Ce qui donne $\omega(z, y) = \omega_z^-$, resp. ω_z^0 , resp. ω_z^+ si $y = z - 1$, , resp. $y = z$, resp. $y = z + 1$, et $\omega(z, y) = 0$ sinon.

Hypothèse 1. Les ω_z sont indépendants et de même loi.

La loi de P est alors entièrement déterminée par la loi de ω_0 .

Pour tout milieu $\omega \in \Omega$, on appelle *marche aléatoire dans le milieu ω* la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ de matrice de transition $Q(z, y) = \omega_z(y)$. Lorsqu'on prend la chaîne de Markov partant du point v , la loi de la marche est notée P_ω^v : il s'agit d'une mesure de probabilité sur $(\mathbb{Z}^{\mathbb{N}}, \mathcal{G})$, où \mathcal{G} est la plus petite tribu sur $\mathbb{Z}^{\mathbb{N}}$ rendant mesurables les fonctions coordonnées. On a donc $P_\omega^v(X_0 = v) = 1$ et $P_\omega^v(X_{n+1} = y | X_n = z) = \omega_z(y)$. On dit que $P_\omega^v(\cdot)$ est la loi *particulière* de la marche aléatoire $(X_n)_{n \in \mathbb{N}}$, car elle fait référence à un milieu particulier. Lorsqu'on tient compte des variations possibles du milieu, on introduit la loi sur $(\Omega \times \mathbb{Z}^{\mathbb{N}}, \mathcal{F} \times \mathcal{G})$ notée $\mathbb{P}^v(\cdot)$ et définie par $\mathbb{P}^v = P \otimes P_\omega^v$, càd. pour $F \in \mathcal{F}$, $G \in \mathcal{G}$:

$$\mathbb{P}^v(F \times G) = \int_F P_\omega^v(G) P(d\omega)$$

(noter que $\omega \rightarrow P_\omega^v(G)$ est mesurable).

La loi *générale* de la marche aléatoire est alors la loi marginale de cette loi sur $\mathbb{Z}^{\mathbb{N}}$. On la note également \mathbb{P}^v :

$$\mathbb{P}^v(G) \stackrel{\text{déf}}{=} \mathbb{P}^v(\Omega \times G) = \int_\Omega P_\omega^v(G) P(d\omega)$$

Cette loi exprime ce qu'il arrive à la marche aléatoire quand on ne sait pas à quel milieu on a affaire.

Remarque 1. Dire qu'une propriété de la marche est vraie \mathbb{P}^v -p.s., c'est donc dire qu'elle est vraie P_ω^v -p.s., pour P -presque tout milieu ω .

Notre travail va consister à trouver certaines propriétés de la MAMA relativement à la loi \mathbb{P}^0 , en fonction de la probabilité P sur Ω .

1.3 Notion de système ergodique

Nous aurons à plusieurs reprises besoin d'utiliser la théorie des systèmes ergodiques, en particulier le théorème ergodique de Birkhoff.

Définition 1. Soit (Ω, \mathcal{T}, P) un espace de probabilité, et $\varphi : \Omega \rightarrow \Omega$ mesurable. On dit que φ *préserve la mesure* si, pour tout A dans \mathcal{T} , on a $P(\varphi^{-1}A) = P(A)$. On dit qu'un événement A est *invariant* pour φ si $\varphi^{-1}A = A$ P -p.p.. On dit enfin que le système $(\Omega, \mathcal{T}, P, \varphi)$ est *ergodique* si φ préserve la mesure et si tout événement φ -invariant a pour probabilité 0 ou 1.

Proposition 1. Pour $\varphi : \Omega \rightarrow \Omega$ mesurable, l'ensemble \mathcal{I} des événements φ -invariants est une sous-tribu de \mathcal{T} .

Démonstration. La démonstration à partir des définitions d'une tribu ne soulève aucune difficulté \square

On a alors le théorème classique suivant, que nous rappelons sans démonstration :

Théorème 2 (Birkhoff). *Soit Ω un espace de probabilité, $\varphi : \Omega \rightarrow \Omega$ préservant la mesure, et $X \in L^1(\Omega)$. Alors*

$$\frac{1}{n} \sum_{m=0}^{n-1} X(\varphi^m(\omega)) \xrightarrow{n \rightarrow \infty} E[X|\mathcal{I}](\omega), \text{ p.s. et dans } L^1.$$

En particulier, si φ est ergodique, la tribu \mathcal{I} est grossière, et donc $E[X|\mathcal{I}]$ est constante, égale à $E[X]$.

Définition 2. Soit (Ω, \mathcal{T}, P) un espace de probabilité, et $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs dans un espace mesurable (E, \mathcal{U}) . $X = (X_n)_{n \in \mathbb{N}}$ est alors une variable aléatoire sur Ω , à valeurs dans l'espace mesurable $(E^{\mathbb{N}}, \mathcal{U}')$, où \mathcal{U}' est la tribu produit. Notons P' sa loi.

Soit θ l'opérateur de décalage sur $E^{\mathbb{N}}$, c.à.d. l'application $(Y_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}} \mapsto (Y_{n+1})_{n \in \mathbb{N}}$. On dit que la suite $(X_n)_{n \in \mathbb{N}}$ est *stationnaire* si θ préserve la mesure P' , et *ergodique* si en outre, le système $(E^{\mathbb{N}}, \mathcal{U}', P', \theta)$ est ergodique.

Un premier intérêt de la notion d'ergodicité dans notre étude est la propriété suivante :

Théorème 3. Notant $\theta : \begin{cases} \Omega \rightarrow \Omega \\ (\omega_z)_{z \in \mathbb{Z}} \mapsto (\omega_{z+1})_{z \in \mathbb{Z}} \end{cases}$ l'opérateur de décalage sur les milieux, $(\Omega, \mathcal{F}, P, \theta)$ est un système ergodique.

Remarque 2. Par symétrie, $(\Omega, \mathcal{F}, P, \theta^{-1})$ est aussi un système ergodique.

Démonstration. D'abord, θ préserve la mesure. En effet, on vérifie immédiatement que $\{A \in \mathcal{F}; P(\theta^{-1}A) = P(A)\}$ est une tribu, et cette tribu contient tous les ensembles de la forme $\{\omega; \omega_z \in A\}$ pour $z \in \mathbb{Z}$, $A \subset M_1(N_z)$ mesurable : en effet, $\theta^{-1}(\{\omega; \omega_z \in A\}) = \{\omega; \omega_{z+1} \in A\}$ et les ω_z sont indépendants et de même loi. Comme ces ensembles engendrent \mathcal{F} par définition, P préserve la mesure.

Passons maintenant à l'ergodicité proprement dite. En fait, on va montrer que θ est un opérateur *mélangeant*, i.e. :

Lemme 4. Soient $A, B \in \mathcal{F}$ des événements mesurables de Ω . Alors

$$P(A \cap \theta^{-j}B) \xrightarrow{j \rightarrow \infty} P(A)P(B)$$

Démonstration. Commençons par le cas où A et B sont des *cylindres*, c.à.d. des événements de la forme $\{\omega; (\omega_{z_1}, \omega_{z_2}, \dots, \omega_{z_n}) \in Y\}$, où Y est une partie mesurable de $\prod_{i=1}^n M_1(N_{z_i})$. Écrivons ainsi

$$\begin{aligned} A &= \{\omega; (\omega_{z_1}, \omega_{z_2}, \dots, \omega_{z_n}) \in Y\} \\ B &= \{\omega; (\omega_{z'_1}, \omega_{z'_2}, \dots, \omega_{z'_m}) \in Z\} \end{aligned}$$

avec $z_1 < z_2 < \dots < z_n$, resp. $z'_1 < \dots < z'_m$. Il est clair que l'événement $\theta^{-j}B$ est $\{\omega; (\omega_{z'_1+j}, \omega_{z'_2+j}, \dots, \omega_{z'_m+j}) \in Z\}$. Pour $j > z_n - z'_1$, $\theta^{-j}B$ est donc dans

$\sigma(\omega_{z'_1+j}, \dots, \omega_{z'_m+j})$, tandis que A est dans $\sigma(\omega_{z_1}, \dots, \omega_{z_n})$. Comme $\{z'_1 + j, z'_2 + j, \dots, z'_m + j\}$ et $\{z_1, z_2, \dots, z_n\}$ sont disjoints et que les ω_z sont indépendants, ces tribus sont indépendantes, et donc

$$P(A \cap \theta^{-j}B) = P(A)P(\theta^{-j}B) = P(A)P(B)$$

puisque θ préserve la mesure. D'où le résultat pour les cylindres.

Nous allons généraliser le résultat à A et B quelconques, *via* le lemme suivant :

Lemme 5. *Soit $A \in \mathcal{F}$. Alors pour tout $\varepsilon > 0$ il existe un cylindre C tel que $P(A \Delta C) \leq \varepsilon$, où Δ note la différence symétrique.*

Remarque 3. Ce lemme exprime que tout ensemble mesurable de Ω peut être approché d'aussi près qu'on veut par un cylindre.

Démonstration. Notons \mathcal{C} l'ensemble des cylindres. Le résultat est trivial pour $A \in \mathcal{C}$. Comme les cylindres engendrent \mathcal{F} , il suffit de montrer que $\{A \in \mathcal{F}; (\exists C \in \mathcal{C})(P(A \Delta C) \leq \varepsilon)\}$ est une tribu. Or c'est bien le cas, en effet :

- \emptyset est un cylindre, donc on peut l'approcher d'aussi près qu'on veut par lui-même.
- Si A est approché à ε près par le cylindre C , alors A^c est approché à ε près par C^c , puisque $A^c \Delta C^c = A \Delta C$. Comme C^c est aussi un cylindre, l'ensemble auquel on s'intéresse est stable par passage au complémentaire.
- Si A est approché à $\varepsilon/2$ près par C , et si B est approché à $\varepsilon/2$ près par C' , alors $A \cup B$ est approché à ε près par $C \cup C'$ (qui est bien un cylindre) : en effet, $(A \cup B) \Delta (C \cup C') \subset (A \Delta C) \cup (B \Delta C')$.
- Enfin, supposons $A_1 \subset A_2 \subset A_3 \subset \dots$, où les A_i peuvent être approchés d'aussi près qu'on veut par des cylindres, et montrons que $A = \bigcup_{i \geq 0} A_i$ est dans le même cas. Soit $\varepsilon > 0$. Puisque $P(A_i) \nearrow_{i \rightarrow \infty} P(A)$, on peut trouver un i_0 tel que $P(A) - P(A_{i_0}) = P(A \setminus A_{i_0}) = P(A \Delta A_{i_0}) \leq \varepsilon/2$. Soit alors C un cylindre approchant A_{i_0} à $\varepsilon/2$ près : on a $A \Delta C \subset (A \Delta A_{i_0}) \cup (A_{i_0} \Delta C)$, d'où $P(A \Delta C) \leq \varepsilon$, ce qu'on voulait

□

fin du lemme 5

Montrons maintenant la propriété de mélange pour A et B quelconques : soient $A, B \in \mathcal{F}$, et soit $\varepsilon > 0$. Soient $A', B' \in \mathcal{C}$ avec $P(A \Delta A') \leq \varepsilon$, resp. $P(B \Delta B') \leq \varepsilon$. On a

$$P(\theta^{-n}B \Delta \theta^{-n}B') = P(\theta^{-n}(B \Delta B')) \leq \varepsilon$$

puisque θ préserve la mesure, d'où

$$P((A \cap \theta^{-n}B) \Delta (A' \cap \theta^{-n}B')) \leq 2\varepsilon$$

et donc

$$|P(A \cap \theta^{-n}B) - P(A' \cap \theta^{-n}B')| \leq 2\varepsilon$$

Or, comme A' et $\theta^{-n}B'$ sont des cylindres,

$$P(A' \cap \theta^{-n}B') \xrightarrow{n \rightarrow \infty} P(A')P(B')$$

et comme

$$|P(A')P(B') - P(A)P(B)| \leq 2\varepsilon + \varepsilon^2$$

on a finalement

$$\limsup_{n \rightarrow \infty} |P(A \cap \theta^{-n}B) - P(A)P(B)| \leq 4\varepsilon + \varepsilon^2$$

ε pouvant être pris aussi petit qu'on veut, le lemme 4 s'ensuit □

fin du lemme 4

Pour finir, la propriété d'ergodicité découle facilement de la propriété de mélange : soit $B \in \mathcal{F}$ avec $\theta^{-1}B = B$. Alors pour tout $A \in \mathcal{F}$,

$$(\forall n \in \mathbb{N})(A \cap \theta^{-n}B = A \cap B)$$

et donc

$$P(A \cap \theta^{-n}B) \xrightarrow{n \rightarrow \infty} \begin{cases} P(A \cap B) \\ P(A)P(B) \end{cases} \quad \text{par la propriété de mélange de } \theta$$

d'où $P(A \cap B) = P(A)P(B)$: B est indépendant de tous les éléments de \mathcal{F} , donc en particulier de lui-même : il est trivial □

Chapitre 2

Réurrence et transience

2.1 Énoncé du théorème

Le but de ce chapitre est de déterminer dans quelle direction part la MAMA à l'infini. Pour éviter d'avoir à traiter le cas de points qui arrêtent la marche, on fait l'hypothèse suivante :

Hypothèse 2. $P(\omega_0^0 = 1) = 0$.

Cette hypothèse signifie que, P -p.s., aucun point ne fixe la marche aléatoire sur lui : en effet, c'est ce que cela signifie *stricto sensu* pour le point 0, mais comme les probabilités de transition ont la même loi en tous les points c'est vrai pour tous les points, et par réunion dénombrable pour tous les points à la fois.

Nous pouvons maintenant énoncer le

Théorème 6. Notons, pour $z \in \mathbb{Z}$, $\rho_z = \frac{\omega_z^-}{\omega_z^+}$, et supposons qu'on puisse définir $E_P(\ln \rho_0)$ dans $\overline{\mathbb{R}}$. Alors :

- Si $E_P(\ln \rho_0) < 0$ alors $X_n \rightarrow_{n \rightarrow \infty} +\infty$ \mathbb{P}^0 -p.s.,
- Si $E_P(\ln \rho_0) > 0$ alors $X_n \rightarrow_{n \rightarrow \infty} -\infty$ \mathbb{P}^0 -p.s.,
- Si $E_P(\ln \rho_0) = 0$ alors $(\limsup_{n \rightarrow \infty} X_n = +\infty$ et $\liminf_{n \rightarrow \infty} X_n = -\infty)$ \mathbb{P}^0 -p.s..

2.2 Démonstration

Démonstration. On commence par traiter le cas le plus général où $\omega_0^+ > 0$ et $\omega_0^- > 0$, P -p.s. : cela signifie que le point 0 est toujours susceptible de renvoyer le mobile aussi bien à gauche qu'à droite, et *via* l'hypothèse 1, c'est vrai aussi pour tous les autres points. Pour les manipulations que nous comptons faire, il est utile de remarquer qu'alors tous les ρ_z sont dans $]0, +\infty[$, P -p.s..

Commençons par raisonner dans un environnement ω fixé (le déplacement du mobile est alors une chaîne de Markov). Soient $m_- < m_+$ et $z \in \{m_-, m_- +$

$1, \dots, m_+$. Nous voulons d'abord calculer la probabilité, notée $\nu_\omega(m_-, z, m_+)$, qu'un mobile, partant de z et évoluant dans le milieu ω , touche le point m_- avant de toucher le point m_+ , càd.

$$\nu_\omega(m_-, m_+, z) \stackrel{\text{déf}}{=} P_\omega^z(\inf\{n \in \mathbb{N}; X_n = m_-\} < \inf\{n \in \mathbb{N}; X_n = m_+\})^1$$

Notons que la chaîne de Markov est irréductible, d'après le cas où nous nous sommes placés, donc le mobile est condamné à quitter presque sûrement l'intervalle $]m_-, m_+[$.

À cause de la propriété de Markov, la fonction $\nu_\omega(m_-, \cdot, m_+)$ est harmonique : on a

$$(\omega_z^+ + \omega_z^-)\nu_\omega(m_-, z, m_+) = \omega_z^- \nu_\omega(m_-, z-1, m_+) + \omega_z^+ \nu_\omega(m_-, z+1, m_+) \quad (2.1)$$

Or on connaît les conditions aux limites de cette relation : on a

$$\begin{cases} \nu_\omega(m_-, m_-, m_+) = 1 \\ \nu_\omega(m_-, m_+, m_+) = 0 \end{cases}$$

(2.1) est une équation de Laplace discrète, dont on connaît les conditions aux limites, sa solution est donc unique. En effectuant des combinaisons linéaires judicieuses, on trouve :

$$\nu_\omega(m_-, z, m_+) = \frac{\rho_z^{1/2} \sum_{i=z+1}^{m_+} \prod_{j=z+1}^{i-1} \rho_j}{\rho_z^{1/2} \sum_{i=z+1}^{m_+} \prod_{j=z+1}^{i-1} \rho_j + \rho_z^{-1/2} \sum_{i=m_-}^{z-1} \prod_{j=i+1}^{z-1} \rho_j^{-1}} \quad (2.2)$$

À partir de là, on va pouvoir obtenir les conclusions du théorème 6 à partir de conditions relativement simples. Soient

$$S(\omega) \stackrel{\text{déf}}{=} \sum_{i=1}^{\infty} \prod_{j=1}^i \rho_j$$

et

$$F(\omega) \stackrel{\text{déf}}{=} \sum_{i=1}^{\infty} \prod_{j=-i}^{-1} \rho_j^{-1}$$

Notons $\mathcal{S}_+ \stackrel{\text{déf}}{=} \{S(\omega) < \infty\}$ et $\mathcal{F}_+ \stackrel{\text{déf}}{=} \{F(\omega) < \infty\}$, et aussi \mathcal{S}_- et \mathcal{F}_- les complémentaires respectifs de ces deux événements. La formule (2.2) donne immédiatement :

1. Si $\omega \in \mathcal{S}_+ \cap \mathcal{F}_-$: $\lim_{m_- \rightarrow -\infty} \lim_{m_+ \rightarrow +\infty} \nu_\omega(m_-, 0, m_+) = 0$;
2. Si $\omega \in \mathcal{S}_- \cap \mathcal{F}_+$: $\lim_{m_+ \rightarrow +\infty} \lim_{m_- \rightarrow -\infty} \nu_\omega(m_-, 0, m_+) = 1$;
3. Si $\omega \in \mathcal{S}_- \cap \mathcal{F}_-$: $(\forall k \in \mathbb{N})(\lim_{m_+ \rightarrow +\infty} \nu_\omega(-k, 0, m_+) = 1)$ et $(\forall k \in \mathbb{N})(\lim_{m_- \rightarrow -\infty} \nu_\omega(m_-, 0, k) = 0)$.

¹L'infimum de l'ensemble vide est pris égal à $+\infty$.

Dans le cas 3, on voit immédiatement que la marche passe P_ω^0 -p.s. par tous les points, donc $\liminf_{n \rightarrow \infty} X_n = -\infty$ et $\limsup_{n \rightarrow \infty} X_n = +\infty$. Il est un peu plus subtil de trouver le comportement de la marche dans les cas 1 et 2 :

Lemme 7. 1. Si $\lim_{m_- \rightarrow -\infty} \lim_{m_+ \rightarrow +\infty} \nu_\omega(m_-, 0, m_+) = 0$, alors $X_n \xrightarrow{n \rightarrow \infty} +\infty$ P_ω^0 -p.s.

2. Si $\lim_{m_+ \rightarrow +\infty} \lim_{m_- \rightarrow -\infty} \nu_\omega(m_-, 0, m_+) = 1$, alors $X_n \xrightarrow{n \rightarrow \infty} -\infty$ P_ω^0 -p.s.

Démonstration. Les deux cas étant symétriques, nous ne traiterons que le cas 1.

Soit $\varepsilon > 0$. Soit m_-^0 tel que $\lim_{m_+ \rightarrow +\infty} \nu_\omega(m_-^0, 0, m_+) \leq \varepsilon$. Cela signifie qu'il y a une P_ω^0 -probabilité d'au moins $1 - \varepsilon$ pour que $(X_n)_{n \in \mathbb{N}}$ soit minorée par m_-^0 et ait $+\infty$ pour limite supérieure. Or, dans un tel cas, $X_n \xrightarrow{n \rightarrow +\infty} +\infty$ P_ω^0 -p.s. : en effet, il est P_ω^0 -presque impossible que la suite (X_n) ait une limite inférieure finie tout en restant minorée par m_-^0 , parce que si était le cas, il existerait un entier M tel que la suite (X_n) passe une infinité de fois en M tout en restant minorée par m_-^0 . Mais cela est absurde, car la chaîne de Markov est irréductible, donc quand elle passe une infinité de fois par M , elle passe aussi presque sûrement une infinité de fois par $m_-^0 - 1$. Finalement il y a une P_ω^0 -probabilité d'au moins $1 - \varepsilon$ que X_n tende vers $+\infty$, et comme ε a été choisi quelconque, $X_n \xrightarrow{n \rightarrow \infty} +\infty$ P_ω^0 -p.s. \square

Pour conclure le cas général de cette démonstration, il nous reste finalement à prouver que $E_P(\ln \rho_0) < 0 \Rightarrow \omega \in \mathcal{S}_+ \cap \mathcal{F}_-$ P -p.s., resp. $E_P(\ln \rho_0) > 0 \Rightarrow \omega \in \mathcal{S}_- \cap \mathcal{F}_+$, resp. $E_P(\ln \rho_0) = 0 \Rightarrow \omega \in \mathcal{S}_- \cap \mathcal{F}_-$. Observons d'abord que \mathcal{S}_+ , \mathcal{F}_+ , \mathcal{S}_- et \mathcal{F}_- sont invariants par l'opérateur de décalage. Par ergodicité de P , on a donc $P(\mathcal{S}_+) \in \{0, 1\}$, resp. $P(\mathcal{F}_+) \in \{0, 1\}$, etc.

Déjà, $P(\mathcal{S}_+) = 0$ ou $P(\mathcal{F}_+) = 0$. En effet, supposons $P(\mathcal{S}_+) = 1$. Alors, P -p.s., $\sum_{i \geq 1} \prod_{j=1}^i \rho_j$ converge, donc $\sum_{i \geq 1} (\prod_{j=1}^i \rho_j)^{-1}$ diverge. Mais comme les ρ_z sont indépendants et de même loi, cela est équivalent à dire que $\sum_{i \geq 1} (\prod_{j=-i}^{-1} \rho_j)^{-1}$ diverge P -p.s., c.à.d. que $P(\mathcal{F}_+) = 0$.

À la lumière de ce résultat, nous allons nous contenter de démontrer les propriétés suivantes : $P(\mathcal{S}_+) = 1 \Leftrightarrow E_P(\ln \rho_0) < 0$ et $P(\mathcal{F}_+) = 1 \Leftrightarrow E_P(\ln \rho_0) > 0$. Ces deux résultats sont symétriques, aussi allons-nous seulement démontrer le premier.

Supposons $E_P(\ln \rho_0) < 0$. Alors la loi des grands nombres entraîne, pour P -presque tout milieu ω : $\frac{1}{i} \sum_{j=1}^i \ln(\rho_j) \xrightarrow{i \rightarrow \infty} E_P(\ln \rho_0) < 0$. En particulier, c notant une constante réelle telle que $E_P(\ln \rho_0) < c < 0$, on peut définir pour presque tout milieu ω un rang $i_0(\omega)$ tel que $i \geq i_0(\omega) \Rightarrow \frac{1}{i} \sum_{j=1}^i \ln(\rho_j) < c$. On en tire

$$S(\omega) \stackrel{\text{déf}}{=} \sum_{i=1}^{i_0(\omega)-1} \prod_{j=1}^i \rho_j + \sum_{i=i_0(\omega)}^{\infty} e^{\sum_{j=1}^i \ln \rho_j} \leq \sum_{i=1}^{i_0(\omega)-1} \prod_{j=1}^i \rho_j + \sum_{i=i_0(\omega)}^{\infty} e^{ci} < +\infty$$

d'où $P(\mathcal{S}_+) = 1$.

Supposons $E_P(\ln \rho_0) > 0$. Alors, toujours par la loi des grands nombres, $\lim_{i \rightarrow \infty} \frac{1}{i} \sum_{j=1}^i \ln(\rho_j) > 0$ P -p.s.. En particulier, c' notant une constante réelle telle que $0 < c' < E_P(\ln \rho_0)$, on peut définir pour presque tout milieu ω un rang $i_0(\omega)$ tel que $i \geq i_0(\omega) \Rightarrow \frac{1}{i} \sum_{j=1}^i \ln(\rho_j) > c'$. On en tire comme précédemment

$$S(\omega) \geq \sum_{i=1}^{i_0(\omega)-1} \prod_{j=1}^i \rho_j + \sum_{i=i_0(\omega)}^{\infty} e^{c'i} = +\infty$$

d'où $P(\mathcal{S}_+) = 0$.

Supposons enfin $E_P(\ln \rho_0) = 0$, et notons σ l'écart-type de $\ln \rho_0$. Si $\sigma = 0$, alors $(\forall z \in \mathbb{Z} \quad \rho_z = 1)$ P -p.s. et $P(\mathcal{S}_+) = 0$ est évident. Si $\sigma > 0$, le théorème central limite donne

$$P\left(\sum_{j=1}^i \ln(\rho_j) \geq \sqrt{i}\right) \xrightarrow{i \rightarrow \infty} p \stackrel{\text{déf}}{=} \frac{1}{\sigma\sqrt{2\pi}} \int_1^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx > 0$$

Notant $p' = \frac{p}{2}$, on a donc $P(\sum_{j=1}^i \ln(\rho_j) \geq \sqrt{i}) \geq p'$ à partir d'un certain rang, d'où $P(S(\omega) \geq e^{\sqrt{i}}) \geq p'$. En faisant tendre i vers ∞ , on a donc $P(\mathcal{S}_-) \geq p'$, d'où $P(\mathcal{S}_+) = 0$ par ergodicité.

Le cas général du théorème étant ainsi démontré, il reste à contrôler le cas particulier où, par exemple, $P(\omega_0^+ = 0) > 0$. Dans un tel cas, puisque $E_P(\ln \rho_0)$ est supposé bien définie, on a $P(\omega_0^- = 0) = 0$. D'autre part, $E_P(\ln \rho_0) = -\infty < 0$. La vérification du théorème consiste alors à démontrer $X_n \xrightarrow{n \rightarrow \infty} -\infty$ \mathbb{P}^0 -p.s.

Soit $\omega \in \Omega$. Appelons *points gauches* les $z \in \mathbb{Z}$ vérifiant $\omega_z^+ = 0$. Alors, puisque tous les ω_z sont indépendants et de même loi, il existe P -p.s. une infinité de points gauches, aussi bien parmi les entiers positifs que parmi les entiers négatifs. Il y a donc P -p.s. un point gauche positif ou nul : notons g_0 de plus petit d'entre eux. Notons d'autre part $g_1 > g_2 > \dots$ les points gauches strictement négatifs. Alors P_ω -p.s. la chaîne de Markov partant de 0 va franchir tous les points gauches jusqu'à $-\infty$: d'une part, en effet, quand la marche aléatoire atteint un point gauche, elle le franchit pour ne plus jamais revenir au-delà P_ω -p.s., d'autre part, il est presque impossible que la chaîne de Markov reste coincée à l'infini entre les points g_{i+1} et g_i : en effet, comme tous les points ont une probabilité non nulle d'envoyer le mobile à gauche, le noyau potentiel U de la chaîne de Markov vérifie $U(x, y) > 0$ pour $x > y$, et donc si la chaîne passe une infinité de fois par un point, elle passe aussi une infinité de fois par tous les points plus à gauche. En conclusion, $X_n \xrightarrow{n \rightarrow \infty} -\infty$ P_ω -p.s..

Remarque 4. Le cas particulier que nous venons de traiter correspond au cas où il existe des points que le mobile ne peut franchir que dans un sens, toujours le même (à gauche ci-dessus). Comme aucun point ne fixe le mobile définitivement, il est clair que le mobile va avoir tendance à aller à l'infini, à mesure qu'il franchit irréversiblement les points gauches (resp. droits).

Le théorème 6 est ainsi démontré □

Chapitre 3

Vitesse de dérive

3.1 Énoncé du théorème

Le fait de savoir dans quelle direction part le mobile à l'infini ne nous renseigne pas sur la vitesse à laquelle il s'éloigne de l'origine. En fait, on ne sait même pas si on peut définir une vitesse moyenne d'éloignement. Nous nous proposons donc dans ce chapitre de déterminer le comportement de la suite $(\frac{X_n}{n})$ quand $n \rightarrow \infty$.

Théorème 8. *Plaçons-nous dans les conditions d'application du théorème 6¹. Alors, outre la conclusion dudit théorème, on a :*

1. *si $E_P(\rho_0) < 1$ alors $\frac{X_n}{n} \xrightarrow{n \rightarrow \infty} \frac{1-E_P(\rho_0)}{E_P(\frac{1}{\omega_0^+})} \mathbb{P}^0$ -p.s.*
2. *si $E_P(\frac{1}{\rho_0}) < 1$ alors $\frac{X_n}{n} \xrightarrow{n \rightarrow \infty} -\frac{1-E_P(\frac{1}{\rho_0})}{E_P(\frac{1}{\omega_0^-})} \mathbb{P}^0$ -p.s.*
3. *si $E_P(\rho_0) \geq 1$ et $E_P(\frac{1}{\rho_0}) \geq 1$ alors $\frac{X_n}{n} \xrightarrow{n \rightarrow \infty} 0 \mathbb{P}^0$ -p.s.*

Remarque 5. La vitesse de la marche sous la mesure \mathbb{P}^0 ne dépend que de la mesure sur le milieu P .

Remarque 6. Si $\omega^+ = 1 - \omega^-$ fixés (marche aléatoire simple) on obtient bien le résultat connu. Par exemple si $\omega^+ > \omega^-$

$$\lim_{n \rightarrow \infty} \frac{X_n}{n} = \frac{1 - \frac{\omega^-}{\omega^+}}{\frac{1}{\omega^+}} = \omega^+ - \omega^-$$

Sinon ce résultat est faux. On obtient par exemple dans le cas $v_P > 0$

$$v_P = \frac{1 - E_P(\rho_0)}{E_P(\frac{1}{\omega_0^+})} \leq \frac{1 - E_P(\rho_0)}{1 + E_P(\rho_0)}$$

Notons $p_0 = \frac{\omega_0^+}{\omega_0^- + \omega_0^+}$. Par l'inégalité de Jensen

$$E_P(\rho_0) = E_P\left(\frac{1}{p_0}\right) - 1 \geq \frac{1}{E_P(p_0)} - 1$$

¹À savoir, aucun point ne fixe la marche aléatoire et $E_P(\ln(\rho_0))$ est bien défini.

d'où

$$v \leq \frac{1 - E_P(\rho_0)}{1 + E_P(\rho_0)} \leq \frac{2 - \frac{1}{E_P(p_0)}}{\frac{1}{E_P(p_0)}} = 2E_P(p_0) - 1 = E_P(\omega_0^+ - \omega_0^-)$$

où la première inégalité est stricte dès lors que $P(\omega_0^0 \neq 0) > 0$, et la seconde dès lors que $\text{var}(\rho_0) > 0$.

Remarque 7. Il existe le cas d'une marche transitoire à vitesse nulle. Par exemple si on prend $P\{\omega^- = (1 - \omega^+) = \frac{1}{3}\} = \frac{1}{2}$ et $P\{\omega^- = (1 - \omega^+) = \frac{3}{5}\} = \frac{1}{2}$ alors $E(\ln \rho_0) = \frac{1}{2} (\ln(\frac{3}{2}) + \ln(\frac{1}{2})) < 0$ mais $E[\rho_0] = 1$.

3.2 Démonstration

Démonstration. Remarquons d'abord qu'on a $\ln(E_P(\frac{1}{\rho_0})) \geq E_P(-\ln(\rho_0))$ par l'inégalité de Jensen, donc, si $E_P(\ln \rho_0) \leq 0$, seuls les cas 1 et 3 sont possibles. Nous allons donc démontrer le théorème sous cette éventualité seulement, sachant que le cas $E_P(\ln \rho_0) \geq 0$ se traite de manière symétrique². L'intérêt de se placer sous cette hypothèse est que le théorème 6 nous permet alors d'affirmer que $\limsup_{n \rightarrow \infty} X_n = +\infty$ \mathbb{P}^0 -p.s., ce qui autorise les notations suivantes :

Posons $T_n \stackrel{\text{déf}}{=} \min\{k; X_k = n\}$, et $\tau_n \stackrel{\text{déf}}{=} T_n - T_{n-1}$. T_n représente le temps que met le mobile, partant de 0, pour atteindre le point n , tandis que τ_n est le temps qui s'écoule entre la première fois où il atteint le point $n - 1$ et la première fois où il atteint le point n ; les T_n et les τ_n sont finis \mathbb{P}^0 -p.s., en vertu de la remarque précédente.

On a :

Lemme 9. *Sous la probabilité \mathbb{P}^0 , la suite $(\tau_i)_{i \geq 1}$ est stationnaire et ergodique.*

Démonstration. La stationnarité est intuitivement claire, car remplacer τ_n par τ_{n+1} , c'est en quelque sorte remplacer n par $n + 1$, c'est-à-dire décaler le milieu d'un cran, ce qui ne change pas la distribution du milieu d'après l'hypothèse 1. Plus précisément, soient $A_1, \dots, A_n \subset \mathbb{N}$. Nous voulons établir que $\mathbb{P}^0(\tau_1 \in A_1, \dots, \tau_n \in A_n) = \mathbb{P}^0(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n)$. Pour ce faire, notons qu'on a par changement d'origine du milieu :

$$P_\omega^0(\tau_1 \in A_1, \dots, \tau_n \in A_n) = P_{\theta^{-1}\omega}^1(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n)$$

D'autre part, dans un milieu fixé, la propriété de Markov forte nous dit que la loi du déplacement du mobile à partir du moment où il a atteint le point 1 est la même que celle qu'il a quand il en part. En particulier, comme le mobile atteint presque certainement le point 1, on a pour presque tout milieu ω :

$$P_\omega^1(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n) = P_\omega^0(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n)$$

²On peut aussi démontrer ce cas en se servant du résultat pour $E_P(\ln \rho_0) \leq 0$ et en inversant gauche et droite.

Donc :

$$\begin{aligned}
& \mathbb{P}^0(\tau_1 \in A_1, \dots, \tau_n \in A_n) \\
&= \int_{\Omega} P_{\omega}^0(\tau_1 \in A_1, \dots, \tau_n \in A_n) \, d\omega \\
&= \int_{\Omega} P_{\theta^{-1}\omega}^1(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n) \, d\omega \\
&= \int_{\Omega} P_{\omega}^1(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n) \, d\omega \\
&\quad \text{(car } \theta^{-1} \text{ préserve la mesure)} \\
&= \int_{\Omega} P_{\omega}^0(\tau_2 \in A_1, \dots, \tau_{n+1} \in A_n) \, d\omega \\
&= \mathbb{P}^0(\tau_1 \in A_1, \dots, \tau_n \in A_n)
\end{aligned}$$

Et la stationnarité est démontrée.

Pour l'ergodicité, posons $\Xi \stackrel{\text{déf}}{=} [0, 1]^{\mathbb{N}}$, $\xi = (\xi_n)_{n \in \mathbb{N}}$ ses éléments, et munissons-le de la mesure U_{Ξ} qui rend toutes les coordonnées ξ_n indépendantes et de loi uniforme sur $[0, 1]$. On peut alors construire $(X_n)_{n \in \mathbb{N}}$, à ω donné, comme une variable aléatoire définie sur Ξ par :

$$X_{n+1} \stackrel{\text{déf}}{=} X_n + \mathbb{1}_{\{\xi_{n+1} < \omega_{X_n}^+\}} - \mathbb{1}_{\{\xi_{n+1} > 1 - \omega_{X_n}^-\}}$$

(C'est d'ailleurs ainsi qu'on procède lors des simulations de marches aléatoires).

Maintenant, soit $A(\tau)$ un événement mesurable par rapport aux τ_i qui est invariant par décalage en τ (nous noterons également θ le décalage en τ , et nous écrirons $\theta A(\tau)$ pour $A(\theta\tau)$ ³). Nous voulons montrer $P \otimes U_{\Xi}(A) \in \{0, 1\}$. Commençons par fixer le milieu ω . Soit alors $k \in \mathbb{N}$. Comme θ est surjective, on a $A = \theta A$, d'où $A = \theta^k A$. Donc $A(\xi)$ ne dépend que de $\tau_{k+1}, \tau_{k+2}, \dots$, et donc que du comportement de la chaîne de Markov à partir du moment où elle atteint le point k . En particulier, A est indépendant de $\sigma(\xi_1, \dots, \xi_k)$, qui est la tribu relative aux k premiers mouvements du mobile. Comme cela est vrai pour tout k , A est indépendant de $\sigma(\xi_1, \xi_2, \dots) = \sigma(\xi)$, donc trivial. Finalement, A ne dépend que de ω . Mais, comme nous l'avons fait remarquer plus haut, on peut décaler les τ_n d'une unité en décalant le milieu d'une unité. Par conséquent, $A(\theta\omega) = \theta A(\omega) = A(\omega)$. A est donc un événement sur $(\omega_z)_{z \in \mathbb{Z}}$ invariant par décalage, et comme le milieu est ergodique d'après le théorème 3, $P(A) \in \{0, 1\}$, CQFD \square

Nous passons maintenant au lemme fondamental suivant :

Lemme 10.

$$E_{\mathbb{P}^0}(\tau_1) \begin{cases} \frac{E_P(\frac{1}{\omega_0^+})}{1 - E_P(\rho_0)} & \text{si } E_P(\rho_0) < 1 \\ +\infty & \text{si } E_P(\rho_0) \geq 1 \end{cases}$$

Démonstration. L'idée consiste à écrire une relation que $E_{\mathbb{P}^0}(\tau_1)$ satisfait vis-à-vis de lui-même, en utilisant l'invariance par décalage du milieu. Pour ce faire, on va avoir recours à la décomposition des temps d'atteinte : notons

$$\tau_1 = \mathbb{1}_{\{X_1=1\}} + \mathbb{1}_{\{X_1=0\}}(1 + \tau_1') + \mathbb{1}_{\{X_1=-1\}}(1 + \tau_0'' + \tau_1'') \quad (3.1)$$

³L'hypothèse d'invariance par décalage en τ s'écrit donc $A = \theta^{-1}(A)$.

où τ'_1 est le temps mis par le mobile pour aller du point 0 au point 1 après être resté immobile au temps 1, τ''_0 est le temps qu'il met pour revenir au point 0 s'il est parti en -1 lors de son premier mouvement, et τ''_1 le temps qu'il met à aller en 1 après être parti à gauche, puis revenu au point 0 plus tard.

Fixons un milieu ω . La propriété de Markov forte nous dit immédiatement que τ'_1 , τ''_0 , et τ''_1 , conditionnés respectivement par les événements $\{X_1 = 0\}$, $\{X_1 = -1\}$ et $\{X_1 = -1\} \cap \{\tau''_0 < \infty\}$, suivent les mêmes lois que τ_1 , $\tau_1(\theta^{-1}\omega)$ et τ_1 .

Supposons d'abord $E_{\mathbb{P}^0}(\tau_1) < \infty$. Alors, P -p.s., on a $E_{\omega}^0(\tau_1) < \infty$ et $E_{\theta^{-1}\omega}^0(\tau_1) < \infty^4$. Passant à l'espérance dans (3.1), on obtient :

$$E_{\omega}^0(\tau_1) = 1 + (\omega_0^0 + \omega_0^-)E_{\omega}^0(\tau_1) + \omega_-^0 E_{\theta^{-1}\omega}^0(\tau_1) \quad (3.2)$$

soit

$$E_{\omega}^0(\tau_1) = \frac{1}{\omega_0^+} + \rho_0 E_{\theta^{-1}\omega}^0(\tau_1) \quad (3.3)$$

après simplification.

Mais $E_{\theta^{-1}\omega}^0(\tau_1)$ ne dépend que du comportement de la chaîne de Markov dans le milieu $\theta^{-1}\omega$ avant qu'elle atteigne le point 1, donc que, dans le milieu $\theta^{-1}\omega$, des points $0, -1, -2, \dots$, autrement dit de $\omega_{-1}, \omega_{-2}, \omega_{-3}, \dots$. Or tous les ω_z sont indépendants vis-à-vis de P , donc $E_{\theta^{-1}\omega}^0(\tau_1)$ est indépendante de $\rho_0 = \frac{\omega_0^-}{\omega_0^+}$. Le passage à l'espérance selon P donne donc :

$$E_{\mathbb{P}^0}(\tau_1) = E_P\left(\frac{1}{\omega_0^+}\right) + E_P(\rho_0) \int_{\Omega} E_{\theta^{-1}\omega}^0(\tau_1) P(d\omega)$$

Mais comme P est θ^{-1} -invariante, $\int_{\Omega} E_{\theta^{-1}\omega}^0(\tau_1) P(d\omega) = \int_{\Omega} E_{\omega}^0(\tau_1) P(d\omega) = E_{\mathbb{P}^0}(\tau_1)$, d'où $E_{\mathbb{P}^0}(\tau_1) = \frac{E_P(\frac{1}{\omega_0^+})}{1 - E_P(\rho_0)}$ puisqu'on a supposé $E_{\mathbb{P}^0}(\tau_1) < \infty$.

Ce résultat montre par ailleurs que $E_{\mathbb{P}^0}(\tau_1) < \infty \Rightarrow E_P(\rho_0) < 1$, il ne reste donc plus qu'à démontrer $E_{\mathbb{P}^0}(\tau_1) = \infty \Rightarrow E_P(\rho_0) \geq 1$

Supposons par l'absurde $E_{\mathbb{P}^0}(\tau_1) = \infty$ et $E_P(\rho_0) < 1$. On peut, en modifiant légèrement le raisonnement ci-dessus, adapter (3.2) en :

$$E_{\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \leq 1 + (\omega_0^0 + \omega_0^-)E_{\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) + \omega_-^0 E_{\theta^{-1}\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \quad (3.4)$$

d'où

$$E_{\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \leq \frac{1}{\omega_0^+} + \rho_0 E_{\theta^{-1}\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \quad (3.5)$$

L'itération de cette équation donne :

$$E_{\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \leq \frac{1}{\omega_0^+} + \frac{\rho_0}{\omega_{-1}^+} + \frac{\rho_0 \rho_{-1}}{\omega_{-2}^+} + \dots + \frac{\prod_{i=0}^{m-1} \rho_{-i}}{\omega_{-m}^+} + \left(\prod_{i=0}^{m-1} \rho_{-i}\right) E_{\theta^{-m}\omega}^0(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}})$$

⁴En effet, θ^{-1} préserve la mesure, donc si un événement est vrai P -p.s., il reste vrai P -p.s. quand on décale le milieu d'un cran.

$$\leq \bar{S} + M \prod_{i=0}^{m-1} \rho_{-i}$$

où \bar{S} note $\sum_{i=0}^{\infty} \frac{1}{\omega_{-i}^+} \prod_{j=0}^{i-1} \rho_{-j} = \frac{1}{\omega_0^+} + \frac{\rho_0}{\omega_{-1}^+} + \frac{\rho_0 \rho_{-1}}{\omega_{-2}^+} + \dots$

En passant à l'espérance selon P , on obtient

$$E_{\mathbb{P}^0}(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \leq E_P(\bar{S}) + M E_P\left(\prod_{i=0}^{m-1} \rho_{-i}\right) = E_P(\bar{S}) + M E_P(\rho_0)^m \quad (3.6)$$

Or on a facilement $E_P(\bar{S}) = \frac{E_P(\frac{1}{\omega_0^+})}{1 - E_P(\rho_0)}$. En faisant tendre m vers $+\infty$ dans (3.6), on obtient donc

$$E_{\mathbb{P}^0}(\tau_1 \mathbb{1}_{\{\tau_1 \leq M\}}) \leq E_P(\bar{S})$$

puis en faisant tendre M vers $+\infty$, par convergence monotone :

$$\infty = E_{\mathbb{P}^0}(\tau_1) = E_{\mathbb{P}^0}(\tau_1 \mathbb{1}_{\{\tau_1 < \infty\}}) \leq E_P(\bar{S}) < \infty$$

Ce qui est absurde et termine la preuve du lemme □

Remarque 8. Pour $E_P(\rho_0) \geq 1$, on a $E_P(\bar{S}) = \infty$, et donc dans tous les cas $E_{\mathbb{P}^0}(\tau_1) = E_P(\bar{S})$.

Puisque la suite $(\tau_i)_{i \geq 1}$ est ergodique (cf. lemme 9) sous la probabilité \mathbb{P}^0 , on a $\frac{T_n}{n} = \frac{\sum_{i=1}^n \tau_i}{n} \xrightarrow{n \rightarrow \infty} E_{\mathbb{P}^0}(\tau_1) = E_P(\bar{S})$ \mathbb{P}^0 -p.s.. Cela suffit pour démontrer le théorème :

Cas 1 : $E_P(\bar{S}) < \infty$. On utilise alors le lemme élémentaire suivant : si $\frac{T_n}{n} \xrightarrow{n \rightarrow \infty} \alpha < \infty$, alors $\frac{X_n}{n} \rightarrow \frac{1}{\alpha}$ ⁵. En effet, soit une suite $(X_n)_{n \in \mathbb{N}}$ avec $\limsup_{n \rightarrow \infty} X_n = +\infty$. Pour $n \in \mathbb{N}$, notons k_n la valeur maximale atteinte par la suite (X_n) à l'instant n : k_n est caractérisé par $T_{k_n} \leq n < T_{k_n+1}$ (en particulier, $k_{T_n} = n$). On a trivialement $X_n \leq n$, et aussi $X_n \geq k_n - (n - T_{k_n})$, car après avoir atteint la valeur k_n à l'instant T_{k_n} , la suite redescend avec une vitesse au plus égale à 1. On en tire la double inégalité :

$$\frac{k_n}{n} - (1 - \frac{T_{k_n}}{n}) \leq \frac{X_n}{n} \leq \frac{k_n}{n} \quad (3.7)$$

Mais $\frac{k_n}{n} = \frac{k_n T_{k_n}}{T_{k_n} n} \leq \frac{k_n}{T_{k_n}}$, et comme $\frac{m}{T_m} \xrightarrow{m \rightarrow \infty} \frac{1}{\alpha}$, on a $\limsup_{n \rightarrow \infty} \frac{k_n}{n} \leq \frac{1}{\alpha}$. Mais on a aussi $\frac{k_n}{n} = \frac{k_n}{k_n+1} \frac{k_n+1}{T_{k_n+1}} \frac{T_{k_n+1}}{n} \sim_{n \rightarrow \infty} \frac{k_n+1}{T_{k_n+1}} \frac{T_{k_n+1}}{n} > \frac{k_n+1}{T_{k_n+1}} \xrightarrow{n \rightarrow \infty} \frac{1}{\alpha}$, d'où $\liminf \frac{k_n}{n} \geq \frac{1}{\alpha}$ et finalement $\frac{k_n}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{\alpha}$.

On en déduit d'autre part que $\frac{T_{k_n}}{n} \xrightarrow{n \rightarrow \infty} 1$, puisque $\frac{T_{k_n}}{n} = \frac{T_{k_n} k_n}{k_n n}$. Donc, en passant à la limite dans (3.7), on obtient bien $\frac{X_n}{n} \rightarrow \frac{1}{\alpha}$, et le lemme est démontré, d'où le résultat.

⁵C'est un lemme parfaitement déterministe : aucune hypothèse particulière n'est faite sur la suite (X_n) .

Cas 2 : $E_P(\overline{S}) = \infty$. Avec les mêmes notations, on a $\frac{X_n}{n} \leq \frac{k_n}{n} = \frac{k_n}{T_{k_n}} = \left(\frac{T_{k_n}}{k_n}\right)^{-1} \rightarrow_{n \rightarrow \infty} 0$, d'où $\limsup_{n \rightarrow \infty} \frac{X_n}{n} \leq 0$. Si on est dans le cas transitoire, cela suffit pour conclure que $\frac{X_n}{n} \rightarrow_{n \rightarrow \infty} 0$; et si on est dans le cas récurrent, on démontre de manière symétrique que $\limsup_{n \rightarrow \infty} \frac{X_n}{n} \leq 0$, d'où aussi $\frac{X_n}{n} \rightarrow_{n \rightarrow \infty} 0$.

Le théorème 8 est donc démontré

□

Chapitre 4

Chaînes auxiliaires

Un des problèmes en étudiant les marches aléatoires en milieu aléatoire est que le processus X_n n'est pas une chaîne de Markov. Les lois de transition dépendent du milieu et donc ne sont pas $\sigma(X_n)$ -mesurables. Pour éviter ce problème nous considérons un processus auxiliaire, celui du changement de milieu vu du mobile. C'est bien une chaîne de Markov, ce qui nous permet d'utiliser les techniques disponibles pour les chaînes de Markov dans notre cadre. À la fin de ce chapitre nous donnerons encore une preuve du théorème 8 concernant la vitesse de la marche. Nous allons nous restreindre au cas 1. de vitesse positive sachant que la preuve du 2. est symétrique. Le cas 3 ne sera pas traité. Dans ce chapitre nous supposons donc toujours que $E(\rho_0) < 1$. On note la vitesse de la marche v_P .

4.1 Chaînes auxiliaires

Nous considérons le processus $\bar{\omega}(n) = \theta^{X_n}\omega$ défini sur l'espace mesurable $(\Omega \times \mathbb{Z}^{\mathbb{N}}, \mathcal{F} \times \mathcal{G})$ muni des mesures P_ω^0 et \mathbb{P}^0 . $\bar{\omega}(n)$ est le milieu que voit le mobile à l'instant n .

Nous généralisons d'abord la notion de chaîne de Markov sur des espaces non dénombrables. On considère un espace polonais (E, τ) muni de sa tribu borélienne et M une probabilité de transition de E dans lui-même. Comme dans le cas discret on définit pour $f : E \rightarrow \mathbb{R}$

$$Mf(\omega) = \int f(\omega')M(\omega, d\omega').$$

Définition 3. Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires à valeurs dans E est appelée *chaîne de Markov* de noyau de transition M si $\forall n \in \mathbb{N}$ la loi conditionnelle de X_{n+1} sachant (X_0, \dots, X_n) est $M(X_n, \bullet)$. Cela signifie que $\forall n \in \mathbb{N}, \forall f : E^{n+1} \rightarrow \mathbb{R}_+, \forall g : E \rightarrow \mathbb{R}_+$

$$E(f(X_0, \dots, X_n)g(X_{n+1})) = E(f(X_0, \dots, X_n)Mg(X_n)).$$

Proposition 11. Soit X_n une MAMA. La suite $(\bar{\omega}(n))_{n \in \mathbb{N}}$ est une chaîne de Markov par rapport à la probabilité de transition

$$M(\omega, \bullet) = \omega_0^+ \delta_{\theta\omega} + \omega_0^- \delta_{\theta^{-1}\omega} + \omega_0^0 \delta_\omega$$

sous P_ω^0 et \mathbb{P}^0 .

Démonstration. Sous P_ω^0 : Soit $n \in \mathbb{N}$, $f : \Omega^{n+1} \rightarrow \mathbb{R}_+$ et $g : \Omega \rightarrow \mathbb{R}_+$ mesurable. On a

$$\begin{aligned} E_\omega^0 \left(f(\bar{\omega}(0), \dots, \bar{\omega}(n)) g(\bar{\omega}(n+1)) \right) &= E_\omega^0 \left(f(\theta^{X_0}\omega, \dots, \theta^{X_n}\omega) g(\theta^{X_{n+1}}\omega) \right) \\ &= E_\omega^0 \left(f(\theta^{X_0}\omega, \dots, \theta^{X_n}\omega) E_\omega^{X_{n-1}}(g(\theta^{X_n}\omega)) \right) \\ &= E_\omega^0 \left(f(\theta^{X_0}\omega, \dots, \theta^{X_n}\omega) [\omega_{X_{n-1}}^+ g(\theta \circ \theta^{X_{n-1}}\omega) \right. \\ &\quad \left. + \omega_{X_{n-1}}^- g(\theta^{-1} \circ \theta^{X_{n-1}}\omega) + \omega_{X_{n-1}}^0 g(\theta^{X_{n-1}}\omega)] \right) \\ &= E_\omega^0 (f(\bar{\omega}(0), \dots, \bar{\omega}(n)) M g(\bar{\omega}(n))) \end{aligned}$$

où la deuxième égalité découle de la propriété de Markov de (X_n) . On obtient le résultat pour \mathbb{P}^0 en intégrant cette égalité par rapport à P . \square

Nous allons chercher une mesure invariante pour cette chaîne dans le cas qui nous intéresse. La construction ressemble fortement à celle dans le cas discret.

On rappelle que $T_1 = \inf\{t > 0 | X_t = 1\}$. Si on se met dans le cas transitoire où $E_{\mathbb{P}} \log \rho_0 < 0$ c.à.d. où $(X_n)_{n \in \mathbb{N}}$ diverge \mathbb{P}^0 -p.s. vers $+\infty$ on a $T_1 < \infty$ \mathbb{P}^0 -p.s.. Dans le lemme 10 nous avons déjà vu que $E_P(\rho_0) < 1$ implique $E_{\mathbb{P}^0}(T_1) < \infty$. On peut donc définir une mesure de probabilité \bar{Q} sur Ω par

$$Q(B) = E_{\mathbb{P}^0} \left(\sum_{i=0}^{T_1-1} \mathbf{1}_{\{\bar{\omega}(i) \in B\}} \right) \quad \bar{Q}(B) = \frac{Q(B)}{Q(\Omega)} = \frac{Q(B)}{E_{\mathbb{P}^0} T_1}.$$

On compte combien de fois la marche X_n partant de 0 voit un milieu dans B avant d'arriver au niveau 1. Le concept de compter combien de fois le processus $\bar{\omega}(n)$ tombe dans B avant un temps dépendant du processus est le même que dans le cas discret. La différence est que ce temps T_1 ne dépend pas directement de $\bar{\omega}(n)$ mais de X_n .

Proposition 12. La mesure $Q(\bullet)$ est M -invariante c.à.d. pour chaque $B \in \mathcal{F}$ on a

$$Q(B) = \int_{\Omega} M(x, B) Q(dx).$$

Remarque 9. Notre mesure \bar{Q} est donc en un sens le milieu typique dans lequel le mobile se trouve pendant son trajet. Cette caractérisation est importante pour comprendre pourquoi elle permet d'obtenir des conclusions concernant la vitesse, en effet la vitesse instantanée du mobile dépend du point où il se trouve.

Démonstration. Nous calculons $\int_{\Omega} M(x, B)Q(dx)$. Par définition de Q on a

$$\begin{aligned}\int_{\Omega} M(x, B)Q(dx) &= E_{\mathbb{P}^0} \left(\sum_{i=0}^{T_1-1} \mathbf{1}_{\{\bar{\omega}(i+1) \in B\}} \right) \\ &= \sum_{k=0}^{\infty} E_{\mathbb{P}^0} \left(\mathbf{1}_{\{k < T_1\}} \mathbf{1}_{\{\bar{\omega}(k+1) \in B\}} \right).\end{aligned}$$

Puis nous divisons l'événement $\{k < T_1\}$ en $\{k+1 = T_1\}$ et $\{T_1 > k+1\}$. Donc

$$\begin{aligned}\int_{\Omega} M(x, B)Q(dx) &= \sum_{k=0}^{\infty} E_{\mathbb{P}^0} \left(\mathbf{1}_{\{T_1=k+1\}} \mathbf{1}_{\{\bar{\omega}(k+1) \in B\}} \right) + \sum_{k=0}^{\infty} E_{\mathbb{P}^0} \left(\mathbf{1}_{\{T_1 > k+1\}} \mathbf{1}_{\{\bar{\omega}(k+1) \in B\}} \right) \\ &= \mathbb{P}^0(T_1 < \infty; \bar{\omega}(T_1) \in B) + \sum_{k=1}^{\infty} \mathbb{P}^0(T_1 > k; \bar{\omega}(k) \in B).\end{aligned}$$

Jusqu'ici la preuve est presque la même que dans le cas discret. Pour évaluer cette expression nous utilisons maintenant des propriétés particulières de la marche. Comme $\mathbb{P}^0(T_1 < \infty) = 1$ et $\bar{\omega}(T_1) = \theta\omega$ le terme de gauche vaut $P(\theta\omega \in B)$ par l'invariance sous décalage de la mesure P c'est égal à $P(\omega \in B)$. On obtient donc

$$\int_{\Omega} M(x, B)Q(dx) = \sum_{k=0}^{\infty} \mathbb{P}^0(\mathbf{1}_{T_1 > k} \bar{\omega}(k) \in B) = Q(B)$$

□

On voit facilement que Q est absolument continue par rapport à P . Calculons la dérivée de Radon-Nikodym. On définit

$$\Lambda(\omega) = \frac{1}{\omega_0^+} \left[1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \rho_j \right]$$

Proposition 13.

$$\frac{dQ}{dP}(\omega) = \Lambda(\omega)$$

Démonstration. Soit $f : \Omega \rightarrow \mathbb{R}$ mesurable bornée. Par définition de Q on a

$$\int f dQ = E_{\mathbb{P}^0} \left(\sum_{i=0}^{T_1-1} f(\bar{\omega}_i) \right)$$

On note $N_i = \#\{k \in \{0, \dots, T_1 - 1\} | X_k = i\}$. Alors on peut réécrire le terme de droite comme

$$E_{\mathbb{P}^0} \left(\sum_{i \leq 0} f(\theta^i \omega) N_i \right)$$

Ceci est égal à

$$\begin{aligned}& \sum_{i \leq 0} E_P(f(\theta^i \omega) E_{\omega}^0 N_i) && \text{par définition de la mesure } \mathbb{P}^0 \\ &= \sum_{i \leq 0} E_P(f(\omega) E_{\theta^{-i}\omega}^0 N_i) && \text{par invariance sous décalage de } P \\ &= E_P(f(\omega) (\sum_{i \leq 0} E_{\theta^{-i}\omega}^0 N_i))\end{aligned}$$

Alors

$$\frac{dQ}{dP} = \sum_{i \leq 0} E_{\theta^{-i}\omega}^0 N_i$$

En particulier cette somme converge.

Afin d'évaluer cette expression nous utilisons un argument de nombres de montées. Remarquons d'abord que nous cherchons à calculer l'espérance des N_i par rapport à la mesure $P_{\theta^{-i}\omega}^0$ càd. dans un milieu fixé. Nous considérons donc la chaîne de Markov X_n pour un milieu ω fixé.

On pose

$$\begin{aligned} \theta_{i,1} &= \min\{k \leq T_1 | X_k = i - 1\} \\ \eta_{i,1} &= \min\{\theta_{i,1} < k \leq T_1 | X_k = i\} \end{aligned}$$

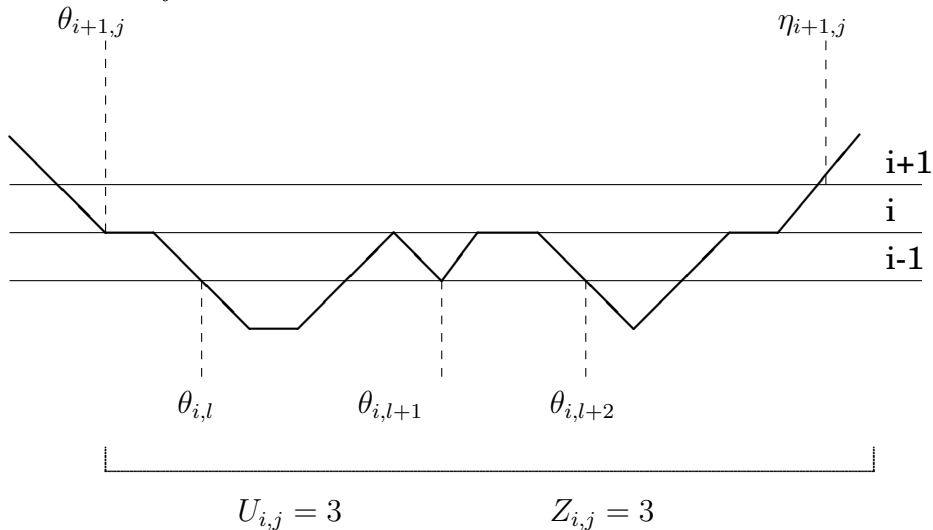
et pour $j \geq 2$

$$\begin{aligned} \theta_{i,j} &= \min\{\eta_{i,j-1} < k \leq T_1 | X_k = i - 1\} \\ \eta_{i,j} &= \min\{\theta_{i,j} < k \leq T_1 | X_k = i\} \end{aligned}$$

$\theta_{i,j}$ est la j -ième fois que la marche atteint le niveau $i - 1$ venant d'au-dessus et $\eta_{i,j}$ est la première fois après cela qu'elle remonte au niveau i . Nous appelons l'intervalle $\{\theta_{i,j}, \dots, \eta_{i,j}\}$ la j -ième excursion partant de i de la marche. Soient

$$\begin{aligned} U_{i,j} &= \#\{l \geq 0 : \theta_{i+1,j} < \theta_{i,l} < \eta_{i+1,j}\} \\ Z_{i,j} &= \#\{k \geq 0 : X_{k-1} = i; X_k = i, \theta_{i+1,j} < k < \eta_{i+1,j}\} \end{aligned}$$

$U_{i,j}$ est le nombre d'excursions partant de i pendant la j -ième excursion partant de $i + 1$ et $Z_{i,j}$ est le nombre de fois que X_n reste en i pendant la même excursion. On définit encore $U_i = \sum_j U_{i,j}$ le nombre total d'excursions partant de i et $Z_i = \sum_j Z_{i,j}$ le nombre total de fois où le mobile reste en i .



On peut exprimer N_i en fonction de U_j, Z_j , plus précisément $N_i = U_i + U_{i+1} + Z_i$. Afin d'évaluer les N_i il suffit donc de calculer la loi des U_i et des Z_i .

Chaque excursion partant de i fait partie d'une excursion partant de $i + 1$. On voit alors que

$$U_i = \sum_{j=0}^{U_{i+1}-1} U_{i,j}$$

Les $U_{i,j}$ sont indépendants et de même loi pour un i fixé. Le processus U_i est donc un processus de branchement.

Pour calculer la loi de $U_{i,j}$ étudions ce qui se passe entre $\theta_{i+1,j}$ et $\eta_{i+1,j+1}$. À $\theta_{i+1,j}$ le mobile monte au niveau $i + 1$ avec une probabilité ω^+ et il descend avec une probabilité ω^- . S'il reste immobile la situation ne change pas. On a donc

$$P_\omega^0(U_{i,j} \geq 1) = \frac{\omega_i^-}{\omega_i^- + \omega_i^+}$$

S'il descend il va forcément revenir en i et on aura la même situation. Comme le processus X_n est markovien ce deuxième choix va être indépendant du premier. On obtient

$$P_\omega^0(U_{i,j} \geq 2) = \left(\frac{\omega_i^-}{\omega_i^- + \omega_i^+} \right)^2$$

En itérant cet argument on voit que $U_{i,j}$ suit une loi géométrique de paramètre $\frac{\omega_i^-}{\omega_i^- + \omega_i^+}$. On peut donc calculer

$$E_\omega^0(U_i | U_{i+1}, \dots, U_0) = U_{i+1} \cdot E(U_{i,j}) = \rho_i U_{i+1}$$

Un argument similaire montre que $Z_{i,j}$ suivent une loi géométrique de paramètre $\frac{\omega_i^0}{\omega_i^0 + \omega_i^+}$. Comme pour les $U_{i,j}$ on écrit

$$Z_i = \sum_{j=0}^{U_{i+1}-1} Z_{i,j}$$

et obtient

$$E_\omega^0(Z_i | U_{i+1}, \dots, U_0) = U_{i+1} \cdot E(Z_{i,j}) = \frac{\omega_i^0}{\omega_i^+} U_{i+1}$$

En utilisant la relation $N_i = U_i + U_{i+1} + Z_i$ on voit que

$$E_\omega^0(N_i | U_{i+1}, \dots, U_0) = E_\omega^0(U_i + U_{i+1} + Z_i | U_{i+1}, \dots, U_0) = \frac{1}{\omega_i^+} E_\omega^0 U_{i+1}$$

Un calcul par récurrence montre que

$$E_\omega^0 N_i = \frac{1}{\omega_i^+} \rho_0 \dots \rho_{i+1}$$

Donc

$$\frac{dQ}{dP} = \frac{1}{\omega_0^+} \left[1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \rho_j \right]$$

et la proposition est démontrée. \square

Avant de passer à la preuve du théorème 8 nous démontrons encore un résultat technique.

Lemme 14. *Sous la probabilité $\bar{Q} \otimes P_\omega^0$ la suite $(\bar{\omega}(n))_{n \in \mathbb{N}}$ est stationnaire et ergodique.*

Démonstration. Fixons d'abord quelques notations. On note \bar{P}_ω la loi de la suite $(\bar{\omega}(n))_{n \in \mathbb{N}}$ avec $\bar{\omega}(0) = \omega$, càd. pour $B_1, \dots, B_l \in \Omega$ mesurables

$$\bar{P}_\omega(\bar{\omega}(i) \in B_i, i = 1, \dots, l) = \int_{B_1} \dots \int_{B_l} M(\omega, d\omega^1) M(\omega^1, d\omega^2) \dots M(\omega^{l-1}, d\omega^l)$$

Puis on note $\bar{\theta}$ l'opérateur de décalage sur $\Omega^{\mathbb{N}}$ c.a.d. $(\bar{\theta}\bar{\omega})(n) = \bar{\omega}(n+1)$ et finalement $\bar{Q} = \bar{Q} \otimes \bar{P}_\omega$. La loi \bar{Q} est celle du processus $(\bar{\omega}(n))_{n \in \mathbb{N}}$.

Montrons que le système $(\Omega^{\mathbb{N}}, \bar{Q}, \bar{\theta})$ est stationnaire ergodique.

D'abord le fait qu'il est stationnaire découle du fait que \bar{Q} est invariante. Les détails sont exactement comme dans la preuve du lemme 9.

Pour l'ergodicité il faut voir que chaque ensemble invariant A a une probabilité triviale. Pour un événement invariant A on pose $\varphi(\omega) = \bar{P}_\omega(A)$ la probabilité qu'un processus commençant à ω soit dans A . Si on regarde la suite de variables aléatoires $(\varphi(\bar{\omega}(n)))_{n \in \mathbb{N}}$ on trouve que c'est une martingale par rapport à la filtration $\mathcal{G}_n = \sigma(\bar{\omega}(0) \dots, \bar{\omega}(n))$. En effet

$$\varphi(\bar{\omega}(n)) = \bar{P}_{\bar{\omega}(n)}(A) = E_{\bar{Q}}(\mathbf{1}_{\bar{\theta}^n A} | \mathcal{G}_n) = E_{\bar{Q}}(\mathbf{1}_A | \mathcal{G}_n)(\bar{\omega}(0))$$

où la deuxième égalité est due à la propriété de Markov et la troisième à l'invariance de A . On voit que cette martingale est fermée et converge donc \bar{Q} p.s. et dans $L^1(\bar{Q})$:

$$\varphi(\bar{\omega}(n)) \rightarrow \mathbf{1}_A \quad \text{quand } n \rightarrow \infty$$

Montrons que $Q(\varphi(\bar{\omega}) \notin \{0, 1\}) = 0$. Regardons un intervalle $[a, b]$ ne contenant pas 0 et 1. Le théorème de Birkhoff nous dit que

$$\frac{1}{n} \sum_0^{n-1} \mathbf{1}_{\{\varphi(\bar{\omega}(n)) \in [a, b]\}} \rightarrow E_{\bar{Q}}(\mathbf{1}_{\{\varphi(\bar{\omega}(n)) \in [a, b]\}} | \mathcal{I})$$

où \mathcal{I} est la tribu invariante. Comme $\varphi(\bar{\omega}(n))$ tend presque sûrement vers une indicatrice le terme à gauche vaut 0 et en intégrant par rapport à \bar{Q} on obtient

$$0 = \bar{Q}(\varphi(\bar{\omega}(0)) \in [a, b]) = \bar{Q}(\varphi(\omega) \in [a, b])$$

Ceci montre que $Q(\varphi(\bar{\omega}) \notin \{0, 1\}) = 0$.

On sait donc que $\varphi(\omega)$ est une fonction indicatrice. Il existe $B \subset \Omega$, $\varphi(\omega) = \mathbf{1}_B$. De plus, la propriété de Markov et l'invariance de A impliquent que $M\mathbf{1}_B = \mathbf{1}_B$, \bar{Q} -p.s., et donc P -p.s.. Mais

$$\mathbf{1}_B = M\mathbf{1}_B \geq \omega_0^+ \mathbf{1}_{\theta B} \quad , P\text{-p.s..}$$

On peut conclure que $\mathbf{1}_B \geq \mathbf{1}_{\theta B}$, P -p.p. et alors $E_P(\mathbf{1}_B) = E_P(\mathbf{1}_{\theta B})$ implique que $\mathbf{1}_B = \mathbf{1}_{\theta B}$, P -p.p.. Mais par l'ergodicité de P , $P(B) \in \{0, 1\}$ et alors $\bar{Q}(B) \in \{0, 1\}$. Comme $\bar{Q}(A) = E_{\bar{Q}}\varphi(\omega)$, on peut conclure \square

4.2 Preuve alternative du théorème 8

Démonstration. Maintenant que nous avons les outils techniques nous nous mettons à prouver encore une fois le théorème concernant la vitesse de la marche. La preuve consiste à couper la variable aléatoire en deux parties et les analyser séparément : une partie martingale et une partie dérive. On définit $d(x, \omega) = E_\omega^x(X_1 - x)$. Alors

$$\begin{aligned} X_n &= \sum_{i=1}^n (X_i - X_{i-1}) = \sum_{i=1}^n (X_i - X_{i-1} - d(X_{i-1}, \omega)) + \sum_{i=1}^n d(X_{i-1}, \omega) \\ &\stackrel{\text{déf}}{=} M_n + \sum_{i=1}^n d(X_{i-1}, \omega) \end{aligned}$$

Regardons d'abord le processus M_n . Sous P_ω^0 c'est une martingale par rapport à la filtration $\mathcal{G}_n = \sigma(M_1, \dots, M_n)$. Démontrons que sa moyenne $\frac{M_n}{n}$ tend vers 0 \mathbb{P}^0 - p.s.. Il découle directement de la définition que $|M_{n+1} - M_n| \leq 2$. Soit $\lambda \in \mathbb{R}$. On obtient alors

$$\begin{aligned} E_\omega^0(e^{\lambda M_n}) &= E_\omega^0(e^{\lambda M_{n-1}} E_\omega^0(e^{\lambda(M_n - M_{n-1})} | \mathcal{G}_{n-1})) \\ &\leq E_\omega^0\left(e^{\lambda M_{n-1}} E_\omega^0\left(\frac{(M_n - M_{n-1}) + 2}{4} e^{2\lambda} + \frac{2 - (M_n - M_{n-1})}{4} e^{-2\lambda} | \mathcal{G}_{n-1}\right)\right) \\ &= E_\omega^0\left(e^{\lambda M_{n-1}} \left(\frac{e^{2\lambda} + e^{-2\lambda}}{2}\right)\right) \\ &\leq E_\omega^0\left(e^{\lambda M_{n-1}} e^{2\lambda^2}\right) \end{aligned}$$

et par récurrence $E_\omega^0(e^{\lambda M_n}) \leq e^{2n\lambda^2}$. On peut donc calculer

$$\begin{aligned} P_\omega^0\left(\left|\frac{M_n}{n}\right| \geq \varepsilon\right) &= P_\omega^0(M_n \geq \varepsilon n) + P_\omega^0(-M_n \geq \varepsilon n) \\ &= P_\omega^0(e^{\lambda M_n} \geq e^{\lambda \varepsilon n}) + P_\omega^0(e^{-\lambda M_n} \geq e^{\lambda \varepsilon n}) \end{aligned}$$

L'inégalité de Markov implique

$$P_\omega^0 \left(\left| \frac{M_n}{n} \right| \leq \varepsilon \right) \leq e^{-\lambda \varepsilon n} (E(e^{\lambda M_n}) + E(e^{-\lambda M_n}))$$

$$\leq 2e^{n(-\varepsilon \lambda + 2\lambda^2)}$$

On pose $\lambda = \frac{\varepsilon}{4}$ et on obtient finalement

$$P_\omega^0 \left(\left| \frac{M_n}{n} \right| \geq \varepsilon \right) \leq 2e^{-n \frac{\varepsilon^2}{8}}.$$

Si on pose alors $\varepsilon = n^{-1/4}$ on voit que

$$\sum_n P_\omega^0 \left(\left| \frac{M_n}{n} \right| \geq n^{-1/4} \right) < \infty.$$

Avec le lemme de Borel-Cantelli on peut conclure que

$$\frac{M_n}{n} \rightarrow 0 \quad , \mathbb{P}^0\text{-p.s.}$$

Pour calculer la vitesse de la marche il suffit donc de calculer le comportement de la dérive en utilisant les résultats concernant les chaînes auxiliaires. On peut écrire

$$\sum_{i=1}^n d(X_{i-1}, \omega) = \sum_{i=1}^n d(0, \bar{\omega}(i-1))$$

Comme le processus $\bar{\omega}(n)$ est ergodique sous la mesure $\bar{Q} \otimes P_\omega^0$ on peut appliquer le théorème de Birkhoff et obtient

$$\frac{1}{n} \sum_{i=1}^n d(0, \bar{\omega}(i-1)) \rightarrow E_{\bar{Q}}(d(0, \bar{\omega}(0))) \quad \bar{Q} \otimes P_\omega^0\text{-p.s.}$$

Calculons cette espérance : Dans un milieu fixé ω la quantité $d(0, \omega)$ vaut

$$d(0, \omega) = E_\omega^x(X_1) = \omega_0^+ - \omega_0^-$$

Et on obtient à la fin

$$E_{\bar{Q}}(d(0, \bar{\omega}(0)))$$

$$= \frac{E_P[\Lambda(\omega)(\omega_0^+ - \omega_0^-)]}{E_P(\Lambda(\omega))} \quad (\text{comme } \frac{dQ}{dP} = \Lambda)$$

$$= \frac{E_P\left(\frac{1}{\omega_0^+} [1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \rho_j](\omega_0^+ - \omega_0^-)\right)}{E_P(\Lambda(\omega))}$$

$$= \frac{1 + E_P(\rho_1 [1 + \sum_{i=2}^{\infty} \prod_{j=2}^i \rho_j]) - E_P(\rho_0 [1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \rho_j])}{E_P(\Lambda(\omega))}$$

$$= \frac{1}{E_P(\Lambda(\omega))} = \frac{1 - E_P(\rho_0)}{E_P\left(\frac{1}{\omega_0^+}\right)}$$

C'est le résultat souhaité. □

Chapitre 5

Comparaison avec les marches aléatoires simples

5.1 Différence conceptuelle

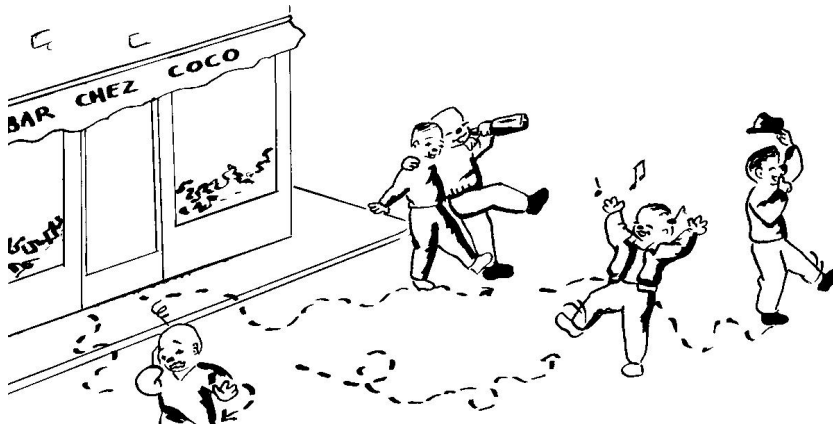


FIG. 5.1 – marche aléatoire simple

Toute la différence entre une MAMA et une marche aléatoire simple (MAS) consiste en la notion de milieu. Dans une MAS, le milieu est homogène : le comportement du point est le même, où qu'il se trouve. Dans une MAMA, en revanche, le milieu n'est pas homogène, et comporte des irrégularités qui influent sur le mouvement du point. On pourrait croire que c'est le non déterminisme du milieu qui crée la différence entre les deux types de marche : en effet, dans une MAMA, le milieu n'est déterminé que par rapport à la probabilité P , et donc le comportement de la MAMA peut *a priori* dépendre du milieu. En fait, il n'en est rien, du moins dans les cas que nous avons étudiés, puisque les propriétés que nous avons énoncées sont vraies \mathbb{P}^0 -p.s., c.à.d. vraies P_ω^0 -p.s. pour P -presque tout milieu ω . Il faut comprendre que, si les différents milieux ω sont certes distincts, leurs propriétés globales sont en revanche identiques : en fait, le tirage au sort du milieu selon la probabilité P ne traduit pas tant le fait qu'on veuille étudier ce qui se passe dans des milieux différents, que

celui qu'on ne sache pas décrire précisément le milieu, juste ses propriétés statistiques : le milieu dans lequel on se place est un milieu « normal » par rapport à la loi de P , qui dit que les ω_z (càd. le comportement de chaque point du milieu vis-à-vis du mobile) sont indépendants et de même loi.

Pour rendre plus pertinente la comparaison entre MAS et MAMA, on associe à chaque MAMA son pendant simple *via* le procédé suivant :

Définition 4. Soit une marche aléatoire en milieu aléatoire, déterminée par la loi de ω_0 . Appelons *marche aléatoire simple associée* à cette MAMA la MAS déterminée par

$$\omega_z = (E_P(\omega_0^-), E_P(\omega_0^0), E_P(\omega_0^+)) \quad \forall z \in \mathbb{Z}$$

Remarque 10. En fait, on peut définir cette MAS de manière très similaire à la MAMA. Si on imagine en effet que les variations du milieu se font non pas dans l'espace mais dans le temps, en suivant exactement la même démarche que dans le paragraphe 1.2, on arrive à la définition suivante :

Définition 5. – On définit un *milieu temporel* $\omega^* = (\omega_n^*)_{n \in \mathbb{N}}$ dont les composantes sont indépendantes et de même loi, chacune suivant la loi de ω_0 . Soit P^* la loi de ω^* .

– Pour un milieu ω^* donné, le comportement de la suite $(X_n)_{n \in \mathbb{N}}$ est alors déterminé par la règle : $P_{\omega^*}^v(X_0 = v) = 1$ et $P_{\omega^*}^v(X_{n+1} = y | X_n = z) = \omega_n^*(y - z)$.

– Puis on définit la loi générale de la marche par $\mathbb{P}_*^v(G) \stackrel{\text{déf}}{=} \int_{\Omega^*} P_{\omega^*}^v(G) P^*(d\omega^*)$.

La définition 5 redonne exactement la même marche que la définition 4, ce qui montre sa naturalité. L'intérêt est de montrer que c'est bien la notion de milieu, càd. d'irrégularité *spatiale*, qui est à l'origine des différences entre les deux types de marches.

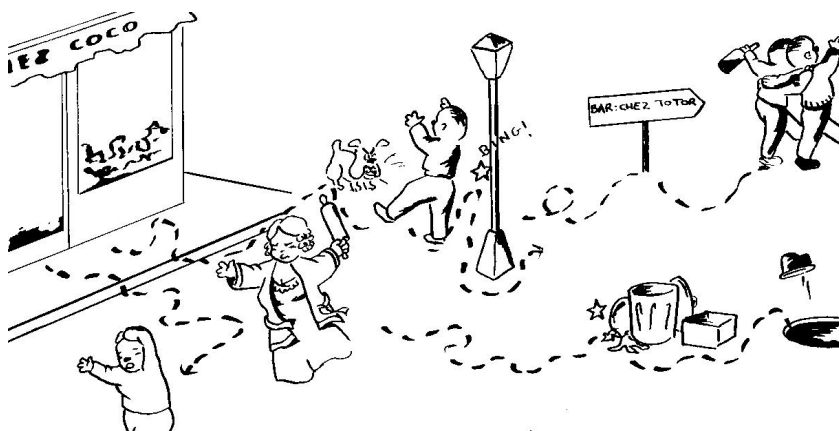


FIG. 5.2 – marche aléatoire en milieu aléatoire

5.2 Comparaison des théorèmes

La loi des grands nombre et le théorème central limite nous permettent de décrire le comportement des MAS :

Théorème 15. *Soit $(Y_n)_{n \in \mathbb{N}}$ la MAS associée à une MAMA (X_n) . Alors :*

1. *Si $E_P(\omega_0^+) > E_P(\omega_0^-)$, la marche (Y_n) tend vers $+\infty$ \mathbb{P}_*^0 -p.s., avec une vitesse de dérive $E_P(\omega_0^+) - E_P(\omega_0^-)$.*
2. *Si $E_P(\omega_0^+) < E_P(\omega_0^-)$, la marche (Y_n) tend vers $-\infty$ \mathbb{P}_*^0 -p.s., avec une vitesse de dérive $E_P(\omega_0^+) - E_P(\omega_0^-)$.*
3. *Si $E_P(\omega_0^+) = E_P(\omega_0^-)$, ($\limsup_{n \rightarrow \infty} Y_n = +\infty$ et $\liminf_{n \rightarrow \infty} Y_n = -\infty$) \mathbb{P}_*^0 -p.s., avec $\frac{Y_n}{n} \rightarrow_{n \rightarrow \infty} 0$.*

Plusieurs observations s'imposent :

1. Si on a pour une MAMA $P(\omega_0 = (p^-, p^0, p^+)) = 1$, alors la MAMA est une marche aléatoire simple, la même que la MAS associée. On vérifie que dans ce cas les théorèmes des chapitres 2 et 3 donnent bien les mêmes résultats que le théorème 15.
2. Contrairement à ce que l'intuition affirmerait, une MAMA peut partir dans une direction opposée à la MAS associée. Prenons ainsi le cas $P(\omega_0 = (0,4; 0; 0,6)) = 0,76$ et $P(\omega_0 = (0,8; 0; 0,2)) = 0,24$. Alors on a $E_P(\omega_0^+) - E_P(\omega_0^-) = 0,008 > 0$, donc la MAS part à droite ; et pourtant $E_P(\ln \rho_0) = 0,024... > 0$, donc la MAMA part à gauche.
3. Nous avons observé dans la remarque 7 qu'une MAMA pouvait dériver à l'infini de plus en plus lentement, avec une vitesse de dérive asymptotique nulle. On voit que ce n'est pas le cas pour les MAS.
4. En vertu de la remarque 6, on s'aperçoit qu'une MAMA est toujours moins rapide (en valeur absolue) que la MAS associée. Au passage, noter aussi que la remarque 6 exclut que la MAMA ait une vitesse de dérive de signe contraire à la MAS associée, ce qui modère le point 2 ci-dessus¹.

Les marches aléatoires en milieu aléatoire ont donc un comportement bien spécifique par rapport aux marches aléatoires simples. Nous allons essayer d'expliquer certaines de ces particularités, en particulier les observations 3 et 4 ci-dessus, à l'aide de la notion de puits de potentiel.

5.3 Une interprétation : notion de puits

5.3.1 Profil de potentiel

Pour bien comprendre les MAMA, il faut voir comment le milieu influe sur le comportement du mobile. Certains points ont plutôt tendance à envoyer le

¹On vérifie bien dans l'exemple proposé que $E_P(\rho_0) = 1,2 \geq 1$ et $E_P(\frac{1}{\rho_0}) = 1,466... \geq 1$; la vitesse de dérive de la MAMA est donc nulle.

mobile vers la droite, d'autres vers la gauche : le théorème 6 nous suggère que c'est la grandeur $\ln(\rho_z)$ qui caractérise bien l'attitude du point z vis-à-vis du mobile. À partir de là, on peut représenter le milieu par un profil de potentiel, où chaque point z correspond à une pente $\ln(\rho_z)$: le mobile aura ainsi tendance à descendre vers les potentiels les plus faibles.



FIG. 5.3 – Un exemple de profil de potentiel : on a pris $P(\omega_0^+ \in [x, x + dx]) = (\frac{1}{2} + x)dx$, et $\omega_0^0 = 0$ p.s.. Sur cette figure 81 points sont représentés.

5.3.2 Effet des puits

On appelle *puits* une région du profil en forme de gorge : dans la partie gauche de cette région, le mobile a tendance à être envoyé à droite, et vice-versa. On comprend donc que le mobile va rester piégé dans le puits, jusqu'à ce que le hasard de sa marche le fasse aller à contre-courant de la pente suffisamment longtemps pour qu'il en sorte. Formalisons cela en termes mathématiques :

Supposons qu'on ait deux points consécutifs z_{F-} et z_{F+} du milieu tels que :

$$\begin{cases} \text{pour } z_{F-} - L_- < z \leq z_{F-} & \ln(\rho_z) < 0 \\ \text{pour } z_{F+} \leq z < z_{F+} + L_+ & \ln(\rho_z) > 0 \end{cases}$$

où L_- et L_+ notent des entiers strictement positifs. C'est la description d'un puits de fond $\frac{z_{F-} + z_{F+}}{2}$, de longueur L_- à gauche, et L_+ à droite. Imaginons qu'à un moment donné, le mobile soit au point z_{F-} . Alors, par la formule (2.2), la probabilité que le mobile sorte du puits par la gauche avant qu'il n'atteigne le point z_{F+} est :

$$\nu(z_{F-} - L_-, z_{F-}, z_{F+}) = \frac{1}{1 + \sum_{i=z_{F-}-L_-+1}^{z_{F-}} e^{-\sum_{j=i}^{z_{F-}} \ln(\rho_j)}}$$

Or, cette valeur est « très petite ». Plus précisément, soit $-\alpha_-$ la moyenne des $\ln(\rho_z)$ qui sont négatifs. Le terme dominant du dénominateur est $e^{-\sum_{j=z_{F-}-L_-+1}^{z_{F-}} \ln(\rho_j)}$, et comme les $\ln \rho_j$ sont de l'ordre de $-\alpha_-$, la probabilité que le mobile sorte du puits par la gauche est de l'ordre de $e^{-\alpha_- L^-}$. En

particulier, si L_- est suffisamment grand, $e^{-\alpha_- L_-} \ll \frac{1}{L_- + 1}$, qui représente la probabilité du même événement dans un profil plat.

De même, une fois que le mobile se sera retrouvé en z_{F+} , il aura une probabilité beaucoup plus forte de revenir en z_{F-} que de sortir du puits par la droite, la probabilité de sortie étant de l'ordre de $e^{-\alpha_+ L^+}$, où α_+ est la moyenne des $\ln(\rho_z)$ qui sont positifs. D'où un mouvement de navette du mobile au fond du puits, qui dure un temps de l'ordre de $e^{\min(\alpha_- L_-, \alpha_+ L^+)}$. Par conséquent, il est clair que l'existence de puits, en piégeant le mobile, va ralentir la marche. Cela explique qu'on ait une vitesse de dérive plus faible dans le cas d'une MAMA que dans celui d'une MAS.

5.3.3 La mesure \bar{Q}

On peut voir ce phénomène d'une autre façon : imaginons une MAMA qui dérive vers la droite. Appelons *points-descentes* les points qui ont tendance à envoyer le mobile vers la droite, et *points-côtes* ceux qui ont tendance à le renvoyer vers la gauche. Si p_c est la proportion de points-côtes, on pourrait s'attendre à ce que le mobile se retrouve avec une probabilité $1 - p_c$ sur un point-descente et une probabilité p_c sur un point-côte au cours de son évolution. Mais cela n'est pas vrai : en effet, puisque l'évolution du mobile se fait de gauche à droite, un point-descente, qui a tendance à laisser passer le mobile vers la droite, ne sera souvent visité qu'une fois. Au contraire, les points-côtes vont généralement renvoyer le mobile en arrière, et celui-ci reviendra pour être renvoyé à nouveau jusqu'à ce qu'il franchisse la côte. Finalement, le milieu que voit le mobile, c.à.d. celui qui est représenté par la probabilité \bar{Q} (cf remarque 9), surreprésente les points-côtes au niveau du mobile par rapport à la probabilité P . Le mobile va donc plus lentement que s'il était dans un milieu homogène, c.à.d. que dans le cas d'une marche aléatoire simple.

Plus généralement, puisque le mobile a tendance à rester piégé au fond des puits, \bar{Q} surreprésente les milieux où le mobile est au fond d'un puits. C'est cette déformation du milieu vu par le point par rapport au milieu vu de l'extérieur qui est à l'origine des différences de comportement entre MAMA et MAS.

5.3.4 Cas transitoire à vitesse nulle

La notion de puits permet aussi de comprendre les cas de dérive à l'infini avec vitesse nulle. Comme nous l'avons vu ci-dessus, plus un puits est profond et plus il ralentit la marche. Plaçons-nous dans un cas où le mobile dérive vers la droite : cela signifie que le profil est globalement en pente (de gauche à droite), et par conséquent l'existence de puits très profonds est improbable, puisqu'elle suppose que le profil soit en montée sur une longue distance. Néanmoins, au cours de l'évolution du mobile, celui-ci va s'aventurer de plus en plus loin et finira donc par rencontrer tous les types de profils locaux, y compris les plus improbables. On s'attend donc à ce que le mobile soit piégé

de plus en plus longtemps dans des puits de plus en plus profonds à mesure de son évolution.

À partir de là, deux cas sont possibles :

- Soit les puits les plus profonds arrivent assez vite pour faire tendre la vitesse vers zéro à l’infini,
- Soit les grands puits sont suffisamment rares pour que le mobile les traverse en conservant une vitesse globale de dérive.

C’est donc la structure des puits qui détermine l’existence de deux cas : une vitesse de dérive globale non nulle lorsqu’il y a peu de puits très profonds, ou une marche de plus en plus lente lorsqu’il y en a beaucoup.

5.3.5 Approche heuristique du théorème 8

Avertissement Cette partie vise simplement à une interprétation intuitive du théorème 8, elle ne prétend pas à la rigueur mathématique.

Nous allons essayer de déterminer heuristiquement le critère du théorème 8 qui détermine si la vitesse de dérive asymptotique est nulle ou non. Plaçons-nous dans le cas où le mobile part vers $+\infty$, et soit p_c la proportion de points-côtes. Cherchons le temps que le mobile met à se rendre au point N : supposons que ce soient les puits de longueur L qui influencent le plus le mouvement du point. La probabilité d’avoir un puits de longueur L est essentiellement limitée par la probabilité d’avoir L points-côtes consécutifs, qui est de p_c^L . Il y a donc environ Np_c^L puits de longueur L entre 0 et N , et chacun de ces puits bloque le mobile pendant un temps de l’ordre de $e^{\alpha L}$, donc finalement le mobile va mettre un temps $N(e^{\alpha}p_c)^L$ à atteindre le point N .

Alors, de deux choses l’une :

- Soit $e^{\alpha}p_c > 1$ et dans ce cas, plus les puits sont profonds et plus ils freinent le mobile efficacement. Au cours de l’évolution du mobile, les puits qu’il faudra prendre en compte seront donc de plus en plus profonds, et le mobile ira de plus en plus lentement, avec une vitesse asymptotique nulle puisque $(e^{\alpha}p_c)^L \rightarrow_{n \rightarrow \infty} +\infty$.
- Soit $e^{\alpha}p_c \leq 1$ et alors les puits les plus profonds ne sont pas plus efficaces que les petits puits, car ils sont trop rares. Donc la vitesse asymptotique de dérive du mobile est non nulle.

Pour savoir dans quel cas on se trouve, il faut donc évaluer $e^{\alpha}p_c$. Or α_+ est l’ordre de grandeur des $\ln \rho_z$ qui sont positifs, donc e^{α_+} est l’ordre de grandeur des ρ_z qui sont plus grands que 1. Comme p_c est la proportion de ces mêmes ρ_z ,

$$e^{\alpha_+}p_c \simeq E_P[\mathbb{1}_{\{\rho_0 > 1\}}\rho_0]$$

On est finalement arrivé à dire que la marche était freinée à l’infini pour $E_P[\mathbb{1}_{\{\rho_0 > 1\}}\rho_0] \gtrsim 1$. D’après le théorème 8, le véritable résultat est $E_P[\rho_0] > 1$. Notre conclusion heuristique n’est donc pas très éloignée de la réalité, ce qui montre que la notion de puits est une bonne interprétation du comportement de la MAMA.

5.4 Exemples numériques

Dans ce chapitre nous comparons les résultats théoriques avec des observations en expérience numérique.

5.4.1 Vitesse dans un milieu simple

On considère un milieu avec seulement deux types de points.

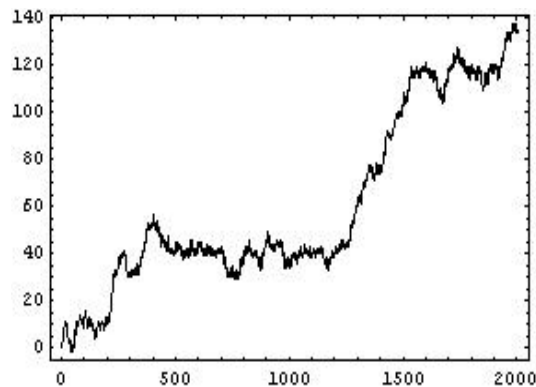
$$P(\omega_0^+ = 1 - \omega_0^- = 0,7) = 1 - P(\omega_0^+ = 1 - \omega_0^- = 0,5) = 0,5$$

D'après les résultats théoriques nous nous attendons à une vitesse $v_P = \frac{1}{6}$. Dans l'expérience numérique nous obtenons pour $n = 2000$ les valeurs suivantes pour X_n

320	418	290	306
364	442	290	328
300	382	356	370
342	336	296	348
322	248	292	318
329,6	365,2	304,8	334,0

*Les valeurs dans une colonne viennent du même milieu
La dernière ligne comporte les valeurs moyennes des milieux*

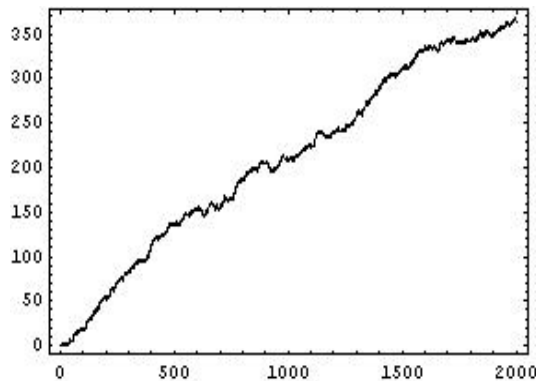
On obtient comme valeur moyenne $\overline{X_{2000}} = 337,4$. La valeur théorique est 333,3.



MAMA pour $n = 2000$

Même dans ce cadre sans grands puits on voit clairement des régions où la marche avance plus lentement qu'ailleurs. Une expérience avec la marche simple associée nous montre que la marche simple avance plus vite. Quelques valeurs pour $n = 2000$

390	466	492	384	324
376	406	348	478	372



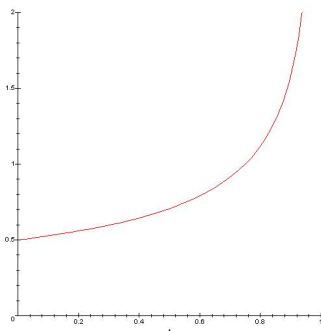
MAS pour $n = 2000$

On obtient comme valeur moyenne $\overline{X_{2000}} = 403,6$. La valeur théorique est 400.

5.4.2 Le cas à vitesse nulle

Nous regardons une mesure P à densité. Posons $P\{\omega_0^0 = 0\} = 1$, et

$$P\{\omega_0^+ < t\} = 1 - \sqrt{1-t}$$



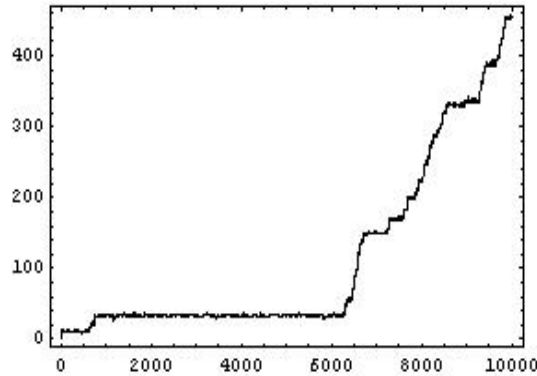
$$t \mapsto \frac{d}{dt}P(\omega_0^+ \leq t)$$

On est bien dans le cas transitoire, car $E[\ln \rho] = -\ln(4) < 0$, mais on a $E[\rho] = E[\rho^{-1}] = +\infty$ donc la vitesse est nulle. Quelques valeurs pour $n = 10000$

49	437	455
37	433	151
19	555	35
31	577	35
5	505	35
39	505	455
49	433	509
49	533	35
34,8	497,3	209,4

Les valeurs dans une colonne viennent du même milieu.

On peut bien observer le phénomène dans l'exemple suivant

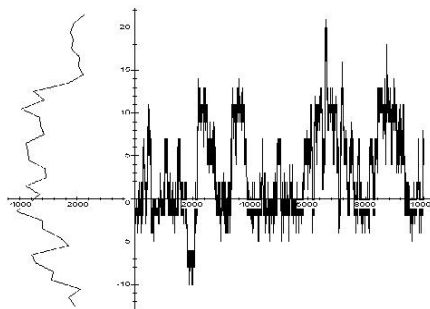


MAMA dans le milieu de la troisième colonne

Il y a un grand obstacle à $X = 35$ ($\omega_{35}^+ = 0.017$ et $\omega_{37}^+ = 0.065$). Dès que la marche a passé cet obstacle elle avance vite mais rencontre un autre obstacle vers $X = 133$ etc. Les obstacles sont extrêmement rares mais quand ils apparaissent ils ralentissent la marche tellement qu'ils rendent l'accroissement sous-linéaire.

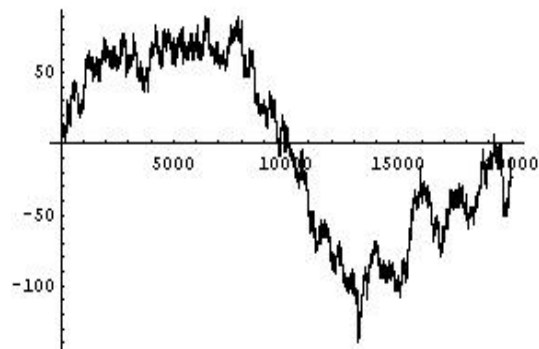
5.4.3 Le cas récurrent

Regardons maintenant le cas récurrent. Supposons que ω^+ suit une loi uniforme sur $[0, 1]$ et $\omega^- = 1 - \omega^+$. Dans ce cadre la marche rencontre très souvent des obstacles et quand elle sort d'un puits elle retombe presque instantanément dans un autre. Le mouvement devient presque statique.



MAMA dans un milieu récurrent pour $n=10000$
À gauche, le profil de potentiel

Si on compare avec la marche simple :



MAS récurrente pour $n=20000$

En fait un théorème de Sinai (cf. [3] page 406) dit que la distance caractéristique de la MAMA en milieu récurrent croît comme $(\ln(n))^2$ (càd $\frac{X_n}{(\ln(n))^2} \xrightarrow{\text{proba}} C(\omega)$) pendant que la MAS croît comme \sqrt{n} d'après le théorème central limite.

Bibliographie

- [1] P. Billingsley, *Ergodic Theory and Information*, John Wiley & Sons, Inc. (1965)
- [2] R. Durrett, *Probability : Theory and example*, 2nd ed., Duxbury Press, Belmont (1996)
- [3] B. Hughes, *Random Walks and Random Environments, Volume 2*, Clarendon Press, Oxford (1996)
- [4] A.S. Sznitman, *Topics in Random Walks in Random Environment*
- [5] O. Zeitouni, *Lecture Notes On Random Walks In Random Environment*, (2003) (disponible sur <http://www-ee.technion.ac.il/~zeitouni/>)

Les figures 5.1 et 5.2 ont été dessinées par Mathilde Peyre.

Document III

Une inégalité de Harnack

Une inégalité de Harnack

Rémi Peyre

6 octobre 2006

1 Introduction

Soit $d \geq 1$ un entier ; on considère l'espace \mathbb{R}^d muni d'un champ de vecteurs \vec{v} localement lipschitzien. On peut alors définir la diffusion brownienne avec dérive \vec{v} par

$$\overrightarrow{dX}_t = \overrightarrow{dB}_t + \vec{v}(X_t) dt, \quad (1)$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien de \mathbb{R}^d , de constante de diffusion σ , c'est-à-dire que pour $t \leq u$:

$$\overrightarrow{B}_t \overrightarrow{B}_u \stackrel{(\text{loi})}{\sim} \mathcal{N} \left(\vec{0}, \sigma(u-t) \mathbf{I}_d \right). \quad (2)$$

L'équation stochastique (1) induit un processus Markovien que nous appellerons *diffusion dérivante* ; on notera $(\mathbb{P}_x)_{x \in \mathbb{R}^d}$ les lois de ce processus.

Tout au long de ce mémoire nous faisons l'hypothèse que le processus \mathbb{P} est récurrent positif au sens de Harris ; \mathbb{P} définit alors une unique probabilité d'équilibre μ . On peut montrer que μ est équivalente à la mesure de Lebesgue : elle s'écrit

$$d\mu = f dx, \quad (3)$$

où $f > 0$ est caractérisée par l'équation différentielle (à prendre au sens des distributions) :

$$\frac{\sigma}{2} \Delta f - \operatorname{div}(f \vec{v}) = 0. \quad (4)$$

Remarque. L'existence de f découle de la théorie générale des équations uniformément elliptiques, pour laquelle on pourra par exemple se référer à [1]. Néanmoins, dans la mesure où notre approche vise précisément à contourner ladite théorie générale, nous nous contenterons ici de partir de l'existence de la mesure μ ; l'existence de la densité f par rapport à la mesure de Lebesgue, en fait, découlera du travail ultérieur.

L'inégalité qui fait l'objet de ce mémoire donne un critère de régularité pour f en fonction de la régularité locale de \vec{v} . Plus précisément, notre énoncé consiste en le

Théorème 1. Soit $o \in \mathbb{R}^d$; on note $V = |\vec{v}(o)|$. On suppose qu'il existe un réel $R > 0$ et une constante $k \geq 0$ tels que, sur $B(o, R)$, \vec{v} soit k -lipschitzienne. Alors $\ln f$ est localement lipschitzienne en o ⁽¹⁾, avec une constante de Lipschitz k' qui peut s'exprimer uniquement en fonction de V , R et k .

Ce genre de théorème est classique, cf. p. ex. [2], pp. 183 et 235, mais ce mémoire vise à le démontrer par une approche différente de celles auxquelles nous sommes habituées. L'enjeu est de faire apparaître une structure probabiliste de couplage, qui exprimera en quelque sorte que « cela coûte pratiquement le même prix d'aller voir quelqu'un que d'aller voir son voisin ».

2 Notations

Introduisons tout d'abord quelques notations en vigueur dans ce mémoire :

- On raisonnera à l'aide de grandeurs physiquement homogènes, ce qui explique l'introduction de la constante de diffusion σ définie en (2).
- Les vecteurs seront notés \vec{a} , \vec{b} , ... et les matrices A , B , ...
- Une probabilité \mathbb{P} étant donnée, l'espérance associée sera indiquée à l'aide de crochets :


$$\mathbb{P}[f] = \int f(\omega) d\mathbb{P}(\omega). \quad (5)$$

- Pour $\vec{u} \in \mathbb{R}^d$ un vecteur, $|\vec{u}|$ notera la norme euclidienne de \vec{u} .
- La mesure de Lebesgue sur \mathbb{R} ou \mathbb{R}^d sera notée dx ; pour $A \subset \mathbb{R}$, resp. \mathbb{R}^d , on notera $|A| = \int_A dx$.
- Pour $x \in \mathbb{R}^d$, $r > 0$, la boule ouverte de rayon r centrée sur x sera notée $B(x, r)$; $B(o, r)$ sera simplement notée B_r .
- Le temps d'atteinte d'un fermé F par un processus aléatoire (continu) sera noté τ_F .
- Le complémentaire d'un ouvert U de \mathbb{R}^d sera noté cU .
- La différence symétrique de deux événements sera notée par Δ :

$$A \Delta B = (A \setminus B) \cup (B \setminus A). \quad (6)$$

- Les constantes importantes qui apparaîtront dans ce texte seront notées K_1, K_2, \dots . Les constantes « muettes » ne servant que provisoirement seront, elles, notées c_1, c_2, \dots et renumérotées à chaque énoncé.

3 Rappels sur les semigroupes à densité

 Le lecteur intéressé par un traitement plus complet de la question pourra se référer par exemple à [3].

¹Ou plus précisément : $\ln f$ admet un représentant continu au voisinage de o , qui est lipschitzien.

Définition 2. Soit $(P_x)_{x \in \mathbb{R}^d}$ un processus sous-markovien à valeurs dans \mathbb{R}^d . On dit que le semigroupe $(P^t)_{t \geq 0}$ engendré par le processus (ou le processus lui-même) est à densité par rapport à la mesure de Lebesgue s'il existe une fonction mesurable $p :]0, \infty[\times (\mathbb{R}^d)^2 \rightarrow [0, \infty[$ vérifiant :

$$\forall t > 0 \quad \forall x \in \mathbb{R}^d \quad \forall A \subset \mathbb{R}^d \text{ mesurable} \quad P_x(X_t \in A) = \int_A p^t(x, y) dy. \quad (7)$$

p est alors appelée la densité de transition du semigroupe $(P^t)_{t \geq 0}$.

Quelques propriétés immédiates des densités de transition :

Proposition 3.

1. Pour tous $t > 0$ et $x \in \mathbb{R}^d$, la fonction $p^t(x, \cdot)$ est unique modulo l'identité y -p.p.
2. Pour tous $t > 0$ et $x \in \mathbb{R}^d$, on a

$$\int_{\mathbb{R}^d} p^t(x, y) dy \leq 1 \quad (8)$$

3. Pour tous $t, u > 0$, $x, z \in \mathbb{R}^d$, p vérifie la relation de Chapman-Kolmogorov :

$$p^{t+u}(x, z) = \int_{\mathbb{R}^d} p^t(x, y) p^u(y, z) dy. \quad (9)$$

On peut alors définir les fonctions de Green :

Définition 4. Soit $(X_t)_{t \geq 0}$ un processus sous-markovien continu, à densité sur \mathbb{R}^d ; soit D un domaine de \mathbb{R}^d . Soit ∂ un point-cimetière; on note X^* le processus X tué en sortant de D , défini par

$$X_t^* = \begin{cases} X_t & \text{si } t < \tau_{cD} ; \\ \partial & \text{sinon.} \end{cases} \quad (10)$$

Il est clair qu'alors $(X_t^*)_{t \geq 0}$ est également un processus sous-markovien à densité; soit p_* la densité de transition associée. Alors la fonction $G_D : (\mathbb{R}^d)^2 \rightarrow [0, \infty]$ définie par

$$G_D(x, y) = \int_0^\infty p_*^t(x, y) dt \quad (11)$$

est appelée la fonction de Green associée au processus X et au domaine D .

De nouveau, $G_D(x, \cdot)$ est unique y -p.p. L'interprétation probabiliste de G_D est donnée par :

$$\forall x \in \mathbb{R}^d \quad \forall A \subset \mathbb{R}^d \text{ mesurable} \quad \int_A G_D(x, y) dy = P_x \left[\int_0^{\tau_{cD}} \mathbb{1}_{X_t \in A} dt \right]. \quad (12)$$

☞ Dans la suite de ce mémoire, nous montrerons que la diffusion dérivante est un processus à densité. Nous notons dès à présent sa densité de transition par p . Par ailleurs, on se fixe $R_2 < R$; la loi de la diffusion dérivante tuée en sortant de B_{R_2} est notée \mathbb{P}^* et sa densité de transition est notée p_* . La fonction de Green associée à la diffusion dérivante et à la boule B_{R_2} est notée $G_{B_{R_2}}$, ou simplement G .

4 Noyau de l'équation stationnaire

La proposition qui est au cœur de notre approche est la suivante :

Proposition 5. *On se fixe $R_1 \in]0, R_2[$. Alors il existe une mesure non nulle θ sur ∂B_{R_1} telle qu'on ait pour tout y de B_{R_1} :*

$$f(y) = \int_{\partial B_{R_1}} G_{B_{R_2}}(x, y) d\theta(x). \quad (13)$$

Preuve. L'idée consiste à s'intéresser aux excursions de la particule X entre les moments où elle rentre dans $\overline{B_{R_1}}$ et ceux où elle sort de B_{R_2} . Concrètement, on définit :

$$\begin{cases} T_0^* = \inf\{t \geq 0; X_t \notin B_{R_2}\}; \\ T_1 = \inf\{t > T_0^*; X_t \in \overline{B_{R_1}}\}; \\ T'_n = \inf\{t > T_n; X_t \notin B_{R_2}\}; \\ T_{n+1} = \inf\{t > T'_n; X_t \in \overline{B_{R_1}}\}. \end{cases} \quad (14)$$

$[T_i, T'_i]$ est alors la i -ème excursion; on note $\mathcal{T} = \{T_i; i \geq 1\}$ l'ensemble des temps de départs d'excursions.

Puisque μ est une mesure d'équilibre, le processus de loi \mathbb{P}_μ est stationnaire. Par conséquent on peut écrire, pour $A \subset B_{R_1}$ mesurable, $c_1 \geq 0$ un temps quelconque :

$$\begin{aligned} \mu(A) &= \mathbb{P}_\mu(X_{c_1} \in A) \\ &= \sum_{i=1}^{\infty} \mathbb{P}_\mu(c_1 \in]T_i, T'_i[\text{ et } X_{c_1} \in A) + \mathbb{P}_\mu(c_1 < T_0^* \text{ et } X_{c_1} \in A). \end{aligned} \quad (15)$$

Comme \mathbb{P}_μ -p.s. on a $T_0^* < \infty$ (par exemple par la proposition 18 *infra*), il s'ensuit que

$$\mu(A) = \lim_{c_1 \rightarrow \infty} \sum_{i=1}^{\infty} \mathbb{P}_\mu(c_1 \in]T_i, T'_i[\text{ et } X_{c_1} \in A). \quad (16)$$

Étudions maintenant l'expression dans la limite. Les T_i sont des temps

d'arrêt, donc on a par la propriété de Markov :

$$\begin{aligned}
& \sum_{i=1}^{\infty} \mathbb{P}_{\mu}(c_1 \in]T_i, T'_i[\text{ et } X_{c_1} \in A) \\
&= \sum_{i=1}^{\infty} \mathbb{P}_{\mu} \left[\mathbb{1}_{T_i < c_1} \mathbb{P}_{X_{T_i}}(c_1 - T_i < \tau_{c_{B_{R_2}}} \text{ et } X_{c_1 - T_i} \in A) \right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}_{\mu} \left[\mathbb{1}_{T_i < c_1} \int_A p_*^{c_1 - T_i}(X_{T_i}, y) dy \right] \\
&= \int_{]0, c_1[\times \partial B_{R_1}} d\mathbb{P}_{\mu} ((\exists i \geq 1)(T_i = t \text{ et } X_{T_i} = x)) \left(\int_A p_*^{c_1 - t}(x, y) dy \right) \\
&\stackrel{\text{(Fubini)}}{=} \int_A dy \int_{]0, c_1[\times \partial B_{R_1}} p_*^{c_1 - t}(x, y) d\mathbb{P}_{\mu} ((\exists i \geq 1)(T_i = t \text{ et } X_{T_i} = x)) \\
&\stackrel{u=c_1-t}{=} \int_A dy \int_{]0, \infty[\times \partial B_{R_1}} p_*^u(x, y) d\mathbb{P}_{\mu} ((\exists i \geq 1)(T_i = c_1 - t \text{ et } X_{T_i} = x)). \quad (17)
\end{aligned}$$

Finalemnt, c_1 n'intervient plus que par la mesure $d\mathbb{P}_{\mu}((\exists i \geq 1)(T_i = c_1 - t \text{ et } X_{T_i} = x))$, que nous noterons $d\nu_{c_1}$ pour alléger. Pour ceux que nos notations laissent perplexes, précisons que $d\nu_{c_1}$ est la mesure sur $]0, \infty[\times \partial B_{R_1}$ définie sur les rectangles mesurables par :

$$\nu_{c_1}(I \times B) = \mathbb{P}_{\mu}[\#\{t \in c_1 - I; t \in \mathcal{T} \text{ et } X_t \in B\}]. \quad (18)$$

Pour étudier ν_{c_1} , nous avons besoin du processus stationnaire bilatère associé au processus stationnaire \mathbb{P}_{μ} . Rappelons qu'il s'agit d'un processus $(X_t)_{t \in \mathbb{R}}$, de loi \mathbb{P}^{st} telle que pour tout $t \in \mathbb{R}$, $(X_{t+u})_{u \geq 0}$ ait pour loi \mathbb{P}_{μ} . La définition de \mathcal{T} peut s'étendre à notre processus bilatère, *via* :

$$t \in \mathcal{T} \Leftrightarrow \left\{ \begin{array}{l} X_t \in \partial B_{R_1} \\ (\forall u < t) \left(((\forall v \in [u, t])(X_v \in B_{R_2})) \Rightarrow ((\forall v \in [u, t])(X_v \notin \overline{B_{R_1}})) \right) \end{array} \right\}. \quad (19)$$

On remarquera que cette définition de \mathcal{T} est invariante par translation.

Les définitions de \mathcal{T} données en (14) et (19) ne sont pas exactement les mêmes, même en restriction à $]0, +\infty[$; convenons donc de noter respectivement par $\mathcal{T}^{>0}$ et $\mathcal{T}^{\mathbb{R}}$ les ensembles définis par (14) et (19). La nuance entre $\mathcal{T}^{\mathbb{R}} \cap]0, +\infty[$ et $\mathcal{T}^{>0}$ est donnée par :

$$t \in \mathcal{T}^{>0} \Leftrightarrow t \in \mathcal{T}^{\mathbb{R}} \text{ et } (\exists u \in [0, t])(X_u \notin B_{R_2}). \quad (20)$$

Dès lors, on écrit pour $B \subset \partial B_{R_1}$ et $t > 0$:

$$\begin{aligned}
\nu_{c_1}(]0, t] \times B) &= \mathbb{P}_{\mu}[\#\{u \in [c_1 - t, c_1[; u \in \mathcal{T}^{>0} \text{ et } X_u \in B\}] \\
&= \mathbb{P}^{\text{st}}[\#\{u \in [c_1 - t, c_1[; u \in \mathcal{T}^{\mathbb{R}} \text{ et } (\exists v \in [0, u])(X_v \notin B_{R_2})\} \text{ et } X_u \in B] \\
&= \mathbb{P}^{\text{st}}[\#\{u \in [-t, 0[; u \in \mathcal{T} \text{ et } (\exists v \in [-c_1, u])(X_v \notin B_{R_2}) \text{ et } X_u \in B\}], \quad (21)
\end{aligned}$$

où la dernière égalité vient par stationnarité.

Or, pour tout $u \in [-t, 0[$, $\mathbb{P}^{\text{st}}((\exists v \in [-c_1, u])(X_v \notin B_{R_2})) \geq \mathbb{P}^{\text{st}}((\exists v \in [-c_1, -t])(X_v \notin B_{R_2})) = \mathbb{P}_\mu((\exists v \in [0, c_1 - t])(X_v \notin B_{R_2})) \xrightarrow{c_1 \rightarrow \infty} \mathbb{P}_\mu((\exists v \in [0, \infty])(X_v \notin B_{R_2})) = 1$, comme nous l'avons déjà signalé *supra* pour la dernière égalité. On en tire finalement :

Avis 6. Soit ν_∞ la mesure sur $]0, \infty[\times \partial B_{R_1}$ définie sur les rectangles mesurables par

$$\nu_\infty(I \times B) = \mathbb{P}^{\text{st}}[\#\{u \in -I; u \in \mathcal{T} \text{ et } X_u \in B\}]. \quad (22)$$

Alors $\nu_{c_1} \xrightarrow{c_1 \rightarrow \infty} \nu_\infty$, au sens où il existe une suite exhaustive $(E_n)_{n \in \mathbb{N}}$ d'ensembles mesurables tels que, pour tout $n \in \mathbb{N}$, $\nu_{c_1}|_{E_n} \xrightarrow{c_1 \rightarrow \infty} \nu_\infty|_{E_n}$ pour tout ensemble mesurable C .

Il s'ensuit par convergence monotone :

$$\mu(A) = \int_A dy \int_{]0, \infty[\times \partial B_{R_1}} p_*^t(x, y) d\nu_\infty(t, x). \quad (23)$$

Finalement, il ne nous reste plus qu'à montrer que $d\nu_\infty$ est de la forme $dt d\theta(x)$ puisqu'alors on aura :

$$\int_A f(y) dy = \mu(A) = \int_A dy \int_0^\infty \int_{\partial B_{R_1}} dt d\theta(x) p_*^t(x, y) \stackrel{\text{(Fubini)}}{=} \int_A \left(\int_{\partial B_{R_1}} G(x, y) d\theta(x) \right) dy, \quad (24)$$

et ce pour tout A , d'où (13) par identification.

Nous allons d'abord nous intéresser, pour $B \subset \partial B_{R_1}$, à la mesure ν_∞^B sur $]0, \infty[$, qu'on définit comme la projection de $\nu_\infty|_{]0, \infty[\times B}$ sur sa première coordonnée. Par stationnarité de X sous \mathbb{P}^{st} , ν_∞^B est invariante par translation. Admettons provisoirement — cela sera prouvé dans quelques instants — que ν_∞^B est finie sur les intervalles bornés. Alors on peut écrire $d\nu_\infty^B = \theta(B) dt$ pour un $\theta(B) < \infty$. Par ailleurs, puisque ν_∞ est une mesure, il s'ensuit que θ est une mesure (finie) sur B et que $d\nu_\infty = dt d\theta(x)$, ce qu'on voulait.

Pour finir, il ne reste plus qu'à vérifier que ν_∞^B est finie sur les intervalles bornés pour tout $B \subset \partial B_{R_1}$, i.e. que $\nu_\infty^{\partial B_{R_1}}$ est finie sur les intervalles bornés. Fixons-nous donc $c_2 > 0$ et majorons $\nu_\infty^{\partial B_{R_1}}(]0, c_2])$. D'abord :

$$\nu_\infty^{\partial B_{R_1}}(]0, c_2]) = \mathbb{P}^{\text{st}}[\#\{\mathcal{T}^{\mathbb{R}} \cap]0, c_2]\}] \leq 1 + \mathbb{P}_\mu[\#\{\mathcal{T}^{>0} \cap]0, c_2]\}]. \quad (25)$$

Ensuite, il est clair qu'il existe une borne $c_3 < 1$ telle que, pour tout $x \in \partial B_{R_1}$, on ait $\mathbb{P}_x(\tau_{cB_{R_2}} \leq c_2) \leq c_3$ ⁽²⁾. En itérant la propriété de Markov, on obtient

²Si le lecteur en doute, qu'il se rassure : de toutes façons cela découlera des résultats du §7.3.

alors que $\mathbb{P}_\mu(T_{i+1} \leq c_2) \leq c_3^i$, d'où $\mathbb{P}_\mu[\#\{\mathcal{T}^{>0} \cap]0, c_2]\} \leq \sum_{i=0}^{\infty} c_3^i < \infty$. \square

Remarque. Dans notre démonstration nous avons volontairement fait l'impasse sur la non-nullité de θ , car celle-ci s'obtient par un raisonnement spécifique aux diffusions uniformément elliptiques. En fait, que θ soit non nulle découle du théorème de support, cf. p. ex. [1], pp. 25-26.

L'outil permettant d'obtenir le théorème 1 à partir de la proposition 5 est le

Lemme 7. Soit E un espace métrique, (Ω, ν) un espace mesuré avec $\nu(\Omega) > 0$, et $g : \Omega \times E \rightarrow]0, \infty[$ une fonction vérifiant pour une certaine constante $k' \geq 0$:

$$\omega\text{-p.p. } x \mapsto \ln g(\omega, x) \text{ est } k'\text{-lipschitzienne.} \quad (26)$$

Alors la fonction $G : E \rightarrow]0, \infty[$ définie par

$$G(x) = \int_{\Omega} g(\omega, x) d\nu(\omega), \quad (27)$$

si elle est finie en au moins un point, est finie partout, et $\ln G$ est k' -lipschitzienne.

Preuve. Soit y un point de E en lequel $G(y) < \infty$. Soit $y' \in E$; notons δ la distance entre y et y' . Alors l'hypothèse (26) implique qu'on a :

$$\omega\text{-p.p. } e^{-k'\delta} g(\omega, y) \leq g(\omega, y') \leq e^{k'\delta} g(\omega, y), \quad (28)$$

d'où par intégration :

$$e^{-k'\delta} G(y) \leq G(y') \leq e^{k'\delta} G(y). \quad (29)$$

Il en résulte, d'une part que $G(y') < \infty$, d'autre part que $|\ln G(y') - \ln G(y)| \leq k'\delta$. Cela suffit pour conclure. \square

Finalement, notre objectif se résume à démontrer la

Proposition 8. Supposons les hypothèses du théorème 1 vérifiées; fixons R_1 et R_2 tels que $0 < R_1 < R_2 < R$. Alors il existe un $\varepsilon > 0$ et une constante $k' \geq 0$ tels que, pour tout $x \in \partial B_{R_1}$, la fonction $\ln G(x, \cdot)$ soit k' -lipschitzienne sur B_ε .

5 Le théorème de Girsanov

Pour calculer la loi d'une diffusion dérivante, une technique consiste à se ramener à un mouvement brownien standard. Notons W_x la loi d'un mouvement brownien issu de x , de constante de diffusion σ . On a la

Proposition 9. *La densité d'une diffusion dérivante par rapport à un mouvement brownien standard est*

$$\frac{d\mathbb{P}_x}{dW_x}((X_u)_{0 \leq u \leq t}) = \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(b_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\}. \quad (30)$$

Preuve. \mathfrak{E} Dans cette preuve nous ne considérerons les processus aléatoires que sur l'intervalle de temps $[0, t]$.

Posons $\mathcal{M}_u = \frac{1}{\sigma} \int_0^u \vec{v}(X_v) \cdot \overrightarrow{dX}_v$, alors, sous W_x , \mathcal{M} est une martingale et la martingale exponentielle associée est $\mathcal{E}_u = \exp \left\{ \frac{1}{\sigma} \int_0^u \left(\vec{v}(X_v) \cdot \overrightarrow{dX}_v - \frac{|\vec{v}(X_u)|^2}{2} dv \right) \right\}$. Par conséquent, il nous faut démontrer que, sous la loi $\mathcal{E}_t dW_x$, $(X_u)_{0 \leq u \leq t}$ est une diffusion brownienne avec dérive \vec{v} . Or le théorème de Girsanov affirme que, sous $\mathcal{E}_t dW_x$, $X_u - \langle X, \mathcal{M} \rangle_u$ est une martingale de variation quadratique $\sigma \mathbf{I}_n du$, donc un mouvement brownien. Par conséquent, puisque $\langle X, \mathcal{M} \rangle_u = \int_0^u \vec{v}(X_v) dv$, cela signifie bien que, sous $\mathcal{E}_t dW_x$, X vérifie l'équation (1), CQFD. \square

6 Explicitation de la densité de transition

Armés de la proposition 9, nous sommes en mesure d'expliciter la densité de transition p associée à la diffusion dérivante. Commençons par quelques rappels sur le pont brownien :

Définition 10. *Soient $x, y \in \mathbb{R}^d$, soit $t > 0$. Le pont brownien issu de x et arrivant au point y au temps t est le processus aléatoire obtenu comme la mesure-image du mouvement brownien W_x par la transformation :*

$$(B_u)_{u \geq 0} \mapsto \left(B_u + \frac{u}{t} \overrightarrow{B_t y} \right)_{0 \leq u \leq t}. \quad (31)$$

On notera $Q_{x \rightsquigarrow y}^t$ la loi de ce processus.

Nous rappelons sans démonstration les trois grandes propriétés du pont brownien, pour lesquelles on pourra par exemple se référer à [4] :

Théorème 11.

1. Soient $x \in \mathbb{R}^d$, $t > 0$. On définit une variable aléatoire $y \stackrel{(loi)}{\sim} \mathcal{N}(x, \sigma t \mathbf{I}_d)$, et un processus $(B_u)_{0 \leq u \leq t}$ tel que, conditionnellement à y , B soit un pont brownien de x à y . Alors la loi déconditionnée de B est la loi W_x d'un mouvement brownien issu de x .
2. La loi du pont brownien est aussi la loi du processus défini par l'équation différentielle stochastique :

$$\begin{cases} X_0 = x & ; \\ \overrightarrow{dX}_u = \overrightarrow{dB}_u + \frac{\overrightarrow{X_u y}}{t-u} du. \end{cases} \quad (32)$$

3. Si $(X_u)_{0 \leq u \leq t}$ est un pont brownien issu de x et arrivant en y au temps t , alors $(X_{t-u})_{0 \leq u \leq t}$ a la loi d'un pont brownien issu de y et arrivant en x au temps t .

Nous pouvons maintenant énoncer la

Proposition 12. Pour $t > 0$, $x, y \in B_{R_2}$, on a :

$$p_*^t(x, y) = \frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q_{x \rightsquigarrow y} \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot d\vec{X}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right]. \quad (33)$$

Preuve. ☞ Pour des raisons typographiques, dans cette preuve nous abrègerons $\vec{v}(X_u) \cdot d\vec{X}_u - \frac{|\vec{v}(X_u)|^2}{2} dt$ en $d\mathcal{A}_u$.

La démonstration n'est qu'un jeu d'écriture à partir de tout ce que nous avons déjà dit. On écrit, pour $t > 0$, $x \in B_{R_2}$ et $A \subset B_{R_2}$ mesurable :

$$\begin{aligned} \int_A p_*^t(x, y) dy &\stackrel{\text{(déf)}}{=} \mathbb{P}_x \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} \mathbb{1}_{X_t \in A} \right] \\ &\stackrel{\text{(prop. 9)}}{=} W_x \left[\exp \left\{ \frac{1}{\sigma} \int_0^t d\mathcal{A}_u \right\} \mathbb{1}_{t < \tau_{cB_{R_2}}} \mathbb{1}_{X_t \in A} \right] \\ &\stackrel{\text{(thm 11-1)}}{=} \int_A \frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q_{x \rightsquigarrow y} \left[\exp \left\{ \frac{1}{\sigma} \int_0^t d\mathcal{A}_u \right\} \mathbb{1}_{t < \tau_{cB_{R_2}}} \right] dy, \end{aligned} \quad (34)$$

d'où le résultat par identification. \square

Notre travail va consister à utiliser (33) pour contrôler les variations de $p_*^t(x, y)$. Plus précisément nous allons démontrer la proposition suivante, qui entraîne la proposition 8 et donc le théorème 1 :

Proposition 13. Supposons \vec{v} k -lipschitzienne sur B_R . Alors on peut trouver un $\varepsilon > 0$ tel que, d'une part il existe deux fonctions $\pi_1(t)$ et $\pi_2(t)$ avec :

$$\forall t > 0 \quad \forall x \in \partial B_{R_1} \quad \forall y \in B_\varepsilon \quad \pi_1(t) \leq p_*^t(x, y) \leq \pi_2(t); \quad (35)$$

$$0 < \int_0^\infty \pi_1(t) dt \leq \int_0^\infty \pi_2(t) dt < \infty, \quad (36)$$

d'autre part il existe une fonction $\kappa(t)$ avec :

$$\forall t > 0 \quad \forall x \in \partial B_{R_1} \quad y \mapsto p_*^t(x, y) \text{ est } \kappa(t)\text{-lipschitzienne sur } B_\varepsilon; \quad (37)$$

$$\int_0^\infty \kappa(t) dt < \infty. \quad (38)$$

Preuve. Montrons que la proposition 13 entraîne la proposition 8. Nous ferons ici abstraction de la question de l'explicitité, car celle-ci, le lecteur le vérifiera, ne pose de problème nulle part dans ce mémoire⁽³⁾.

Déjà, la condition (36) assure que G est uniformément strictement positive et finie sur $\partial B_{R_1} \times B_\varepsilon$, d'où il s'ensuit, puisque la mesure θ introduite à la proposition 5 est non nulle et finie, que f est (uniformément) strictement positive et finie sur B_ε . Maintenant, on a pour $y, y' \in B_\varepsilon$, notant $\delta = |\overrightarrow{yy'}|$:

$$\left| \frac{G(y')}{G(y)} - 1 \right| \leq \frac{\int_0^\infty \delta \kappa(t) dt}{\int_0^\infty \pi_1(t) dt} \leq c_1 \delta, \quad (39)$$

où $c_1 = \frac{\int_0^\infty \kappa(t) dt}{\int_0^\infty \pi_1(t) dt}$ est une constante finie. Or, pour $c_2 > 1$ une constante arbitraire, il existe $c_3 > 0$ telle qu'on ait pour tous $a, b > 0$:

$$\left| \frac{b}{a} - 1 \right| \leq c_3 \quad \Rightarrow \quad |\ln b - \ln a| \leq c_2 \left| \frac{b}{a} - 1 \right|. \quad (40)$$

Par connexité de B_ε , on en tire que $\ln G$ est $c_2 c_1$ -lipschitzienne. \square

7 Quelques estimées

Le but de cette partie est d'obtenir des estimées convenables pour les fonctions $\pi_1(t)$ et $\pi_2(t)$ introduites dans la proposition 13, ainsi que d'obtenir quelques majorations qui nous serviront pour obtenir une valeur pour $\kappa(t)$. Pour l'instant, nous supposons simplement que \vec{v} est bornée sur B_R ; nous notons $W = \sup_{B_R} |\vec{v}|$ ⁽⁴⁾. ε est supposé choisi tel que $\varepsilon < R_1$.

7.1 Majoration, I

☞ Dans ce paragraphe, on se fixe $t > 0$ et $x, y \in B_{R_2}$.

Tout part de la constatation suivante :

³Attirons simplement l'attention sur deux points à ce sujet : *primo*, les expressions qu'on obtient font intervenir R_1 , R_2 et ε , mais ce n'est pas un problème car toutes ces valeurs peuvent être rendues explicites — par exemple, on a défini R_1 et R_2 comme deux valeurs arbitraires vérifiant $0 < R_1 < R_2 < R$, par conséquent pour rendre ces valeurs explicites il suffit de les choisir respectivement égales, mettons, à $R/3$ et $2R/3$. *Secundo*, pour établir l'existence de certaines constantes on est amené à choisir des durées arbitraires non triviales. Dans ce cas, l'explicitité peut s'établir en prenant comme durée de référence R^2/σ , qui est bien homogène à un temps. On arrive de même à construire des constantes non triviales de toutes homogénéités.

⁴On observe que, sous les hypothèses du théorème 1, on peut bien majorer W par l'expression explicite $V + kR$.

Avis 14. Soit $(X_u)_{0 \leq u \leq t}$ un pont brownien de x à y . Soit $(\mathcal{M}_u)_{0 \leq u \leq t/2}$ le processus aléatoire défini par

$$\mathcal{M}_u = \frac{1}{\sigma} \int_0^{u \wedge \check{\tau}_{cB_{R_2}}} \vec{v}(X_v) \cdot \left(\overrightarrow{dX}_v - \frac{1}{t-v} \overrightarrow{X}_v y \, dv \right), \quad (41)$$

alors, sous $Q_{x \rightsquigarrow y}$, \mathcal{M} est une martingale continue issue de 0, de variation quadratique vérifiant :

$$Q_{x \rightsquigarrow y}\text{-p.s.} \quad \forall u \quad d\langle \mathcal{M} \rangle_u \leq \frac{W^2}{\sigma} du. \quad (42)$$

On a un résultat analogue par retournement du temps : notons $\check{\tau}_{cB_{R_2}}$ l'instant de sortie hors de B_{R_2} pris en remontant le temps, i.e. $\check{\tau}_{cB_{R_2}} = \sup\{u \leq t; X_u \notin B_{R_2}\}$, et soit $(\mathcal{N}_u)_{0 \leq u \leq t/2}$ le processus aléatoire défini par

$$\mathcal{N}_u = \frac{1}{\sigma} \int_{(t-u) \vee \check{\tau}_{cB_{R_2}}}^t \vec{v}(X_v) \cdot \left(\overrightarrow{dX}_v - \frac{1}{v} x \overrightarrow{X}_v \, dv \right), \quad (43)$$

alors \mathcal{N} est une martingale continue issue de 0, de variation quadratique vérifiant :

$$Q_{x \rightsquigarrow y}\text{-p.s.} \quad \forall u \quad d\langle \mathcal{N} \rangle_u \leq \frac{W^2}{\sigma} du. \quad (44)$$

Remarque. Nous avons donc écrit $\int_0^t \vec{v}(X_u) \overrightarrow{dX}_u$ comme somme d'un terme de martingale et d'un terme de dérive. Mais pourquoi donc s'encombrer de deux expressions différentes ? Eh bien, nous connaissons des techniques générales efficaces pour contrôler le terme de martingale, mais le terme de dérive tel qu'il apparaît dans (41) devient difficile à contrôler quand $u \rightarrow t$. D'où l'intérêt d'avoir deux écritures du pont brownien qui permettent d'avoir un terme de dérive « raisonnable » quelle que soit la valeur de u , selon qu'on utilise \mathcal{M} ou \mathcal{N} .

On rappelle le résultat qui définit les martingales exponentielles :

Théorème 15. Soit $(\mathcal{L}_u)_{u \geq 0}$ une martingale locale continue. Alors

$$e^{\mathcal{L}_u - \langle \mathcal{L} \rangle_u / 2} \quad (45)$$

est une martingale locale.

On en tire immédiatement le contrôle suivant sur $\mathcal{M}_{t/2}$, resp. $\mathcal{N}_{t/2}$:

Proposition 16.

$$\forall \lambda \in \mathbb{R} \quad Q_{x \rightsquigarrow y} \left[e^{\lambda \mathcal{M}_{t/2}} \right] \leq e^{\left(\frac{W^2 t}{4\sigma} \right) \lambda^2}, \quad (46)$$

$$\text{resp.} \quad Q_{x \rightsquigarrow y} \left[e^{\lambda \mathcal{N}_{t/2}} \right] \leq e^{\left(\frac{W^2 t}{4\sigma} \right) \lambda^2}. \quad (47)$$

D'où

Proposition 17.

$$p_*^t(x, y) \leq \frac{e^{-|\overrightarrow{xy}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} e^{W^2 t/\sigma + 4WR_2/\sigma}. \quad (48)$$

Preuve. On part de (33). Tout d'abord, on utilise la majoration triviale

$$\exp \left\{ - \int_0^\infty \frac{|\overrightarrow{v}(X_u)|^2}{2\sigma} dt \right\} \leq 1, \quad (49)$$

puis on écrit que, sur l'événement $t < \tau_{cB_{R_2}}$, on a la relation :

$$\frac{1}{\sigma} \int_0^t \overrightarrow{v}(X_u) \cdot d\overrightarrow{X}_u = \mathcal{M}_{t/2} + \mathcal{N}_{t/2} + \int_0^{t/2} \frac{\overrightarrow{X}_u \cdot \overrightarrow{v}(X_u)}{\sigma(t-u)} du + \int_{t/2}^t \frac{\overrightarrow{xX}_u \cdot \overrightarrow{v}(X_u)}{\sigma u} du, \quad (50)$$

où on a pour tout $u : |\overrightarrow{X}_u \cdot \overrightarrow{y}|, |\overrightarrow{xX}_u| \leq 2R_2$, d'où

$$\int_0^{t/2} \frac{\overrightarrow{X}_u \cdot \overrightarrow{v}(X_u)}{\sigma(t-u)} du + \int_{t/2}^t \frac{\overrightarrow{xX}_u \cdot \overrightarrow{v}(X_u)}{\sigma u} du \leq \frac{4R_2W}{\sigma}. \quad (51)$$

D'autre part

$$Q_{x \rightsquigarrow y} [e^{\mathcal{M}_{t/2} + \mathcal{N}_{t/2}}] \leq e^{\frac{W^2 t}{\sigma}}, \quad (52)$$

d'après (46) et (47) utilisées avec $\lambda = 2$ et l'inégalité de Cauchy-Schwartz⁽⁵⁾. Le résultat découle alors de la combinaison des formules (49) à (52). \square

Hélas, la majoration (48) ainsi obtenue est trop grossière ! En particulier, elle ne peut convenir pour définir π_2 puisque (36) n'est pas vérifiée. Nous allons donc chercher un majorant de $p_*^t(x, y)$ qui décroisse exponentiellement quand $t \rightarrow \infty$.

7.2 Majoration, II

Commençons par une propriété classique des diffusions uniformément elliptiques :

Proposition 18. *Il existe deux constantes $K_1 \geq 1$ et $K_2 > 0$ telles qu'on ait :*

$$\forall x \in B_{R_2} \quad \forall t \geq 0 \quad \mathbb{P}_x(t < \tau_{cB_{R_2}}) \leq K_1 e^{-K_2 t}, \quad (53)$$

i.e.

$$\forall x \in B_{R_2} \quad \forall t \geq 0 \quad \mathbb{P}_x^*(X_t \neq \partial) \leq K_1 e^{-K_2 t}. \quad (54)$$

⁵Attention : on ne peut pas concaténer \mathcal{M} et \mathcal{N} pour obtenir une troisième martingale.

Preuve. Soit $c_1 > 0$ une durée arbitraire. Nous allons montrer l'existence d'une constante $c_2 < 1$ telle qu'on ait

$$\forall x \in B_{R_2} \quad \mathbb{P}_x^*(X_{c_1} \neq \partial) \leq c_2. \quad (55)$$

En effet, si (55) est vraie alors on a par la relation de Chapman-Kolmogorov itérée :

$$\forall x \in B_{R_2} \quad \forall n \in \mathbb{N} \quad \mathbb{P}_x^*(X_{nc_1} \neq \partial) \leq c_2^n, \quad (56)$$

d'où, toujours par la relation de Chapman-Kolmogorov, en écrivant $t = (t - \lfloor \frac{t}{c_1} \rfloor c_1) + \lfloor \frac{t}{c_1} \rfloor c_1$:

$$\forall x \in B_{R_2} \quad \forall t \geq 0 \quad \mathbb{P}_x^*(X_t \neq \partial) \leq c_2^{\lfloor \frac{t}{c_1} \rfloor} \leq c_2^{(t/c_1)-1}, \quad (57)$$

ce qui donne bien (54) avec $K_1 = c_2^{-1}$ et $K_2 = \ln(c_2^{-1})/c_1$.

Reste à montrer (55). Soit $x \in B_{R_2}$. Soit \vec{v} un vecteur unitaire arbitraire ; pour $X \in \mathbb{R}^d$, notons $X \cdot \vec{v}$ la coordonnée de X relative à \vec{v} dans un repère orthonormé d'origine o .

Rappelons que le processus $(X_t)_{t \geq 0}$ de loi \mathbb{P}_x est défini par

$$\begin{cases} X_0 = x ; \\ \overrightarrow{dX}_t = \overrightarrow{dB}_t + \vec{v}(X_t) dt. \end{cases} \quad (58)$$

On définit alors un processus à valeurs réelles $(Y_t)_{t \geq 0}$ par

$$\begin{cases} Y_0 = -R_2 ; \\ dY_t = \overrightarrow{dB}_t \cdot \vec{v} - W dt, \end{cases} \quad (59)$$

où B est le même mouvement brownien que dans (58). Il est alors clair que, sur l'événement $\{c_1 < \tau_{cB_{R_2}}\}$, on a

$$\forall t \in [0, c_1] \quad Y_t < X_t \cdot \vec{v}. \quad (60)$$

Or, comme

$$\{c_1 < \tau_{cB_{R_2}}\} \subset \{X_{c_1} \in B_{R_2}\} \subset \{X_{c_1} \cdot \vec{v} < R_2\}, \quad (61)$$

il nous suffit finalement de montrer que $\mathbb{P}(Y_{c_1} \geq R_2) > 0$ pour conclure. Et justement Y est un mouvement brownien unidimensionnel avec dérive constante, dont on connaît parfaitement la loi :

$$\mathbb{P}(Y_{c_1} \geq R_2) = \mathbb{P}\left(\mathcal{N}(0, \sigma c_1) \geq W c_1 + 2R_2\right) > 0, \quad (62)$$

et la démonstration est terminée. \square

Nous allons nous servir de la proposition 18 pour obtenir un majorant de p_*^t à décroissance exponentielle en t . Dans toute la suite du texte, $K_3 > 0$ est une durée fixée. La décroissance exponentielle résulte alors de la relation de Chapman-Kolmogorov :

Proposition 19. *Pour $t \geq K_3$, on a pour tous $x, y \in B_{R_2}$:*

$$p_*^t(x, y) \leq \frac{K_1 e^{W^2 K_3 / \sigma + 4W R_2 / \sigma}}{(2\pi\sigma K_3)^{d/2}} e^{-K_2(t-t_3)}. \quad (63)$$

Par la suite nous noterons simplement $p_*^t(x, y) \leq K_4 e^{-K_2(t-t_3)}$.

Preuve. On écrit la relation de Chapman-Kolmogorov avec $t = (t - K_3) + K_3$:

$$\begin{aligned} p_*^t(x, y) &\stackrel{(9)}{=} \int_{B_{R_2}} p_*^{t-K_3}(x, x') p_*^{K_3}(x', y) dx' \\ &\stackrel{(48)}{\leq} \frac{e^{W^2 K_3 / \sigma + 4W R_2 / \sigma}}{(2\pi\sigma K_3)^{d/2}} \int_{B_{R_2}} p_*^{t-K_3}(x, x') dx' \\ &= \frac{e^{W^2 K_3 / \sigma + 4W R_2 / \sigma}}{(2\pi\sigma K_3)^{d/2}} \mathbb{P}_x^*(X_{t-K_3} \neq \partial) \\ &\stackrel{(54)}{\leq} K_4 e^{-K_2(t-K_3)}. \end{aligned} \quad (64)$$

□

Nous en déduisons une expression valable pour $\pi_2(t)$: on prend

$$\pi_2(t) = \begin{cases} \frac{1}{(2\pi\sigma t)^{d/2}} e^{(R_1 - \varepsilon)^2 / 2\sigma t + W^2 t / \sigma + 4W R_2 / \sigma} & \text{si } t < K_3 ; \\ K_4 e^{-K_2(t-K_3)} & \text{sinon.} \end{cases} \quad (65)$$

Ainsi l'expression de $\pi_2(t)$ est intégrable, et aussi bornée. En vue d'utilisations ultérieures, notons donc $K_5 = \sup_{t>0} \pi_2(t)$.

7.3 Minoration

☞ Dans ce paragraphe on suppose $x \in \partial B_{R_1}$, $y \in B_\varepsilon$.

Nous commençons par la minoration suivante sur le pont brownien :

Proposition 20. *Il existe deux constantes $K_6 > 0$ et $K_7 > 0$ telles qu'on ait :*

$$\forall x \in \partial B_{R_1} \quad \forall y \in B_\varepsilon \quad \forall t > 0 \quad Q_{x \rightsquigarrow y}^t(t < \tau_{cB_{R_2}}) \geq K_6 e^{-K_7 t}. \quad (66)$$

Preuve. Soient $x \in \partial B_{R_1}$, $y \in B_\varepsilon$. Par la définition 10 du pont brownien, on s'aperçoit que, notant W_x la loi du mouvement brownien standard issu de x :

$$Q_{x \rightsquigarrow y}^t(t < \tau_{cB_{R_2}}) \geq W_x(t < \tau_{cB(x, R_2 - R_1)}). \quad (67)$$

Par conséquent, nous allons simplement démontrer que le membre de droite de (67) est minoré par $K_6 e^{-K_7 t}$.

Pour alléger les notations, nous notons dans cette démonstration $r = R_2 - R_1$. Soit donc $(B_u)_{u \geq 0}$ un mouvement brownien issu de x . En nous plaçant dans un repère orthornormé d'origine x , les différentes coordonnées de B sont indépendantes et de même loi ; par conséquent, pour établir notre minoration exponentielle on peut se contenter de travailler sur un mouvement brownien unidimensionnel $(b_u)_{u \geq 0}$ issu de 0, dont nous noterons la loi w_0 , et montrer qu'il existe deux constantes K'_6 et K'_7 telles qu'on ait

$$w_0 \left(\sup_{u \in [0, t]} |b_u| \leq \frac{r}{\sqrt{d}} \right) \geq K'_6 e^{-K'_7 t}, \quad (68)$$

(67) venant alors avec $K_6 = K'_6{}^d$ et $K_7 = dK'_7$.

Montrons donc (68). Fixons-nous $c_1 \in]0, \frac{r}{\sqrt{d}}[$ arbitraire. Il nous suffit en fait de démontrer qu'il existe un temps $c_2 > 0$ et une constante $c_3 > 0$ tels que

$$\forall x' \in [-c_1, c_1] \quad w_{x'} \left(c_2 < \tau_{\mathfrak{q} - \frac{r}{\sqrt{d}}, \frac{r}{\sqrt{d}}} \text{ et } b_{c_2} \in [-c_1, c_1] \right) \geq c_3 : \quad (69)$$

ensuite, on appliquera la relation de Chapman-Kolmogorov itérée, et le résultat tombera avec $K'_6 = c_3$ et $K'_7 = \ln(c_3^{-1})/c_2$.

Soit donc $x' \in [-c_1, c_1]$; pour fixer les idées nous supposons $x' \geq 0$. On se fixe $c_4 > 0$; par autosimilarité du mouvement brownien on a pour tout $u \geq 0$:

$$w_{x'}(b_u \in [x' - c_4 u^{1/2}, x']) = c_5, \quad (70)$$

où c_5 est une probabilité strictement positive ne dépendant pas de u . Pour u suffisamment petit, $\forall x' \in [0, c_1]$ $[x' - c_4 u^{1/2}, x'] \subset [-c_1, c_1]$.

D'autre part, $w_{x'} \left(c_2 < \tau_{\mathfrak{q} - \frac{r}{\sqrt{d}}, \frac{r}{\sqrt{d}}} \right) \geq w_0 \left(c_2 < \tau_{\mathfrak{q} - \frac{r}{\sqrt{d}}, \frac{r}{\sqrt{d}} - c_1} \right) \xrightarrow{c_2 \rightarrow 0} 1$, donc, fixant $c_6 \in]0, c_5[$ arbitraire, si c_2 est pris suffisamment petit $\forall x' \in [0, c_1]$ $w_{x'} \left(c_2 < \tau_{\mathfrak{q} - \frac{r}{\sqrt{d}}, \frac{r}{\sqrt{d}}} \right) \geq 1 - c_6$, d'où (69) avec $c_3 = c_5 - c_6$. \square

Maintenant, nous allons énoncer et prouver un lemme de concentration pour parvenir à notre minoration. Mais d'abord, nous avons besoin d'un lemme préparatoire qui permet d'obtenir des résultats de concentration en espérance à partir de concentration sur les fonctions de répartition :

Lemme 21. *Soit $f : \mathbb{R} \rightarrow \mathbb{R}_+$ une fonction croissante. Soit \mathcal{X} une variable aléatoire réelle vérifiant :*

$$\forall x \in \mathbb{R} \quad \mathbb{P}(\mathcal{X} \leq x) \leq \Phi(x) \quad (71)$$

pour une fonction $\Phi : \mathbb{R} \rightarrow [0, 1]$ croissante ⁽⁶⁾. On définit l'inverse généralisée de Φ par $\Phi^{-1}(u) = \sup\{x \in \mathbb{R} ; \Phi(x) \leq u\}$. Alors, si \mathcal{A} est un événement de probabilité p , on a :

$$\mathbb{P}[\mathbb{1}_{\mathcal{A}} f(\mathcal{X})] \geq \int_0^p f(\Phi^{-1}(u)) du. \quad (72)$$

⁶À noter que n'importe quelle fonction $\Phi : \mathbb{R} \rightarrow [0, +\infty[$ peut être transformée en une fonction $\tilde{\Phi}$ convenable, en définissant $\tilde{\Phi}(x) = 1 \wedge \inf_{y \geq x} \Phi(y)$.

Preuve. Appelons (Ω, \mathcal{F}) l'espace de probabilité sur lequel on travaille. Soit $\Psi(x) = \mathbb{P}(\mathcal{X} \leq x)$ la véritable fonction de répartition de \mathcal{X} , et Ψ^{-1} son inverse généralisée; on a $\Psi \leq \Phi$, d'où $\Psi^{-1} \geq \Phi^{-1}$. Quitte à agrandir \mathcal{F} , on peut supposer qu'il existe une variable aléatoire u uniformément distribuée sur $[0, 1]$ telle que $\mathcal{X} = \Psi^{-1}(u)$. Définissons maintenant

$$\begin{aligned} g : \mathcal{A} &\longrightarrow [0, 1] \\ \omega &\mapsto \mathbb{P}(\omega' \in \mathcal{A} \text{ et } u(\omega') \leq u(\omega)). \end{aligned} \quad (73)$$

g est une fonction mesurable, et vérifie clairement $g(\omega) \leq u(\omega)$, d'où $f(\Psi^{-1}(g(\omega))) \leq f(\mathcal{X}(\omega))$ par croissance de Ψ^{-1} et de f . En outre, g a été définie de sorte que $g(\omega)$ soit distribuée uniformément sur $[0, p]$ (regarder les fonctions de répartition). Finalement, on a :

$$\begin{aligned} \mathbb{P}[\mathbb{1}_{\mathcal{A}} f(\mathcal{X})] &= \int_{\mathcal{A}} f(\mathcal{X}(\omega)) d\omega \\ &\geq \int_{\mathcal{A}} f(\Psi^{-1}(g(\omega))) d\omega = \int_0^p f(\Psi^{-1}(u)) du \geq \int_0^p f(\Phi^{-1}(u)) du. \end{aligned} \quad (74)$$

□

Remarque. Dans le cas où Φ est de classe C^1 , le changement de variables $x = \Phi^{-1}(u)$ dans l'intégrale donne :

$$\mathbb{P}[\mathbb{1}_{\mathcal{A}} f(\mathcal{X})] \geq \int_{-\infty}^{\Phi^{-1}(p)} f(x) \Phi'(x) dx, \quad (75)$$

expression souvent plus simple à manipuler.

Du lemme 21, on tire

Lemme 22. Soit \mathcal{X} une variable aléatoire suivant une loi P . On suppose que, pour une certaine constante $c_1 \geq 0$, on a le contrôle suivant sur la log-Laplace de \mathcal{X} :

$$\forall \lambda \in \mathbb{R} \quad \ln P[e^{\lambda \mathcal{X}}] \leq c_1 \frac{\lambda^2}{2}. \quad (76)$$

Soit $p \in [0, 1]$ et soit \mathcal{A} un événement avec $P(\mathcal{A}) \geq p$. Alors :

$$\forall \lambda \geq 0 \quad P[\mathbb{1}_{\mathcal{A}} e^{\lambda \mathcal{X}}] \geq \frac{p^2}{2} e^{-c_1 \lambda^2 / 2}. \quad (77)$$

Preuve. L'inégalité de Chernoff donne pour tout $\xi \leq 0$:

$$P(\mathcal{X} \leq \xi) \leq e^{-\xi^2 / 2c_1}. \quad (78)$$

Maintenant, soit $\lambda \geq 0$. Remarquons d'abord qu'on ne perd rien à supposer qu'on a exactement $P(\mathcal{A}) = p$. Via le lemme 21 et la remarque qui le suit, (78) donne :

$$P[\mathbb{1}_{\mathcal{A}} e^{\lambda \mathcal{X}}] \geq \frac{1}{c_1} \int_{-\infty}^{-\sqrt{2c_1 \ln(\frac{1}{p})}} (-\xi) e^{-\xi^2 / 2c_1 + \lambda \xi} d\xi. \quad (79)$$

Maintenant, on utilise l'inégalité élémentaire

$$-\frac{\xi^2}{2c_1} + \lambda\xi \geq -\frac{\xi^2}{c_1} - \frac{c_1\lambda^2}{2} \quad (80)$$

pour obtenir

$$P[\mathbb{1}_{\mathcal{A}}e^{\lambda\mathcal{X}}] \geq \frac{e^{-c_1\lambda^2/2}}{c_1} \int_{-\infty}^{-\sqrt{2c_1 \ln(\frac{1}{p})}} (-\xi)e^{-\xi^2/c_1} d\xi = \frac{p^2}{2} e^{-c_1\lambda^2/2}. \quad (81)$$

□

Finalement, on obtient une minoration convenable pour définir $\pi_2(t)$:

Proposition 23.

$$p_*^t(x, y) \geq \frac{K_6^2}{2(2\pi\sigma t)^{d/2}} \exp\left\{-\frac{2R_1^2}{\sigma t} - 2K_7 t - \frac{3W^2 t}{2\sigma} - \frac{4R_2 W}{\sigma}\right\}. \quad (82)$$

Preuve. On part de l'expression (33) pour, *via* quelques minoration triviales, obtenir :

$$p_*^t(x, y) \geq \frac{e^{-2R_1^2/\sigma t}}{(2\pi\sigma t)^{d/2}} e^{\frac{1}{\sigma}(-4R_2 W - W^2 t/2)} Q_{x \rightsquigarrow y}^t \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} e^{\mathcal{M}_{t/2} + \mathcal{N}_{t/2}} \right]. \quad (83)$$

Puis la proposition 20 nous donne qu'on a $Q_{x \rightsquigarrow y}^t (t < \tau_{cB_{R_2}}) \geq K_6 e^{-K_7 t}$. Comme par ailleurs la formule (52) donne le contrôle $Q_{x \rightsquigarrow y}^t [e^{\lambda(\mathcal{M}_{t/2} + \mathcal{N}_{t/2})}] \leq e^{\lambda^2 W^2 t/\sigma}$, le lemme 22 appliqué avec $c_1 = 2W^2 t/\sigma$, $p = K_6 e^{-K_7 t}$ et $\lambda = 1$ donne

$$Q_{x \rightsquigarrow y}^t \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} e^{\mathcal{M}_{t/2} + \mathcal{N}_{t/2}} \right] \geq \frac{K_6^2 e^{-2K_7 t}}{2} e^{-W^2 t/\sigma}. \quad (84)$$

Il ne reste plus qu'à combiner (83) et (84) pour conclure. □

8 Couplage

Dans cette partie, on suppose les hypothèses du théorème 1 vérifiées ; nous allons voir comment comparer $p_*^t(x, y)$ et $p_*^t(x, y')$ pour $y, y' \in B_\varepsilon$. Soient donc $\varepsilon > 0$ suffisamment petit — plus précisément, on prend $\varepsilon > 0$ quelconque tel que $\varepsilon \leq R_1$, $R_2 + 2\varepsilon \leq R$ et $R_2 - 2\varepsilon > R_1$ —, $y, y' \in B_\varepsilon$, $t > 0$ et $x \in \partial B_{R_1}$. Notons $\delta = |\overrightarrow{yy'}|$; dans la suite du texte, $O(\delta)$ notera toute une expression explicite en fonction de V, k, R (et des grandeurs définies à partir de ces données) et δ telle qu'à V, k, R fixés, on ait $\limsup_{\delta \rightarrow 0} \frac{|O(\delta)|}{\delta} < \infty$.

8.1 Principe

Le couplage repose sur le

Lemme 24. *Soit*

$$\begin{aligned} \psi : \mathcal{C}([0, t], \mathbb{R}^d) &\longrightarrow \mathcal{C}([0, t], \mathbb{R}^d) \\ (X_u)_{0 \leq u \leq t} &\mapsto (X'_u)_{0 \leq u \leq t} = \left(X_u + \frac{u}{t} \overrightarrow{yy'} \right)_{0 \leq u \leq t}, \end{aligned} \quad (85)$$

alors la mesure-image de $Q_{x \rightsquigarrow y}$ par ψ est $Q_{x \rightsquigarrow y'}$.

Preuve. Soit $(B_u)_{0 \leq u \leq t}$ un mouvement brownien de loi W_x , et soient

$$\varphi : (B_u)_{0 \leq u \leq t} \mapsto \left(B_u + \frac{u}{t} \overrightarrow{B_t y} \right)_{0 \leq u \leq t}; \quad (86)$$

$$\varphi' : (B_u)_{0 \leq u \leq t} \mapsto \left(B_u + \frac{u}{t} \overrightarrow{B_t y'} \right)_{0 \leq u \leq t}. \quad (87)$$

Alors par la définition 10 du pont brownien, $\varphi(B)$ a pour loi $Q_{x \rightsquigarrow y}$ et $\varphi'(B)$ a pour loi $Q_{x \rightsquigarrow y'}$. Comme $\psi \circ \varphi = \varphi'$, on a bien que la mesure-image de $Q_{x \rightsquigarrow y}$ par ψ est $Q_{x \rightsquigarrow y'}$. \square

Grâce au lemme 24 nous pouvons donc exprimer la loi $Q_{x \rightsquigarrow y'}$ en fonction de la loi $Q_{x \rightsquigarrow y}$. La seule loi de pont brownien à laquelle nous aurons recours sera ainsi $Q_{x \rightsquigarrow y}$, que nous ne noterons plus que Q dans la suite pour alléger. Par ailleurs, nous noterons systématiquement $(X'_u)_{0 \leq u \leq t}$ pour $\psi((X_u)_{0 \leq u \leq t})$, et pour F un fermé de \mathbb{R}^d , τ'_F notera le temps d'atteinte de F par X' . Le couplage va nous permettre de donner une expression de $\kappa(t)$ satisfaisant (37) et (38).

☞ Dans un premier temps de notre raisonnement, nous ne supposons pas qu'on a $x \in \partial B_{R_1}$, mais simplement $x \in B_{R_2}$.

Par (33) et par couplage, on a :

$$\begin{aligned} p_*^t(x, y') - p_*^t(x, y) &= \\ &= \frac{e^{-|\overrightarrow{xy'}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\overrightarrow{v}(X'_u) \cdot d\overrightarrow{X}_u + \overrightarrow{v}(X'_u) \cdot \overrightarrow{yy'} \frac{du}{t} - \frac{|\overrightarrow{v}(X'_u)|^2}{2} du \right) \right\} \right] \\ &\quad - \frac{e^{-|\overrightarrow{xy}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\overrightarrow{v}(X_u) \cdot d\overrightarrow{X}_u - \frac{|\overrightarrow{v}(X_u)|^2}{2} du \right) \right\} \right]. \end{aligned} \quad (88)$$

8.2 Demi-tours au bord

L'équation (88) nous laisse à penser qu'on va pouvoir écrire $p_*^t(x, y') - p_*^t(x, y)$ comme un $O(\delta)$, puisque son second membre peut se réécrire comme l'espérance, sous Q , de deux quantités qui deviennent très proches quand

$\delta \rightarrow 0$. À une exception près toutefois : $\mathbb{1}_{\tau'_{cB_{R_2}}}$ peut être égale à 0 tandis que $\mathbb{1}_{\tau_{cB_{R_2}}}$ serait égale à 1, ou vice-versa. En effet, si la particule X vient à s'approcher très près de ∂B_{R_2} sans toutefois franchir cette frontière, il se peut que X' , elle, la franchisse, auquel cas une des deux particules sera tuée et pas l'autre — le phénomène inverse est également possible. Nous voulons montrer que néanmoins, ces phénomènes de « demi-tour au bord » deviennent suffisamment improbables quand δ tend vers 0.

Nous avons d'abord besoin d'introduire quelques notations :

Définition 25. On appelle B'_{R_2} la boule B_{R_2} décalée du vecteur $\vec{y}'y$; on note $B^*_{R_2}$ l'enveloppe convexe de $B_{R_2} \cup B'_{R_2}$, et $B^{**}_{R_2} = B_{R_2} \cap B'_{R_2}$. On note $\partial^+ B_{R_2} = B_{R_2} \setminus B'_{R_2}$, et $\partial^- B_{R_2} = B^*_{R_2} \setminus B_{R_2}$ (⁷). On note enfin $\partial^* B_{R_2} = \partial^+ B_{R_2} \cup \partial^- B_{R_2} = B^*_{R_2} \setminus B^{**}_{R_2}$. Par souci de légèreté, en l'absence d'ambiguïté nous omettrons les indices R_2 .

On a alors immédiatement :

Avis 26.

$$\begin{cases} \{t < \tau_{cB}\} \setminus \{t < \tau'_{cB}\} \subset \{\tau_{\overline{\partial^+ B}} < t\} \cap \{t < \tau_{cB^*}\} \\ \{t < \tau'_{cB}\} \setminus \{t < \tau_{cB}\} \subset \{\tau_{\overline{\partial^- B}} < t\} \cap \{t < \tau_{cB^*}\}, \end{cases} \quad (89)$$

donc

$$\{t < \tau_B\} \Delta \{t < \tau'_B\} \subset \{\tau_{\overline{\partial^* B}} < t\} \cap \{t < \tau_{cB^*}\}. \quad (90)$$

Notons $^{1/2}\text{-tr au b.}$ l'événement $\{\tau_{\overline{\partial^* B}} < t\} \cap \{t < \tau_{cB^*}\}$. On écrit que :

$$\begin{aligned} & \left| \frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau_{cB_{R_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right] \right. \\ & \quad \left. - \frac{e^{-|\vec{x}\vec{y}'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right] \right| \\ & \leq \underbrace{\frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{^{1/2}\text{-tr au b.}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right]}_{(*)}. \quad (91) \end{aligned}$$

Un raisonnement analogue à celui du § 6 nous montre que (*) est une sorte de « densité de demi-tour au bord », i.e. que (*) vérifie pour tout $A \subset B_{R_\varepsilon}$ mesurable :

$$\int_A (*) dy = \mathbb{P}_x(^{1/2}\text{-tr au b. et } X_t \in A). \quad (92)$$

Cela va nous permettre d'obtenir un majorant de (*) sur B_ε : on va montrer qu'il existe une constante c_1 indépendante de A telle que

$$\mathbb{P}_x(^{1/2}\text{-tr au b. et } X_t \in A) \leq c_1 |A|, \quad (93)$$

⁷À noter que $\partial^+ B$ et $\partial^- B$ ne jouent pas tout-à-fait des rôles symétriques.

d'où on tirera $(*) \leq c_1$ pour tout $y \in B_\varepsilon$.

En utilisant la propriété de Markov avec les temps d'arrêts $\tau_{\overline{\partial^* B}}$ et $\inf\{u \geq \tau_{\overline{\partial^* B}}; X_u \in \overline{B_{R_1}}\}$, on s'aperçoit que

$$\mathbb{P}(1/2\text{-tr au b. et } X_t \in A) \leq \sup_{x' \in \overline{\partial^* B}} \mathbb{P}_{x'} \left(\tau_{\overline{B_{R_1}}} < \tau_{cB^*} \right) \cdot \sup_{\substack{x'' \in \partial B_{R_1} \\ u > 0}} \mathbb{P}_{x''} (X_u \in A). \quad (94)$$

Dans le second membre de (94), le facteur de droite est majoré par $K_5|A|$ d'après les majorations du § 7.2. Il nous reste donc à majorer le facteur de gauche.

On commence par une observation géométrique :

Avis 27. Soit $x' \in \overline{\partial^* B}$. Soit \vec{v} le vecteur unitaire $\overrightarrow{ox'}/|ox'|$. Alors, pour $z \in \mathbb{R}^d$:

$$\begin{cases} \vec{x'z} \cdot \vec{v} \geq \delta & \Rightarrow z \notin B^*; \\ z \in \overline{B_{R_1}} & \Rightarrow \vec{x'z} \cdot \vec{v} \leq -(R_2 - R_1 - 2\varepsilon). \end{cases} \quad (95)$$

Cette observation va nous permettre de nous ramener à un mouvement brownien unidimensionnel, par une technique de couplage analogue à celle de la démonstration de la proposition 18, et que par conséquent nous ne détaillerons pas ici. Toujours est-il que, notant $w_{0,-W}$ la loi d'un mouvement brownien unidimensionnel issu de 0 et soumis à une dérive constante $-W$, il vient que :

$$\mathbb{P}_{x'} \left(\tau_{\overline{B_{R_1}}} < \tau_{cB^*} \right) \leq w_{0,-W} \left(\tau_{\{-(R_2 - R_1 - 2\varepsilon)\}} < \tau_{\{\delta\}} \right). \quad (96)$$

Or, sous la loi $w_{0,-W}$, $e^{\frac{2W}{\sigma}x_u}$ est une martingale, comme on le vérifie par la formule d'Itô. Il s'ensuit, puisque le mouvement brownien avec dérive finit presque-sûrement par quitter $]- (R_2 - R_1 - 2\varepsilon), \delta[$, que

$$w_{0,-W} \left(\tau_{\{-(R_2 - R_1 - 2\varepsilon)\}} < \tau_{\{\delta\}} \right) = \frac{e^{\frac{2W}{\sigma}\delta} - 1}{e^{\frac{2W}{\sigma}\delta} - e^{-\frac{2W}{\sigma}(R_2 - R_1 - 2\varepsilon)}} = O(\delta). \quad (97)$$

Finalement, pour $y \in B_\varepsilon$:

$$(*) \leq K_5 O(\delta) = O(\delta). \quad (98)$$

8.3 Termes de couplage

Maintenant, nous voulons majorer

$$\begin{aligned}
& \left| \frac{e^{-|\vec{xy}'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau'_{c_{B_{R_2}}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X'_u) \cdot \overrightarrow{dX}_u + \vec{v}(X'_u) \cdot \frac{\vec{yy}'}{t} - \frac{|\vec{v}(X'_u)|^2}{2} du \right) \right\} \right. \right. \\
& \quad \left. \left. - \frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau'_{c_{B_{R_2}}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right] \right| \\
& \leq Q \left[\mathbb{1}_{t < \tau'_{c_{B_{R_2}}}} \left| \frac{e^{-|\vec{xy}'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X'_u) \cdot \overrightarrow{dX}_u + \vec{v}(X'_u) \cdot \frac{\vec{yy}'}{t} - \frac{|\vec{v}(X'_u)|^2}{2} du \right) \right\} \right. \right. \\
& \quad \left. \left. - \frac{e^{-|\vec{x}\vec{y}|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right| \right]. \quad (99)
\end{aligned}$$

Pour borner ce genre d'expression, on utilise un lemme élémentaire très pratique, qui permet de séparer la majoration en plusieurs termes :

Lemme 28. Soient $a, a', b, b' \geq 0$, on a :

$$|a'b' - ab| \leq |a' - a|b + |b' - b|a'. \quad (100)$$

Preuve. On utilise simplement l'identité $a'b' - ab = (a' - a)b + (b' - b)a$. \square

Avant de nous lancer dans les calculs, remarquons que si, comme nous l'avons suggéré, ε est suffisamment petit pour avoir $R_2 + 2\varepsilon \leq R$, on a sous l'événement $\{t < \tau'_{c_{B_{R_2}}}\}$ que $X_u, X'_u \in B_R$ pour tout $u \in [0, t]$.

Pour commencer, on observe que

$$|e^{-|\vec{xy}'|^2/2\sigma t} - e^{-|\vec{x}\vec{y}|^2/2\sigma t}| \leq e^{-|\vec{xy}'|^2/2\sigma t} |e^{4R_2\delta/\sigma t} - 1| : \quad (101)$$

en effet $|\vec{xy}'|^2 - |\vec{x}\vec{y}|^2 = (\vec{x}\vec{y} + \vec{xy}') \cdot \vec{yy}'$, d'où $|\vec{xy}'|^2 - |\vec{x}\vec{y}|^2 \leq (|\vec{x}\vec{y}| + |\vec{xy}'|)\delta \leq 4R_2\delta$. Comme, pour $|x| \leq a$, on a $|e^x - 1| \leq e^a - 1$, (101) s'ensuit en factorisant par $e^{-|\vec{xy}'|^2/2\sigma t}$.

De (101), on déduit par le lemme 28 :

$$\begin{aligned}
|p_*^t(x, y') - p_*^t(x, y)| & \leq O(\delta) + \frac{e^{-|\vec{xy}'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \cdot (e^{4R_2\delta/\sigma t} - 1) \cdot e^{W^2t/\sigma + 4WR/\sigma} \\
& \quad + \frac{e^{-|\vec{xy}'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \times \\
& \quad Q \left[\mathbb{1}_{t < \tau'_{c_{B_{R_2}}}} \left| \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X'_u) \cdot \overrightarrow{dX}_u + \vec{v}(X'_u) \cdot \frac{\vec{yy}'}{t} - \frac{|\vec{v}(X'_u)|^2}{2} du \right) \right\} \right. \right. \\
& \quad \left. \left. - \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX}_u - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right| \right], \quad (102)
\end{aligned}$$

où on a utilisé implicitement la majoration

$$Q \left[\mathbb{1}_{t < \tau'_{c_{BR_2}}} \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX_u} - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right] \leq e^{W^2 t / \sigma + 4WR / \sigma}, \quad (103)$$

qui résulte d'un travail tout-à-fait similaire à celui du § 7.1.

Passons à l'étape suivante. On aura remarqué que le couplage fait apparaître un terme en plus dans l'intégrale en exponentielle, dû à la dérive supplémentaire que doit subir la particule pour arriver en y' . En utilisant la majoration triviale

$$t < \tau'_{c_{BR_2}} \quad \Rightarrow \quad \left| \frac{1}{\sigma} \int_0^t \vec{v}(X'_u) \cdot \frac{\vec{y}y'}{t} du \right| \leq \frac{W\delta}{\sigma}, \quad (104)$$

on obtient que

$$\begin{aligned} |p_*^t(x, y') - p_*^t(x, y)| &\leq O(\delta) \\ &+ \frac{e^{-|\vec{x}y'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} (e^{4R_2\delta/\sigma t} + e^{W\delta/\sigma} - 2) e^{W^2 t / \sigma + 4WR / \sigma} \\ &+ \frac{e^{-|\vec{x}y'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \times Q \left[\mathbb{1}_{t < \tau'_{c_{BR_2}}} \left| \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X'_u) \cdot \overrightarrow{dX_u} - \frac{|\vec{v}(X'_u)|^2}{2} du \right) \right\} \right. \right. \\ &\quad \left. \left. - \exp \left\{ \frac{1}{\sigma} \int_0^t \left(\vec{v}(X_u) \cdot \overrightarrow{dX_u} - \frac{|\vec{v}(X_u)|^2}{2} du \right) \right\} \right| \right]. \quad (105) \end{aligned}$$

Au terme suivant ! Nous observons qu'on a :

$$\begin{aligned} &||\vec{v}(X'_u)|^2 - |\vec{v}(X_u)|^2| \\ &= |\vec{v}(X_u) + \vec{v}(X'_u)| \cdot |\vec{v}(X'_u) - \vec{v}(X_u)| \leq 2W \cdot k |\overrightarrow{X_u X'_u}| \leq 2Wk\delta, \quad (106) \end{aligned}$$

d'où toujours par les mêmes arguments :

$$\begin{aligned} |p_*^t(x, y') - p_*^t(x, y)| &\leq O(\delta) + \\ &\frac{e^{-|\vec{x}y'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} (e^{4R_2\delta/2\sigma t} + e^{W\delta/\sigma} + e^{2Wk\delta/\sigma} - 3) e^{W^2 t / \sigma + 4WR / \sigma} \\ &+ \frac{e^{-|\vec{x}y'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} Q \left[\mathbb{1}_{t < \tau'_{c_{BR_2}}} \left| e^{\frac{1}{\sigma} \int_0^t \vec{v}(X'_u) \cdot \overrightarrow{dX_u}} - e^{\frac{1}{\sigma} \int_0^t \vec{v}(X_u) \cdot \overrightarrow{dX_u}} \right| \right]. \quad (107) \end{aligned}$$

(on a utilisé implicitement que $e^{-\int_0^t \frac{|\vec{v}(X_u)|^2}{2\sigma} du} \leq 1$).

Il nous reste à majorer

$$Q \left[\mathbb{1}_{t < \tau'_{c_{BR_2}}} \left| e^{\frac{1}{\sigma} \int_0^t \vec{v}(X'_u) \cdot \overrightarrow{dX_u}} - e^{\frac{1}{\sigma} \int_0^t \vec{v}(X_u) \cdot \overrightarrow{dX_u}} \right| \right]. \quad (108)$$

Cette fois-ci, on ne peut plus utiliser le lemme 28 car les majorants à l'intérieur de l'intégrale ne sont plus constants. Qu'à cela ne tienne ! On écrit que par l'inégalité de Cauchy-Schwartz :

$$(108) \leq Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} e^{\frac{2}{\sigma} \int_0^t \vec{v}(X_u) \cdot \overrightarrow{dX_u}} \right]^{1/2} \cdot Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} \left(e^{\frac{1}{\sigma} \int_0^t (\vec{v}(X'_u) - \vec{v}(X_u)) \cdot \overrightarrow{dX_u}} - 1 \right)^2 \right]^{1/2}. \quad (109)$$

Ainsi nous voyons apparaître des semimartingales dans l'exponentielle, que nous allons contrôler par les mêmes méthodes qu'au § 7.1.

Comme nous nous sommes arrangés pour avoir $X_u, X'_u \in B_R$ sous $\{t < \tau'_{cB_{R_2}}\}$, on a alors que sous cette condition $|\vec{v}(X_u)|, |\vec{v}(X'_u)| \leq W$ et $|\vec{v}(X'_u) - \vec{v}(X_u)| \leq k\delta$. On utilise alors exactement les mêmes arguments que dans le § 7.1 pour aboutir à :

Proposition 29. 1. Sur l'événement $\{t < \tau'_{cB_{R_2}}\}$, la variable aléatoire $\frac{2}{\sigma} \int_0^t \vec{v}(X_u) \cdot \overrightarrow{dX_u}$ coïncide avec une décomposition $\mathcal{X}_d + \mathcal{X}_m$ ⁽⁸⁾, où \mathcal{X}_d et \mathcal{X}_m sont deux variables aléatoires vérifiant respectivement :

$$Q\text{-p.s.} \quad |\mathcal{X}_d| \leq 8 \frac{RW}{\sigma} \quad (110)$$

et

$$\forall \lambda \in \mathbb{R} \quad Q[e^{\lambda \mathcal{X}_m}] \leq e^{\left(4 \frac{W^2 t}{\sigma}\right) \lambda^2}. \quad (111)$$

2. De même, sur $\{t < \tau'_{cB_{R_2}}\}$, l'intégrale $\frac{1}{\sigma} \int_0^t (\vec{v}(X'_u) - \vec{v}(X_u)) \cdot \overrightarrow{dX_u}$ s'écrit $\mathcal{Y}_d + \mathcal{Y}_m$, avec

$$Q\text{-p.s.} \quad |\mathcal{Y}_d| \leq 4 \frac{Rk\delta}{\sigma} \quad (112)$$

et

$$\forall \lambda \in \mathbb{R} \quad Q[e^{\lambda \mathcal{Y}_m}] \leq e^{\frac{k^2 \delta^2 t}{\sigma} \lambda^2}. \quad (113)$$

On en déduit dans un premier temps :

$$Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} e^{\frac{2}{\sigma} \int_0^t \vec{v}(X_u) \cdot \overrightarrow{dX_u}} \right]^{1/2} \leq e^{4 \frac{RW}{\sigma} + 2 \frac{W^2 t}{\sigma}}. \quad (114)$$

Pour contrôler le facteur de droite du second membre de (109), on a besoin d'un lemme *ad hoc* :

Lemme 30. Soit \mathcal{Y} une variable aléatoire de loi P , vérifiant

$$\forall \lambda \in \mathbb{R} \quad P[e^{\lambda \mathcal{Y}}] \leq e^{c_1 \frac{\lambda^2}{2}} \quad (115)$$

pour une constante $c_1 \geq 0$. Alors :

$$P[(e^{\mathcal{Y}} - 1)^2] \leq e^{2c_1} + 2e^{c_1/2} - 3. \quad (116)$$

⁸ d comme « dérive » et m comme « martingale », Nda.

Preuve. Il suffit de vérifier l'inégalité analytique

$$\forall y \in \mathbb{R} \quad (e^y - 1)^2 \leq e^{2y} + 2e^{-y} - 3 \quad (117)$$

et d'utiliser (115) avec $\lambda = 2$ et $\lambda = -1$. \square

On écrit alors

$$\begin{aligned} Q \left[(e^{\mathcal{Y}_d + \mathcal{Y}_m} - 1)^2 \right] &= Q \left[(e^{\mathcal{Y}_d} (e^{\mathcal{Y}_m} - 1) + (e^{\mathcal{Y}_d} - 1))^2 \right] \\ &\leq 2Q \left[e^{2\mathcal{Y}_d} (e^{\mathcal{Y}_m} - 1)^2 + (e^{\mathcal{Y}_d} - 1)^2 \right] \end{aligned} \quad (118)$$

pour en déduire dans un deuxième temps :

$$\begin{aligned} Q \left[\mathbb{1}_{t < \tau'_{cB_{R_2}}} \left(e^{\frac{1}{\sigma} \int_0^t (\vec{v}(X'_u) - \vec{v}(X_u)) \cdot \overrightarrow{dX_u}} - 1 \right)^2 \right]^{1/2} &\leq \\ \sqrt{2} \left(e^{8 \frac{kR\delta}{\sigma}} \left(e^{4 \frac{k^2 t \delta^2}{\sigma}} + 2e^{\frac{k^2 t \delta^2}{\sigma}} - 3 \right) + \left(e^{\frac{4kR\delta}{\sigma}} - 1 \right)^2 \right)^{1/2}. \end{aligned} \quad (119)$$

Donc, en récapitulant tout, on a enfin notre formule de couplage :

$$\begin{aligned} |p_*^t(x, y') - p_*^t(x, y)| &\leq O(\delta) + \\ &\frac{e^{-|xy'|^2/2\sigma t}}{(2\pi\sigma t)^{d/2}} \left\{ \left(e^{\frac{4R_2\delta}{2\sigma t}} + e^{\frac{W\delta}{\sigma}} + e^{\frac{2Wk\delta}{\sigma}} - 3 \right) e^{\frac{4WR}{\sigma} + \frac{W^2 t}{\sigma}} \right. \\ &\left. + \sqrt{2} e^{4 \frac{RW}{\sigma} + 2 \frac{W^2 t}{\sigma}} \left(e^{8 \frac{kR\delta}{\sigma}} \left(e^{4 \frac{k^2 t \delta^2}{\sigma}} + 2e^{\frac{k^2 t \delta^2}{\sigma}} - 3 \right) + \left(e^{\frac{4kR\delta}{\sigma}} - 1 \right)^2 \right)^{1/2} \right\}. \end{aligned} \quad (120)$$

Hélas, nous rencontrons le même écueil qu'avec la formule (48) que nous avons obtenue au § 7.1, à savoir que notre majorant va mal se comporter quand $t \rightarrow \infty$, ce qui nous interdit de l'utiliser pour définir $\kappa(t)$, car (38) ne serait pas vérifiée.

À même mal, même remède! Nous allons utiliser la décroissance exponentielle de \mathbb{P}^* jointe à la relation de Chapman-Kolomogorov pour transformer (120) en un résultat exponentiellement décroissant en t , donc convergent par intégration.

8.4 Contrôle aux temps grands

D'abord, commençons par vérifier que la formule (120) se comporte bien aux temps petits, i.e. qu'elle donne une borne uniforme en x et t pour $x \in \partial B_{R_1}$ et $t \leq K_3$. Nous avons besoin pour ce faire de deux lemmes analytiques élémentaires :

Lemme 31.

1. Pour $a > 0$:

$$\sup_{t \geq 0} \frac{e^{-a/t}}{t^{d/2}} = \left(\frac{d}{2ea} \right)^{d/2}. \quad (121)$$

2. Pour $a > b > 0$:

$$\sup_{t \geq 0} e^{-a/t} (e^{b/t} - 1) \leq b \left(\frac{d+2}{2e(a-b)} \right)^{d/2+1}. \quad (122)$$

Preuve. Le n° 1 vient simplement d'une étude de fonction ; le n° 2 s'obtient à partir du n° 1 et de l'inégalité $e^c - 1 \leq ce^c$ pour $c \geq 0$. \square

On tire alors de (120), pour $x \in \partial B_{R_1}$, $y, y' \in B_\varepsilon$, $t \leq K_3$, pour δ suffisamment petit :

$$\begin{aligned} |p_*^t(x, y') - p_*^t(x, y)| &\leq O(\delta) + \left\{ \frac{1}{(2\pi\sigma)^{d/2}} \cdot \frac{2R_2\delta}{\sigma} \left(\frac{d+2}{2e((R_1-\varepsilon)^2/2\sigma - 2R_2\delta/\sigma)} \right)^{d/2+1} \right. \\ &+ \left. \left(\frac{d\sigma}{e(R_1-\varepsilon)^2} \right)^{d/2} \left(e^{W\delta/\sigma} + e^{2Wk\delta/\sigma} - 2 \right) \right\} \times e^{4WR/\sigma + W^2K_3/\sigma} + \left(\frac{d\sigma}{e(R_1-\varepsilon)^2} \right)^{d/2} \times \\ &\sqrt{2} e^{4RW/\sigma + 2W^2K_3/\sigma} \left\{ e^{8kR\delta/\sigma} \left(e^{4k^2K_3\delta^2/\sigma} + 2e^{k^2K_3\delta^2/\sigma} - 3 \right) + \left(e^{4kR\delta/\sigma} - 1 \right)^2 \right\}^{1/2} \\ &= O(\delta). \quad (123) \end{aligned}$$

Comme, pour δ suffisamment petit, il existe une constante K_8 telle que $|O(\delta)| \leq K_8|\delta|$, il s'ensuit que $p_*^t(x, \cdot)$ est K_8 -lipschitzienne sur B_ε , et ce pour tout $x \in \partial B_{R_1}$; on peut donc prendre $\kappa(t) = K_8$ pour $t \in]0, K_3]$.

Maintenant, pour traiter le cas où $t > K_3$, on va utiliser la relation de Chapman-Kolmogorov. L'astuce consiste à remarquer que, pour t fixé égal à K_3 , on peut obtenir une majoration de $p^t(x, y)$ indépendante de $x \in B_{R_2}$:

$$\begin{aligned} \forall x \in B_{R_2} \quad \forall y \in B_\varepsilon \quad |p_*^{K_3}(x, y') - p_*^{K_3}(x, y)| &\leq O(\delta) + \\ &\frac{1}{(2\pi\sigma K_3)^{d/2}} \left\{ \left(e^{\frac{4R_2\delta}{2\sigma K_3}} + e^{\frac{W\delta}{\sigma}} + e^{\frac{2Wk\delta}{\sigma}} - 3 \right) e^{4\frac{WR}{\sigma} + \frac{W^2K_3}{\sigma}} \right. \\ &\left. \sqrt{2} e^{4\frac{RW}{\sigma} + 2\frac{W^2K_3}{\sigma}} \left\{ e^{8\frac{kR\delta}{\sigma}} \left(e^{4\frac{k^2K_3\delta^2}{\sigma}} + 2e^{\frac{k^2K_3\delta^2}{\sigma}} - 3 \right) + \left(e^{4\frac{kR\delta}{\sigma}} - 1 \right)^2 \right\}^{1/2} \right\} \\ &= O(\delta). \quad (124) \end{aligned}$$

On écrit alors $t = (t - K_3) + K_3$ pour obtenir, par la relation de Chapman-Kolmogorov :

$$\begin{aligned} \forall t > K_3 \quad \forall x \in \partial B_{R_1} \quad \forall y, y' \in B_\varepsilon \\ |p_*^t(x, y') - p_*^t(x, y)| &= \left| \int_{B_{R_2}} p_*^{t-K_3}(x, x') \left(p_*^{K_3}(x', y') - p_*^{K_3}(x', y) \right) dx' \right| \\ &\stackrel{(54)}{\leq} K_1 e^{-K_2 t} \sup_{x' \in B_{R_2}} |p_*^{K_3}(x', y') - p_*^{K_3}(x', y)| \stackrel{(124)}{\leq} K_1 e^{-K_2 t} O(\delta). \quad (125) \end{aligned}$$

Puisque $|O(\delta)| \leq K_9|\delta|$ pour δ suffisamment petit, $p_*^t(x, \cdot)$ est $K_1K_9e^{-K_2t}$ -lipschitzienne sur B_ε pour tous $x \in \partial B_{R_1}$, $t > K_3$, et donc on peut prendre $\kappa(t) = K_1K_9e^{-K_2t}$ pour $t > K_3$. Finalement, on a bien obtenu une expression de $\kappa(t)$ vérifiant (37) et (38), ce qui achève notre démonstration du théorème 1.

Références

- [1] R. F. Bass, *Diffusions and Elliptic Operators*, Springer, 1998.
- [2] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of second order*, seconde édition, Springer-Verlag, 1983.
- [3] W. Feller, *An Introduction to probability theory and its applications*, J. Wiley, 1968.
- [4] D. Revuz & M. Yor, *Continuous Martingales and Brownian Motion*, Springer, 1991.

Document IV
Ramification des graphes

Ramification des graphes : arbres et graphes de Cayley

Rémi Peyre

23 novembre 2005

1 Résumé

Dans tout cet exposé, tous les graphes sont des graphes simples infinis supposés connexes et localement finis.

L'objectif de cet exposé consiste à étudier la vitesse à laquelle un graphe se ramifie lorsqu'on s'éloigne d'un point-origine o . Pour cela, on crée une marche aléatoire biaisée partant du point o , où le mobile a λ fois moins de chances d'emprunter une arête qui s'éloigne de o qu'une arête qui s'en rapproche ; et on se demande à partir de quelle valeur de λ la marche devient récurrente. Lorsque le graphe est un arbre, cette valeur coïncide avec un paramètre plus structural de l'arbre, le *nombre de branchement*, qui mesure la vitesse de ramification de l'arbre en termes de flots. Le *nombre de croissance* donne également une idée de la vitesse de ramification d'un arbre, et est plus facile à calculer en général, mais il ne constitue qu'une mesure grossière du nombre de branchement.

Nous allons néanmoins montrer que pour une certaine classe d'arbres, les arbres *sous-périodiques*, nombres de branchement et de croissance coïncident, ce qui facilitera le calcul du nombre de branchement.

Un graphe est dit *transitif* lorsque tous ses sommets « voient » la même chose. Une classe particulière de graphes transitifs est constituée par les *graphes de Cayley*, dont la structure dérive de celle d'un groupe. Nous allons voir que, pour les graphes de Cayley, on a un résultat du même type que pour les arbres sous-périodiques, à savoir que la valeur critique de λ pour la marche aléatoire biaisée s'obtient en calculant un nombre de croissance du graphe.

2 Arbres

2.1 Nombre de branchement

Soit A un arbre de racine o . La structure de l'arbre induit une notion naturelle de distance par rapport à o , aussi bien que les sommets que pour les arêtes : la distance d'un sommet s à o , notée $|s|$, est la longueur du plus court chemin de G reliant o à s , et la distance à o de l'arête a , notée $|a|$, est la longueur du plus court chemin partant de o dont la dernière arête est a (noter qu'en particulier, si a est adjacente à o , alors $|a| = 1$). On définit la *couronne* de rayon n , notée C_n , comme l'ensemble des arêtes situées à la distance n de o .

☞ *Toutes ces définitions s'étendent au cas d'un graphe quelconque G dans lequel on a privilégié un point-origine o .*

On cherche à contrebalancer la ramification de l'arbre lorsqu'on s'éloigne de l'origine. Deux méthodes sont envisageables :

1. On attribue à chaque arête une conductance $c(a) = \lambda^{-|a|}$, où λ est un paramètre ≥ 1 , et on fait partir la marche aléatoire associée de o ; on regarde à partir de quelle valeur de λ la marche devient récurrente. Notons λ_c^1 la valeur critique de λ associée.
2. On attribue à chaque arête une section $c(a) = \lambda^{-|a|}$, et on tente de faire passer un flot de o à l'infini; on regarde à partir de quelle valeur de λ cela n'est plus possible. Notons λ_c^2 la valeur critique de λ associée.

On a alors :

Théorème 1

$$\lambda_c^1 = \lambda_c^2$$

Définition 2 *On appelle nombre de branchement de l'arbre A et on note $br(A)$ la valeur commune de λ_c^1 et λ_c^2 .*

Démonstration : Supposons $\lambda < \lambda_c^1$. Alors il est possible de faire passer un courant de o à l'infini dans le réseau électrique A doté des conductances $\lambda^{-|a|}$. Considérons le courant qui passe lorsque les potentiels sont $V(o) = 1$ et $V(\infty) = 0$. On a alors pour toute arête a : $i(a) = (v(a^-) - v(a^+))c(a)$. Mais $0 \leq v(a^-), v(a^+) \leq 1$ d'où $|i(a)| \leq c(a) = \lambda^{-|a|}$. Comme i est un flot de o à l'infini, on a trouvé un flot de débit non nul et borné sur chaque arête par $\lambda^{-|a|}$, et donc $\lambda \leq \lambda_c^2$. Par conséquent, $\lambda_c^1 \leq \lambda_c^2$.

Réciproquement, supposons $\lambda < \lambda_c^2$. Il existe donc un flot de débit non nul de o à l'infini pour les sections $\lambda^{-|a|}$; soit θ un tel flot. Comme A est un arbre, il est clair qu'un tel flot peut être choisi de telle sorte que sur toute arête le flot s'éloigne de o . (en effet, on peut enlever les arêtes orientées à contresens sans

diminuer le débit). Soit $\lambda' < \lambda$. On a $0 \leq \theta(a) \leq \lambda^{-|a|} = (\frac{\lambda'}{\lambda})^{|a|} \lambda'^{-|a|}$. Calculons la puissance dissipée par ce flot dans le réseau électrique associé à la valeur λ' :

$$\begin{aligned} \mathcal{E}(\theta) &= \sum_{n \geq 1} \sum_{a \in C_n} \lambda'^n \theta(a)^2 \\ &\leq \sum_{n \geq 1} \sum_{a \in C_n} \left(\frac{\lambda'}{\lambda}\right)^n \theta(a) = \sum_{n \geq 1} \left(\frac{\lambda'}{\lambda}\right)^n \text{Débit}(\theta) = \text{Débit}(\theta) \sum_{n=1}^{\infty} \left(\frac{\lambda'}{\lambda}\right)^n < \infty \end{aligned}$$

Ainsi, il existe un flot de débit non nul et d'énergie finie dans le réseau associé à la valeur λ' , donc $\lambda' \leq \lambda_c^1$. On en déduit que $\lambda \leq \lambda_c^1$, puis que $\lambda_c^2 \leq \lambda_c^1$, ce qui achève la démonstration \square

Remarque 3 *On ne sait pas ce qui se passe à la valeur $\lambda = br(A)$. A priori, la marche aléatoire associée à la valeur $br(A)$ peut tout aussi bien être récurrente que transitoire, de même que le flot peut passer ou pas à la valeur critique. Tout ce qu'on sait par la démonstration ci-dessus, c'est que lorsque la marche aléatoire est transitoire pour une certaine valeur de λ , en l'occurrence pour $\lambda = br(A)$, alors le flot passe pour cette même valeur. Et de fait, à $\lambda = br(A)$ on peut trouver des exemples pour les trois autres cas possibles.*

2.2 Nombre de croissance

On constate aisément que le nombre de branchement de l'arbre k -aire est k . Cela suggère d'introduire la définition suivante :

Définition 4 *On appelle nombre de croissance de l'arbre A , et on note $cr(A)$, la limite de la suite $(\#C_n)^{1/n}$, si elle existe. (rappelons que C_n note l'ensemble des arêtes situées à la distance n de o). Plus généralement, on pose $\underline{cr}(A) = \liminf_{n \rightarrow \infty} (\#C_n)^{1/n}$ et $\overline{cr}(A) = \limsup_{n \rightarrow \infty} (\#C_n)^{1/n}$.*

Ainsi, pour l'arbre k -aire, nombres de croissance et de branchement coïncident.

Le nombre de croissance apparaît comme plus facile à calculer que le nombre de branchement, car il ne fait intervenir que la taille des couronnes, et pas la structure fine de l'arbre. Néanmoins, il existe des contre-exemples où $cr(A) \neq br(A)$. Notre travail va donc consister à rechercher des conditions suffisantes pour avoir l'égalité entre nombres de branchement et de croissance.

Notons d'abord la propriété suivante :

Proposition 5

$$br(A) \leq \underline{cr}(A)$$

Démonstration : Soit $\lambda > \underline{\text{cr}}(A)$; on va montrer que $\lambda \geq \text{br}(A)$. Puisque $\lambda > \underline{\text{cr}}(A)$, il existe une suite $(n_k)_{k \geq 1}$ avec $\#C_{n_k} \leq \lambda^{n_k}$. Comme les C_{n_k} sont des plans de coupe disjoints, le critère de Nash-Williams s'applique :

$$\begin{aligned} \mathcal{R}(a \leftrightarrow \infty) &\geq \sum_k \left(\sum_{a \in C_{n_k}} c(a) \right)^{-1} \\ &= \sum_k (\#C_{n_k} \lambda^{-n_k})^{-1} \geq \sum_{k=1}^{\infty} 1 = \infty \end{aligned}$$

Ce qui suffit pour conclure que la marche aléatoire associée à la valeur λ est récurrente \square

Remarque 6 *Lorsque le graphe est à symétrie sphérique par rapport à o , c'est-à-dire que tous les points situés à la même distance de o jouent le même rôle, on peut facilement démontrer que l'égalité est toujours atteinte dans la proposition précédente.*

2.3 Arbres périodiques et sous-périodiques

Nous allons maintenant nous intéresser à une forme d'invariance par translation des arbres.

Définition 7 *Pour s un sommet de l'arbre A d'origine o , le sous-arbre de A de racine s est l'arbre constitué par s et l'ensemble des points qui en descendent ; on le note A^s .*

Définition 8 *On dit qu'un arbre A est périodique lorsqu'il existe une partie finie S_0 de l'ensemble S des sommets telle que, pour tout $x \in S$, il existe $x_0 \in S_0$ avec $A^x \simeq A^{x_0}$. On dit que l'arbre est 0-périodique si on peut prendre $S_0 = \{o\}$ dans la phrase précédente.*

En plus d'avoir des propriétés d'invariance par translation intéressantes, les arbres périodiques peuvent s'interpréter comme le revêtement orienté de multigraphes finis : on appelle *revêtement orienté* d'un multigraphe¹ orienté fini M l'arbre constitué par les chemins finis du multigraphe issus d'un point-origine o de M .

On démontre alors facilement :

Proposition 9 *Le revêtement orienté d'un multigraphe orienté fini est un arbre périodique ; réciproquement, tout arbre périodique est le revêtement orienté d'un multigraphe orienté fini (il n'y a pas unicité).*

1. rappelons qu'un *multigraphe* est un graphe qui contient éventuellement des arêtes multiples et des boucles.

La proposition 9 permet donc de ramener l'étude d'un arbre périodique à celle d'un graphe orienté fini. En particulier, pour A un arbre périodique, on peut démontrer que si A est associé au multigraphe M , et que tous les points de M peuvent être atteints à partir de o , alors le nombre de croissance de A est le rayon spectral de la matrice d'adjacence de M — qu'on sait calculer facilement.

On définit également la notion plus générale de graphe *sous-périodique* :

Définition 10 *On dit qu'un arbre A est sous-périodique lorsqu'il existe une partie finie S_0 de l'ensemble S des sommets telle que, pour tout $x \in S$, il existe $x_0 \in S_0$ tel qu'on ait un morphisme de graphes injectif de A^x dans A^{x_0} . On dit que l'arbre est 0-sous-périodique si on peut prendre $S_0 = \{o\}$ dans la phrase précédente.*

Le théorème essentiel sur les arbres sous-périodiques est le suivant :

Théorème 11 *Pour un arbre sous-périodique, le nombre de croissance existe et vérifie $cr(A) = br(A)$.*

Démonstration :

On sait déjà que $br(A) \leq \underline{cr}(A)$ par la proposition 5, il ne reste donc qu'à montrer que $\overline{cr}(A) \leq br(A)$.

☞ *Pour commencer, on suppose l'arbre 0-sous-périodique et sans feuille*².

Fixons $\lambda > br(A)$ et montrons que $\overline{cr}(A) \leq \lambda$. Dans le graphe correspondant à la valeur de λ , il existe un plan de coupe Π_0 telle que la section de Π_0 , notée $\|\Pi_0\|_\lambda$, soit inférieure ou égale à 1. Quitte à remplacer Π_0 par son minimalisé³, on peut supposer que Π_0 est fini. Soit alors d la distance à la racine de l'arête le plus en aval de Π_0 .

On va montrer que, pour tout n , on peut construire à partir de Π_0 un plan de coupe Π_n de section ≤ 1 dont toutes les arêtes a vérifient $n < |a| \leq n + d$. Pour ce faire, considérons un plan de coupe Π et une arête a de Π . Notons x le point-fils de a . Alors le rôle de l'arête a dans le plan de coupe Π consiste à isoler de o les descendants lointains de x . Pour ce faire, on peut tout aussi bien remplacer a par un plan de coupe de l'arbre A^x ; en particulier, comme A^x est s'injecte morphiquement dans A , on peut prendre pour isoler les descendants de x de la racine le plan de coupe translaté de Π_0 dans l'arbre A^x . Le point consiste alors à remarquer que, comme $\|\Pi_0\|_\lambda \leq 1$, le translaté de Π_0 aura une section $\leq \lambda^{-|x|}$, et donc qu'on n'augmente pas la section de Π en remplaçant a par le translaté de Π_0 dans l'arbre A^x .

2. c'est-à-dire que tout sommet de l'arbre a au moins un descendant.

3. on rappelle que le *minimalisé* d'un plan de coupe Π est le bord de la composante connexe de o délimitée par ce plan de coupe ; c'est aussi un plan de coupe, et c'est un sous-ensemble fini de Π .

Cette constatation établie, il n'y a plus qu'à appliquer l'opération décrite ci-dessus aux arêtes de Π les plus proches de o et à recommencer jusqu'à ce que toutes les arêtes du plan de coupe obtenu aient une distance à o qui soit $> n$, on obtient alors un plan de coupe dont toutes les arêtes sont à une distance de o comprise entre n et $n + d$.

Mais alors, il y a au moins $\#C_n$ arêtes dans le plan de coupe Π_n , car il faut une arête de Π_n différente pour isoler de la racine les descendants de chaque arête de C_n (en effet, d'une part l'arbre est supposé sans feuille, donc il faut au moins une arête pour isoler de o les descendants d'un point, d'autre part les arêtes de Π_n sont situées au-delà de la couronne C_n , donc il en faut deux différentes pour isoler les descendants de deux arêtes de C_n distinctes). On a donc :

$$1 \geq \|\Pi_n\|_\lambda \geq \#C_n \lambda^{-(n+d)}$$

d'où $(\#C_n)^{1/n} \leq \lambda^{1+d/n}$, et en passant à la limite :

$$\overline{\text{cr}}(A) \leq \lambda$$

ce qu'on voulait.

On traite maintenant le cas où l'arbre a des feuilles.

Si l'arbre A a des feuilles, notons A' l'arbre A dans lequel on a placé un « rai » au bout de chaque feuille, c'est-à-dire une branche infinie sans ramification. On obtient alors un arbre sans feuille de même nombre de branchement (car il n'est pas possible de faire passer un flot dans un rai dès lors que $\lambda > 1$), et qui est encore 0-sous-périodique. Alors, en appliquant le résultat précédent :

$$\overline{\text{cr}}(A) \leq \overline{\text{cr}}(A') = \text{br}(A') = \text{br}(A)$$

d'où le résultat pour les arbres 0-périodiques à feuilles.

On démontre maintenant le cas général.

Soit A un arbre sous-périodique quelconque. On sait qu'il existe S_0 une sous-famille finie de S tels que les A^t , $t \in S_0$ contiennent à morphisme près tous les A^s , $s \in S$. Soit une telle famille S_0 . On considère alors l'arbre Γ constitué par le copie disjointe de tous les A^t , $t \in S_0 \cup \{o\}$ avec leur racines identifiées. Γ est alors un graphe 0-sous-périodique et donc $\text{br}(\Gamma) = \text{cr}(\Gamma)$.

Or on a $\text{br}(\Gamma) = \text{br}(A)$. Pour voir cela, posons $N = \max_{t \in S_0} |t|$ et considérons les sections associées à une valeur de $\lambda > \text{br}(A)$. Alors, pour tout $\varepsilon > 0$, il existe un plan de coupe Π de A de section $\leq \varepsilon$, en vertu du théorème maxiflot-minicoupe. On peut par ailleurs supposer que Π ne contient aucune arête de distance à $o \leq N$. Mais alors la restriction de Π à A^t , $t \in S_0$ est un plan de coupe de A^t , donc on peut transporter ce plan de coupe dans l'arbre Γ où il aura alors une section $\leq \lambda^N \varepsilon$. En faisant de même pour tous les points de

$S_0 \cup \{o\}$, on obtient un plan de coupe Π' de Γ de section

$$\|\Pi'\|_\lambda \leq (\#C_n + 1)\lambda^N \varepsilon$$

Par conséquent,

$$\inf_{\Pi' \text{ plan de coupe de } \Gamma} \|\Pi'\|_\lambda = 0$$

et donc $\lambda \geq \text{br}(\Gamma)$ en vertu du théorème maxiflot-minicoupe. Ainsi, $\text{br}(\Gamma) \leq \text{br}(A)$, et l'inégalité inverse est triviale.

Alors

$$\overline{\text{cr}}(A) \leq \overline{\text{cr}}(\Gamma) = \text{br}(\Gamma) = \text{br}(A)$$

ce qu'on voulait \square

Remarque 12 *En modifiant légèrement la première partie de la démonstration, on peut affiner le résultat : pour $\lambda = \text{br}(\Gamma)$, le flot passe, et même il existe un flot de débit ≥ 1 .*

3 Graphes de Cayley

3.1 Préambule : graphes transitifs

Nous venons d'étudier les arbres, qui sont des graphes possédant une propriété d'ordre. Nous allons maintenant étudier des graphes ayant une propriété d'invariance. On définit les graphes *transitifs* comme des graphes dans lesquels tous les sommets jouent le même rôle :

Définition 13 *Un graphe G est dit transitif lorsque, pour tous $s, t \in G$, il existe un automorphisme de G qui envoie s sur t .*

Dans un graphe transitif, on peut donc particulariser n'importe quel point o comme point-origine, ce que nous ferons toujours par la suite ; le choix de o n'a aucune importance.

3.2 Graphes de Cayley

Une classe importante de graphes transitifs est constituée par les *graphes de Cayley* : pour G un groupe infini, et B une partie génératrice finie de G stable par inversion et ne contenant pas e , on définit le graphe de Cayley sur G associé à B comme le graphe dont les sommets sont les points de G et où l'arête $a = [xy]$ appartient au graphe si et seulement si $x^{-1}y \in B$, c.à.d. si et seulement s'il existe $s \in S_0$ tel que $y = xs$. Il est clair qu'on définit ainsi un graphe transitif infini, connexe, localement fini. L'origine d'un tel graphe sera toujours prise en l'élément neutre e de G .

Bien qu'il existe des contre-exemples, un très grand nombre de graphes transitifs peuvent s'écrire comme des graphes de Cayley, ce qui justifie l'étude

de ces graphes.

La notion de nombre de branchement ne s'étend pas aux graphes autres que les arbres, car le théorème 1 ne s'y applique pas. (en effet, sa démonstration utilise l'orientation naturelle des arêtes de l'arbre, ce qui n'a pas de sens en général). Néanmoins, il demeure intéressant d'étudier les marche aléatoires biaisées partant de e et en particulier de calculer la valeur de λ_c^1 — que nous noterons ici simplement λ_c .

Il existe un analogue du théorème 11 pour les graphes de Cayley. Pour cela, on définit la *sphère de rayon n* d'un graphe de Cayley G , notée S_n , comme l'ensemble des sommets situés à la distance n de e . On définit alors le *nombre de croissance* du graphe de Cayley par :

Définition 14

$$cr(G) = \lim_{n \rightarrow \infty} (\#S_n)^{1/n}$$

Cela sous-entend :

Proposition 15 *La suite $((\#S_n)^{1/n})_{n \geq 1}$ converge.*

Démonstration : C'est une application du lemme de Fekete⁴ : on a en effet $\#S_{n+m} \leq \#S_n \#S_m$ à cause de la propriété de transitivité du graphe. (en effet, pour atteindre un point situé à la distance $n+m$ de e , on peut d'abord chercher un point à distance n de e , puis un point à distance m de ce point intermédiaire — qui « voit » la même chose que e) \square

Avis 16 *On peut aussi exprimer $cr(G)$ en termes d'arêtes : notant C_n l'ensemble des arêtes situées à la distance n de o , on a $cr(G) = \lim_{n \rightarrow \infty} (\#C_n)^{1/n}$*

Preuve : Si β note $\#B$, on voit que $\#C_n \leq \beta \#S_{n-1}$ et $\#S_n \leq 2\#C_n$, d'où le résultat \square

L'équivalent de la proposition 5 est donnée par la

Proposition 17

$$\lambda_c \leq cr(G)$$

Démonstration : Il n'y a qu'à recopier la démonstration de la proposition 5 \square

Les jalons sont maintenant posés pour démontrer le théorème-phare de cet exposé :

Théorème 18 *Si G est un graphe de Cayley, alors $\lambda_c = cr(G)$.*

4. Le lemme de Fekete énonce que, si $(\mu_n)_{n \geq 1}$ est une suite de valeurs réelles positives sous-multiplicative, c.à.d. telle que $\mu_{n+m} \leq \mu_n \mu_m$ pour tous n, m , alors μ_n converge quand $n \rightarrow \infty$.

Démonstration : On va construire un arbre sous-périodique qui approche en un certain sens le graphe de Cayley. Pour ce faire, on classe par ordre alphabétique les éléments de B . Comme B est un système générateur de G , tout élément de G peut s'écrire comme un produit d'éléments de B . On retient comme telle écriture d'un élément g son *écriture minimale*, à savoir l'écriture qui fait intervenir le moins de facteurs possibles, et où, parmi les ex-æquo éventuels, on retient celui qui vient le premier dans l'ordre alphabétique. On définit alors l'*arbre minimal* du graphe de Cayley comme le graphe A constitué par les sommets de G , où deux sommets sont reliés si et seulement si l'écriture minimale de l'un se déduit de celle de l'autre en rajoutant une lettre à droite.

On vérifie alors immédiatement :

Lemme 19 *A est un arbre 0-sous-périodique d'origine e ; et c'est un arbre couvrant de G , c.à.d. que tous les sommets de G sont dans A .*

D'autre part, la distance d'un point à l'origine est la même dans A que dans G , et par conséquent le cardinal de la couronne de rayon n de A , qui aussi le nombre de points à la distance n de e dans A , est égal au cardinal de la sphère de rayon n de G . On donc :

Lemme 20

$$cr(G) = cr(A)$$

On peut alors conclure : A est un sous-réseau électrique de G^5 , donc $\lambda_c^1(A) \leq \lambda_c(G)$ en vertu du critère de monotonie de Rayleigh. On peut donc écrire :

$$\lambda_c(G) \leq cr(G) = cr(A) = br(A) = \lambda_c^1(A) \leq \lambda_c(G)$$

et donc $\lambda_c(G) = cr(G)$, CQFD \square

Références

- [1] Russel LYONS, *Probabilities on trees and networks*, version du 17 avril 2005, §2.7 à 2.9.
<http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html> ,
 © Russel Lyons & Yuval Peres.

5. noter que la distance à e des arêtes est la même dans A et dans G , donc les conductances sont bien les mêmes.

Document V

Approche probabiliste du problème d'équilibre

Approche probabiliste du problème d'équilibre

Rémi Peyre

3 avril 2006

1 Introduction

Nous avons au cours des exposés précédents introduit un semigroupe sous-markovien d'opérateurs $R_t^{U,V}$ sur les espaces $L^p(U)$, $1 \leq p \leq \infty$, où U est un domaine de \mathbb{R}^d et V une fonction potentiel positive dans un espace fonctionnel judicieux K_d^{loc} . Nous allons maintenant nous intéresser à une approche probabiliste de ce semigroupe, qui nous permettra de comprendre en quoi il est intéressant et comment l'interpréter. En particulier, nous allons donner une interprétation probabiliste au problème d'équilibre, qui lors des exposés précédents n'avait été abordé que sous l'angle différentiel.

Nous montrerons également le lien entre les définitions différentielle et probabiliste du problème d'équilibre ; définirons la notion de capacité et en étudierons une application au calcul du volume de la saucisse de Wiener.

2 Rappels

2.1 Hypothèses et notations

$d \geq 1$ est un entier naturel, et U un domaine de \mathbb{R}^d , i.e. un ouvert non vide et connexe.

$((Z_t)_{t \geq 0}, W_x)$ est un mouvement brownien standard de \mathbb{R}^d partant de x . On note $p(t, x, y)$ la densité de probabilité de transition du mouvement brownien : $dW_x(Z_t = y) = p(t, x, y)dy$ ¹. On note $W_{x,y}^t$ la loi de W_x conditionnée par rapport à l'événement $Z_t = y$.

Pour B un fermé de \mathbb{R}^d , on introduit $H_B = \min\{s \geq 0; Z_s \in B\}$ qui est le *temps d'entrée* dans B . On note également $\tilde{H}_B = \min\{s > 0; Z_s \in B\}$ le *temps de retour* en B , ainsi que pour $r > 0$: $H_r = \min\{s \geq 0; |Z_s - Z_0| \geq r\}$. Pour un

¹On connaît la valeur explicite de p : $p(t, x, y) = (2\pi t)^{-d/2} e^{-|x-y|^2/2t}$.

ouvert U , on introduit de même le *temps de sortie de U* : $T_U = \min\{s \geq 0; Z_s \notin U\} = H_{cU}$.

On définit sur U un *potentiel*, noté V ; c'est une fonction mesurable, positive et qui vérifie :

$$\forall N < \infty \sup_{|x| \leq N} \mathbb{E}^{W_x} \left[\int_0^{H_r} V(Z_s) ds \right] \xrightarrow{r \rightarrow 0} 0 \quad (1)$$

On écrira pour simplifier $V \in K_d^{\text{loc}}$. La condition $V \in K_d^{\text{loc}}$ est une condition d'intégrabilité *ad hoc* pour l'étude qui suit; elle donne en particulier :

Lemme 1.

$$W_x\text{-p.s.} \quad \forall t < \infty \quad \int_0^t V(Z_s) ds < \infty$$

On définit alors l'opérateur $R_t^{U,V}$ par

$$R_t^{U,V} f(x) = \mathbb{E}^{W_x} \left[f(Z_t) \exp \left\{ - \int_0^t V(Z_s) ds \right\} \mathbb{1}_{T_U > t} \right] \quad (2)$$

Il est aisé de se rendre compte que $R_t^{U,V}$ est un opérateur à densité : on a

$$R_t^{U,V} f(x) = \int f(y) r_{U,V}(t, x, y) dy \quad (3)$$

où $r_{U,V}(t, x, y)$ est défini par

$$r_{U,V}(t, x, y) = p(t, x, y) \mathbb{E}^{W_{x,y}^t} \left[\exp \left\{ - \int_0^t V(Z_s) ds \right\} \mathbb{1}_{T_U > t} \right] \quad (4)$$

On utilisera aussi l'intégrale en t de $r_{U,V}$, notée $g_{U,V}$:

$$g_{U,V}(x, y) = \int_0^\infty r_{U,V}(s, x, y) ds \quad (5)$$

$g_{U,V}$ s'appelle la *fonction de Green* associée au domaine U et au potentiel V .

2.2 Résultats sur les fonction de Green

Les exposés précédents nous ont donné plusieurs résultats sur les fonctions de Green, qui nous seront très utiles par la suite. Notons d'abord la propriété immédiate suivante :

Proposition 2. $r_{U,V}(t, x, y)$ et $g_{U,V}(x, y)$ sont symétriques par rapport à x et y .

On a vu également que, du fait que U est connexe :

Proposition 3.

$$\begin{aligned} \forall t > 0 \quad \forall z, z' \in U \quad r_{U,V}(t, z, z') > 0 \\ \forall z, z' \in U \quad g_{U,V}(z, z') > 0 \end{aligned}$$

On a ensuite un résultat de continuité sur $g_{U,V}$:

Théorème 4. $g_{U,V}$ est continue sur $U \times U$ (à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$).

☞ **On supposera par ailleurs tout au long de cet exposé :**

Hypothèse 5. $\exists z_0, z'_0 \in U \quad g_{U,V}(z_0, z'_0) < \infty$

Cette hypothèse est plus puissante qu'il n'y paraît, puisqu'on a :

Théorème 6. Sous l'hypothèse 5, alors pour tous $z, z' \in U$ avec $z \neq z'$, on a $g_{U,V}(z, z') < \infty$. Plus précisément, on a : $g_{U,V}(z, z') = \infty \Leftrightarrow (z = z' \text{ et } d \geq 2)$.

Enfin, $g_{U,V}$ satisfait l'identité de la résolvante :

$$g_{U,V+\lambda}(x, y) = g_{U,V+\mu}(x, y) + (\mu - \lambda) \int_U g_{U,V+\lambda}(x, z) g_{U,V+\mu}(z, y) dz \quad (6)$$

pour tous $\lambda, \mu \geq 0$.

2.3 Objectif

L'objet de notre étude est le *problème d'équilibre* : pour $K \subset U$ un compact, on considère la fonction v continue sur \mathbb{R}^d qui satisfait sur $U \setminus K$ l'équation différentielle :

$$\frac{1}{2} \Delta v - Vv = 0 \quad (7)$$

avec pour conditions aux bords

$$\begin{cases} v \equiv 1 & \text{sur } K \\ v \equiv 0 & \text{sur } {}^cU \end{cases} \quad (8)$$

Résoudre le problème d'équilibre signifie alors mettre v sous la forme :

$$v(x) = \int_K g(x, y) \mu(dy) \quad (9)$$

pour x dans U , avec μ une mesure finie sur K .

Au cours des exposés précédents nous avons résolu ce problème dans le cas où K et U étaient lisses et V de classe C^∞ , à l'aide du calcul différentiel. Notre objectif est d'interpréter cette relation dans un cadre plus général, par une approche probabiliste.

3 Marche aléatoire associée au problème

3.1 Définition du mouvement brownien en terrain miné

L'objet important pour la compréhension de notre opérateur $R_t^{U,V}$ est le « mouvement brownien en terrain miné ». Intuitivement, il s'agit d'un mouvement brownien ordinaire, mais dont la durée de vie est éventuellement finie : le mouvement brownien $(Z_t)_{t \geq 0}$ « meurt » avec une densité de probabilité $V(Z_t)$ lorsqu'il est dans U , et meurt instantanément dès qu'il sort de U . Cela se formalise de la façon suivante :

Définition 7. Soit ∂ un point-cimetière. On se place dans l'espace des trajectoires dans U à durée de vie éventuellement finie :

$$C_\partial(\mathbb{R}_+, U) = \left\{ w : \mathbb{R}_+ \rightarrow U \cup \{\partial\}; \exists \tau \in [0, \infty] \quad \begin{array}{l} w|_{[0, \tau[} \text{ est continue dans } U \\ w|_{[\tau, \infty[} \equiv \partial \end{array} \right\}$$

Le processus Z_t est le processus canonique $Z_t(w) = w(t)$. La loi \mathbb{P}_x sur $C_\partial(\mathbb{R}_+, U)$ est définie de la manière suivante : on considère l'espace produit $C(\mathbb{R}_+, \mathbb{R}^d) \times \mathbb{R}_+$, doté de la loi $\hat{W}_x = W_x \otimes e^{-\xi} d\xi$, i.e. on tire indépendamment un mouvement brownien $(Z_t)_{t \geq 0}$ issu de x et une variable exponentielle ξ de paramètre 1. On définit alors l'heure de la mort du mouvement brownien par

$$\tau((Z_t)_{t \geq 0}, \xi) = \min \left(T_U((Z_t)_{t \geq 0}), \inf \left\{ s \geq 0; \int_0^s V(Z_u) du \geq \xi \right\} \right)$$

La loi \mathbb{P}_x est alors définie comme l'image de \hat{W}_x par l'application mesurable $((Z_t)_{t \geq 0}, \xi) \mapsto (t \mapsto Z_t \cdot \mathbb{1}_{t < \tau} + \partial \cdot \mathbb{1}_{t \geq \tau})$.

Remarque 8. Grâce au lemme 1, on a $\int_0^s V(Z_u) du \xrightarrow{s \rightarrow 0} 0$ \mathbb{P}_x -p.s.. Comme la loi de ξ ne charge pas 0 et que, par continuité du mouvement brownien, on a $T_U > 0$ \mathbb{P}_x -p.s., on en déduit :

$$\forall x \in U \quad \mathbb{P}_x\text{-p.s.} \quad \tau > 0$$

☞ Dorénavant, on supposera U et V fixés, et on travaillera toujours sur le processus $(Z_t)_{t \geq 0}$ muni de la loi \mathbb{P}_x . Les symboles U et V seront enlevés lorsqu'il n'y aura pas d'ambiguïté.

Proposition 9. $(Z_t)_{t \geq 0}$ vérifie la propriété de Markov forte par rapport à la filtration naturelle de $(Z_t)_{t \geq 0}$, notée $(\mathcal{G}_t)_{t \geq 0}$, à savoir $\mathcal{G}_t = \sigma(Z_s; 0 \leq s \leq t)$.

Démonstration. La propriété de Markov n'est pas surprenante, compte tenu de la définition intuitive du mouvement brownien en terrain miné. Dans la mesure où la preuve est tout-à-fait similaire à la preuve de la propriété de Markov forte pour un mouvement brownien standard, nous ne la détaillerons pas ici □

D'autre part, on a le lien fondamental suivant entre $(Z_t)_{t \geq 0}$ et les opérateurs $R_t^{U,V}$:

Théorème 10. *Pour $f \in L^\infty(U)$, si on pose $f(\partial) = 0$, on a :*

$$\mathbb{E}_z[f(t)] = R_t^{U,V} f(z)$$

Démonstration. On a :

$$\begin{aligned} \mathbb{E}_z[f] &= \mathbb{E}^{\hat{W}_z} [f(Z_t) \mathbb{1}_{t < \tau}] = \mathbb{E}^{\hat{W}_z} \left[f(Z_t) \mathbb{1}_{t < T_U} \mathbb{1}_{\xi > \int_0^t V(Z_s) ds} \right] \\ &\stackrel{(\text{idpce})}{=} \mathbb{E}^{\hat{W}_z} \left[f(Z_t) \mathbb{1}_{t < T_U} \exp \left\{ - \int_0^t V(Z_s) ds \right\} \right] = R_t^{U,V} f(z) \quad (10) \end{aligned}$$

□

3.2 Propriété de transience

Pour un mouvement brownien général, on sait qu'on a récurrence (au sens topologique) en dimensions 1 et 2, et transience en dimension ≥ 3 . On peut se demander si la possibilité pour le mouvement brownien d'être tué au cours de son évolution lui interdit d'avoir un comportement récurrent, même en petite dimension. L'hypothèse 5 assure que c'est effectivement le cas :

Théorème 11. *Sous l'hypothèse 5, pour $K \subset U$ compact, pour $x \in U$ on a \mathbb{P}_x -p.s. $\exists T < \infty \quad \forall t > T \quad Z_t \notin K$.*

Remarque 12. *En dimension 1 ou 2, cela implique que le mouvement brownien se fait tuer avec probabilité 1 ; par contre, en dimension 3, le théorème ne nous apprend rien de plus que la théorie usuelle du mouvement brownien (il est clair que, si le mouvement brownien est transitoire sans être tué, il l'est a fortiori en étant tué).*

Démonstration. Le point-clef de la démonstration est la relation :

Lemme 13.

$$\forall x \in U \quad \int_K g(x, z) dz < \infty \quad (11)$$

Démonstration. Pour le voir, on applique l'indentité de la résolvante (formule (6)) avec $\lambda = 0$ et $\mu > 0$:

$$\begin{aligned} g_{U,V}(x, y) &= g_{U, V+\mu}(x, y) + \mu \int_K g_{U,V}(x, z) g_{U, V+\mu}(z, y) dz \\ &\geq \mu \cdot \inf_{z \in K} (g_{U, V+\mu}(z, y)) \cdot \int_K g_{U,V}(x, z) dz \quad (12) \end{aligned}$$

Mais $g_{U,V+\mu}$, comme toute fonction de Green, est strictement positive et continue dans $\mathbb{R} \cup \{+\infty\}$, donc l'infimum de $g_{U,V+\mu}(\cdot, y)$ sur le compact K est > 0 . On est donc en droit d'écrire :

$$\int_K g_{U,V}(x, z) dz \leq \frac{1}{\mu \cdot \inf_{z \in K} (g_{U,V+\mu}(z, y))} g_{U,V}(x, y)$$

Mais d'après l'hypothèse 5 et le théorème 6, on a $g_{U,V}(x, y) < \infty$ dès lors que $y \neq x$ (ce qui est possible puisque U ne peut être réduit à un point). Par conséquent $\int_K g_{U,V}(x, y) < \infty$ \square

L'idée est maintenant d'introduire un compact intermédiaire K' entre K et U tel que $K \subset \overset{\circ}{K}' \subset K' \subset U$; et de regarder les allées et venues du processus de part et d'autre de l'« anneau » $\overset{\circ}{K}' \setminus K$. On définit ainsi :

Définition 14.

$$\begin{aligned} R_1 &= \inf\{s \geq 0; Z_s \in K\} \\ D_1 &= \inf\{s \geq R_1; Z_s \notin \overset{\circ}{K}'\} \\ R_{n+1} &= \inf\{s \geq D_n; Z_s \in K\} \\ D_{n+1} &= \inf\{s \geq R_{n+1}; Z_s \notin \overset{\circ}{K}'\} \end{aligned}$$

R_n est appelé n -ième temps de retour en K ; D_n est le n -ième temps de départ de $\overset{\circ}{K}'$. Ces valeurs peuvent éventuellement être infinies à partir d'un certain rang.

On écrit alors :

$$\begin{aligned} \int_K g(x, z) dz &\stackrel{(\text{déf de } g)}{=} \int_K \left(\int_0^\infty r(t, x, z) dt \right) dz \\ &\stackrel{(\text{Fubini})}{=} \int_0^\infty \int_K r(t, x, z) dz dt \\ &\stackrel{(\text{thm 10})}{=} \int_0^\infty \mathbb{E}_x [\mathbb{1}_K(Z_s)] ds \\ &\stackrel{(\text{Fubini})}{=} \mathbb{E}_x \left[\int_0^\infty \mathbb{1}_K(Z_s) ds \right] = \mathbb{E}_x \left[\sum_{i \geq 1} \int_{R_i}^{D_i} \mathbb{1}_K(Z_s) ds \right] \\ &\stackrel{(\text{Fubini})}{=} \sum_{i \geq 1} \mathbb{E}_x \left[\mathbb{1}_{R_i < \infty} \int_{R_i}^{D_i} \mathbb{1}_K(Z_s) ds \right] \\ &\stackrel{(\text{Markov fort})}{=} \sum_{i \geq 1} \mathbb{E}_x \left[\mathbb{1}_{R_i < \infty} \mathbb{E}_{Z_{R_i}} \left[\int_0^{D_1} \mathbb{1}_K(Z_s) ds \right] \right] \quad (13) \end{aligned}$$

Maintenant, on va minorer $\mathbb{E}_{Z_{R_i}} \left[\int_0^{D_1} \mathbb{1}_K(Z_s) ds \right]$. Pour tout $z \in K^2$, en ob-

²Noter que pour $n \geq 2$, si $R_n < \infty$, on a $Z_{R_n} \in \partial K$ par continuité du mouvement brownien. Par contre, si le point de départ x du processus est dans $K \setminus \partial K$, alors $R_1 = 0$ et $Z_{R_n} = x \notin \partial K$.

servant que sous \mathbb{P}_z on a $D_1 = T_{\mathring{K}'}$, le théorème 10 assure que :

$$\mathbb{E}_z \left[\int_0^{D_1} \mathbb{1}_K(Z_s) ds \right] = \int_K g_{\mathring{K}', V}(z, z') dz'$$

qui est > 0 d'après la proposition 3, quitte à agrandir K pour qu'il soit de mesure non nulle. Or la fonction positive $z \mapsto g(z, z')_{\mathring{K}', V}$ est continue, donc par le lemme de Fatou $z \mapsto \int_K g_{\mathring{K}', V}(z, z') dz'$ est semi-continue inférieurement. Par conséquent, son infimum \mathcal{A} est atteint sur K , et donc est > 0 ; d'où en injectant ce résultat dans (13) :

$$\infty > \int_K g(x, z) dz \geq \mathcal{A} \sum_{i \geq 1} \mathbb{P}_x[R_i < \infty]$$

donc $\sum_{i \geq 1} \mathbb{P}_x[R_i < \infty] < \infty$, et on a par Borel-Cantelli :

$$\mathbb{P}_x - \text{p.s.} \quad \exists i \quad R_i = \infty$$

Cela suffit à conclure qu'au bout d'un certain temps le mobile quitte définitivement K , car sinon cela lui imposerait de rester indéfiniment dans \mathring{K}' à partir d'un certain moment, ce qui est presque impossible puisque K' est borné. \square

4 Problème d'équilibre

4.1 La dernière visite

Nous allons maintenant nous attaquer pour de bon au problème d'équilibre. Notre approche paraît peu orthodoxe : Au lieu de nous intéresser à la solution v de l'équation différentielle (7), nous allons *définir* v comme un objet probabiliste, puis résoudre le problème d'équilibre pour celui-ci, avant de constater *a posteriori* que ce v est bien celui qu'on cherche. En fait, il n'y a pas lieu de s'en étonner, puisque l'interprétation physique naturelle de (7) consiste précisément en ce qui va suivre.

Soit $K \subset U$ un compact non vide. Pour w une trajectoire du mouvement brownien en terrain miné, on définit

$$L_K(w) = \sup\{t \in]0, \infty[; Z_t(w) \in K\}$$

D'après le théorème 11, $L_K(w) < \infty$ p.s.. Le *potentiel d'équilibre* de K (relativement à U et V) est la fonction

$$x \in U \mapsto v(x) = \mathbb{P}_x[L_K > 0] = \mathbb{P}_x[\tilde{H}_K < \infty]$$

D'après la définition du mouvement brownien en terrain miné, cela revient à poser :

$$v(x) = \mathbb{E}^{W_x} \left[\mathbb{1}_{\tilde{H}_K < \infty} \exp \left\{ - \int_0^{\tilde{H}_K} V(Z_s) ds \right\} \right] \quad (14)$$

On introduit maintenant des fonctions ψ_ε qui mesurent la probabilité que la dernière visite en K soit tout proche :

$$\psi_\varepsilon(z) = \frac{\mathbb{P}_z[0 < L_K \leq \varepsilon]}{\varepsilon} \quad (15)$$

Ce sont ces fonctions qui vont nous permettre de résoudre le problème d'équilibre, puisque, comme nous allons le constater, les mesures $\psi_\varepsilon(z)dz$ convergent vers la mesure d'équilibre.

Nous devons également introduire la *distribution de dernière visite en K* : pour tout $x \in U$, elle est définie par

$$L(x, dy) = \mathbb{E}_x[L_K > 0 \text{ et } Z_{L_K^-} \in dy] \quad (16)$$

$L(x, \cdot)$ exprime la répartition de l'endroit où se trouve un mouvement brownien en terrain miné issu de x lorsqu'il touche K pour la dernière fois. C'est clairement une mesure de sous-probabilité, portée par K .

Remarque 15. *Noter qu'on ne prend pas la position au temps L_K , mais sa limite à gauche $Z_{L_K^-}$. En effet, il faut tenir compte du cas où le mobile est tué alors qu'il était dans K : dans ce cas en effet, on a $Z_{L_K} = \partial$, mais le « dernier point de K visité par le mobile » est $Z_{L_K^-}$.*

4.2 La mesure $e(dy)$

$L(x, \cdot)$ va nous permettre de comprendre vers quelle limite tend ψ_ε . On a en effet, notant $C_b(U)$ l'ensemble des fonctions sur U continues et bornées :

Lemme 16.

$$\forall f \in C_b(U) \quad \int_U g(x, y) \psi_\varepsilon(y) f(y) dy \xrightarrow{\varepsilon \rightarrow 0} \int_K L(x, dy) f(y)$$

Démonstration. On peut se ramener au cas $f \geq 0$. On écrit alors :

$$\begin{aligned} & \int_U g(x, y) \psi_\varepsilon(y) f(y) dy \stackrel{\text{(thm 10)}}{=} \int_0^\infty \mathbb{E}_x [f(Z_s) \psi_\varepsilon(Z_s)] ds \\ &= \frac{1}{\varepsilon} \int_0^\infty \mathbb{E}_x [f(Z_s) \mathbb{E}_{Z_s}[0 < L_K \leq \varepsilon]] ds \\ & \stackrel{\text{(Markov)}}{=} \frac{1}{\varepsilon} \int_0^\infty \mathbb{E}_x [f(Z_s) \mathbb{1}_{s < L_K \leq s + \varepsilon}] ds \\ &= \mathbb{E}_x \left[\mathbb{1}_{0 < L_K \leq \varepsilon} \cdot \frac{1}{\varepsilon} \int_0^{L_K} f(Z_s) ds \right] + \mathbb{E}_x \left[\mathbb{1}_{L_K > \varepsilon} \cdot \frac{1}{\varepsilon} \int_{L_K - \varepsilon}^{L_K} f(Z_s) ds \right] \\ & \stackrel{\text{(CV dominée)}}{\xrightarrow{\varepsilon \rightarrow 0}} 0 + \mathbb{E}_x [\mathbb{1}_{L_K > 0} f(Z_{L_K^-})] \\ &= \int_K L(x, dy) f(y) \end{aligned}$$

CQFD

□

Nous en déduisons la convergence étroite des ψ_ε :

Théorème 17.

$$\psi_\varepsilon(y)dy \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{g(x, y)}L(x, dy) \quad (17)$$

x étant un élément quelconque de U .

On note $e(dy)$ la limite des $\psi_\varepsilon(y)dy$.

Démonstration. D'après le paragraphe 2.2, $g(x, y)$ est une fonction continue à valeurs dans $]0, +\infty]$, finie pour $y \neq x$. Pour f continue à support compact dans U , $\frac{1}{g(x, y)}f(y)$ est donc également continue à support compact ; d'après le lemme 16, on a donc

$$\begin{aligned} \int_U f(y)\psi_\varepsilon(y)dy &= \int_U g(x, y)\psi_\varepsilon(y) \left(\frac{1}{g(x, y)}f(y) \right) dy \\ &\xrightarrow{\varepsilon \rightarrow 0} \int_K \frac{1}{g(x, y)}f(y)L(x, dy) \end{aligned}$$

ce qui est bien la convergence étroite énoncée □

Remarque 18. *Attention ! La relation $g(x, y)\frac{1}{g(x, y)} = 1$ n'est pas automatique, car $g(x, y)$ peut être infini. Néanmoins on a $g(x, y)\frac{1}{g(x, y)} = 1$ pour tout $y \neq x$ et donc y -presque partout, ce qui justifie la première égalité ci-dessus.*

Proposition 19. $e(dy)$ est une mesure finie portée par K .

Démonstration. $L(x, dy)$ est portée par K , donc $e(dy)$ également. Pour voir que e est finie, soit $x \in U$ arbitraire. La mesure $L(x, dy)$ est finie car c'est une sous-probabilité. De plus, la fonction $\frac{1}{g(x, y)}$ est finie et continue sur le compact K finie, donc bornée. Par conséquent, $e(dy) = \frac{1}{g(x, y)}L(x, dy)$ est une mesure finie. □

4.3 Solution du problème d'équilibre

La mesure e résout le problème d'équilibre :

Théorème 20. 1.

$$\forall x \in U \quad v(x) = \int_K g(x, y)e(dy) \quad (18)$$

2. $e(dy)$ est la seule mesure sur K solution du problème d'équilibre (18).

$e(dy)$ est appelée *mesure d'équilibre* de K (relativement à U et V) ; la masse totale de K pour e est appelée *capacité* de K .

Démonstration. (1). La relation (18) s'obtient en intégrant par rapport à y la relation suivante :

$$\forall x \in U \quad g(x, y)e(dy) = L(x, dy) \quad (19)$$

Cette relation semble immédiate de par l'équation (17) qui définit $e(dy)$, mais comme nous le rappelait déjà la remarque 18, il faut prendre garde à ce que g est infini pour $y = x$ et $d \geq 2$. On a donc besoin de :

Lemme 21. *Si $d \geq 2$ alors $\forall x \in K \quad L(x, \{x\}) = 0$.*

Démonstration. Ce lemme est évident en revenant à la définition de $L(x, dy)$. En effet, pour $d \geq 2$ on a une propriété de transience au sens strict (i.e. non topologique) pour le mouvement brownien standard qui nous dit que \mathbb{P}_x -p.s. $\forall t > 0 \quad Z_{t-} \neq x$ et qui vaut *a fortiori* pour le mouvement brownien en terrain miné. Par conséquent, $L(x, \{x\}) = \mathbb{P}_x(L_K > 0 \text{ et } Z_{L_K-} = x) = 0$ et le lemme est démontré. D'où (1). \square

(2). Soit $\tilde{e}(dy)$ une mesure sur K satisfaisant (18). Déjà, \tilde{e} est une mesure finie. En effet, pour $x \in U$ arbitraire, $g(x, y)$, qui est strictement positive et continue, est minorée par un $m > 0$ sur K . Par conséquent,

$$\tilde{e}(K) \leq \frac{1}{m} \int_K g(x, y)\tilde{e}(dy) \stackrel{(\text{par hyp.})}{=} \frac{1}{m} v(x) \leq \frac{1}{m} < \infty$$

Soit maintenant f continue et bornée sur U . Pour $\mu > 0$, on définit

$$G_\mu f(z) = \int_U g_{U, V+\mu}(z, z')f(z')dz' \\ \stackrel{(\text{thm } 10)}{=} \mathbb{E}^{W_z} \left[\int_0^{T_U} f(Z_s) \exp \left\{ -\mu s - \int_0^s V(Z_u)du \right\} ds \right]$$

$G_\mu f$ apparaît comme une régularisée de f : en effet, on a

$$\forall z \in U \quad \mu G_\mu f(z) \xrightarrow{\mu \rightarrow \infty} f(z)$$

car on a

$$f(Z_s) \exp \left\{ - \int_0^s V(Z_u)du \right\} ds \xrightarrow{s \rightarrow 0} f(z) \quad W_z\text{-p.s.}$$

grâce au lemme 1. Soit $x \in U$. On a par hypothèse

$$\int_K g(x, y)\tilde{e}(dy) = \int_K g(x, y)e(dy)$$

d'où, en appliquant à g l'identité de la résolvante avec $\lambda = 0$:

$$\int_K \left(g_{U, V+\mu}(x, y) + \mu \int_U g_{U, V+\mu}(x, z)g_{U, V}(z, y)dz \right) \tilde{e}(dy) = \\ \int_K \left(g_{U, V+\mu}(x, y) + \mu \int_U g_{U, V+\mu}(x, z)g_{U, V}(z, y)dz \right) e(dy)$$

ce qui se réécrit, en appliquant le théorème de Fubini et en utilisant à nouveau (18) :

$$\begin{aligned} \int_K g_{U,V+\mu}(x, y) \tilde{e}(dy) + \mu \int_U g_{U,V+\mu}(x, z) v(z) dz \\ = \int_K g_{U,V+\mu}(x, y) e(dy) + \mu \int_U g_{U,V+\mu}(x, z) v(z) dz \end{aligned}$$

Donc en simplifiant

$$\int_K g_{U,V+\mu}(x, y) \tilde{e}(dy) = \int_K g_{U,V+\mu}(x, y) e(dy) \quad (20)$$

On en tire

$$\begin{aligned} \int_K \mu G_\mu f(y) \tilde{e}(dy) \stackrel{(\text{déf de } G_\mu)}{=} \mu \int_K \int_U g_{U,V+\mu}(y, x) f(x) \tilde{e}(dy) dx \\ \stackrel{(\text{Fubini} + (20))}{=} \mu \int_K \int_U g_{U,V+\mu}(y, x) f(x) e(dy) dx = \int_K \mu G_\mu f(y) e(dy) \end{aligned}$$

d'où par convergence dominée³ quand $\mu \rightarrow \infty$:

$$\int_K f(y) \tilde{e}(dy) = \int_K f(y) e(dy)$$

et ce pour toute f continue bornée, d'où $\tilde{e} = e$ □

4.4 Un résultat plus précis

On peut raffiner le résultat donné par (18) :

Théorème 22.

$$\mathbb{P}_x(L_K \in ds \text{ et } Z_{L_K-} \in dy) = r(s, x, y) e(dy) ds$$

³noter en effet que $\|\mu G_\mu f\|_\infty \leq \|f\|_\infty$.

Démonstration. Soit $t \geq 0$ et $f \in L^\infty(U)$ une fonction-test. On a :

$$\begin{aligned}
\mathbb{E}_z[\mathbb{1}_{L_K > t} f(Z_{L_K^-})] &\stackrel{\text{(Markov)}}{=} \mathbb{E}_x [E_{Z_t}[\mathbb{1}_{L_K > 0} f(Z_{L_K^-})]] \\
&= \mathbb{E}_x \left[\int_K L(Z_t, dy) f(y) \right] \\
&\stackrel{(19)}{=} \mathbb{E}_x \left[\int_K g(Z_t, y) f(y) e(dy) \right] \\
&\stackrel{\text{(thm 10)}}{=} \int_U r(t, x, z) \int_K g(z, y) f(y) e(dy) dz \\
&= \int_U \int_K \int_0^\infty r(t, x, z) r(u, z, y) f(y) e(dy) dz du \\
&\stackrel{v=t+u}{=} \int_t^\infty \int_K \left(\int_U r(t, x, z) r(v-t, z, y) dz \right) f(y) e(dy) dv \\
&\stackrel{\text{(semigroupe)}}{=} \int_t^\infty \int_K r(v, x, y) f(y) e(dy) dv
\end{aligned}$$

ce qu'on voulait □

5 Lien avec les ÉDP

5.1 Énoncé du théorème

Le « potentiel d'équilibre » $v(x)$ que nous venons de définir peut également être caractérisé par une équation aux dérivées partielles, ce qui montre le lien entre le problème d'équilibre différentiel résolu dans les exposés précédents et le problème probabiliste soulevé ici. L'intérêt de travailler avec des concepts probabilistes, outre le nouvel angle d'attaque du problème que cela procure, est que cela dispense des conditions de régularité qui étaient demandées précédemment.

Théorème 23. *Supposons V continue. Alors*

1. *Sur $U \setminus K$, v est de classe C^2 et*

$$\frac{1}{2} \Delta v = Vv \tag{21}$$

2. *En les points réguliers de K , resp. de ${}^c U$, v est continue de valeur 1, resp. tend vers 0.*

Démonstration. 5.2 L'identité différentielle

Montrons d'abord que v est continue sur $U \setminus K$. Soit $x \in U \setminus K$, soit O un voisinage de x avec \bar{O} compact et contenu dans $U \setminus K$. On a :

$$v(x) = \int_K g(x, y) e(dy)$$

avec g continue sur $\bar{O} \times K$, et finie (car $\bar{O} \cap K = \emptyset$), donc bornée par compacité, ce qui donne la continuité de l'intégrale par convergence dominée.

Soit L l'opérateur défini sur $C_c^\infty(U)$ par :

$$Lf = \frac{1}{2} \Delta f - Vv$$

On peut montrer sans grande difficulté que L est le générateur du semigroupe $(R_t^{U,V})_{t \geq 0}$, ce qui ici signifiera simplement que pour $\varphi \in C_c^\infty(U)$, prolongée par 0 en ∂ , on a :

$$\frac{d}{dt} \Big|_{t=0} \mathbb{E}_x[\varphi(X_t)] = L\varphi(x)$$

Remarque 24. Ici nous ne considérons que l'action de $R_t^{U,V}$ sur des fonctions de $C_c^\infty(U)$. En réalité, comme nous l'avons vu dans les exposés précédents, le cadre naturel pour étudier l'action du semigroupe est de se placer sur les fonctions de $L^2(U)$. Dans ce cas, l'idée générale reste valable, mais d'un point de vue technique on affirme l'existence d'une extension \hat{L} de L opérant sur $L^2(U)$ tout entier telle que pour toute fonction φ de $L^2(U)$, on ait :

$$\frac{R_t^{U,V} \varphi - \varphi}{t} \xrightarrow[t \rightarrow 0]{L^2(U)} \hat{L}\varphi$$

Remarque 25. Le formalisme de la remarque ci-dessus permet également de définir un générateur \hat{L} de $(R_t^{U,V})_{t \geq 0}$ prolongeant L dans le cas où on a seulement $V \in L_{loc}^2 \cap K_d^{loc}$, ce qui permet d'étendre le théorème 23 dans le cas général $V \in K_d^{loc}$. Néanmoins, les outils techniques mis en jeu étant un peu compliqués, nous n'aborderons pas cette question ici.

On a le lemme fondamental suivant :

Lemme 26. Supposons l'existence d'un $\alpha > 0$ minorant V . Alors pour $\varphi \in C_c^\infty(U)$, on a

$$\varphi(x) = - \int_U L\varphi(y)g(x, y)dy \quad (22)$$

Démonstration.

$$\begin{aligned} \int_U L\varphi(y)g(x, y)dy &= \int_U \frac{d}{du} \Big|_{u=0} \mathbb{E}_y[\varphi(X_u)]g(x, y)dy \\ &= \int_0^\infty \int_U \frac{d}{du} \Big|_{u=0} \mathbb{E}_y[\varphi(X_u)]r(t, x, y)dydt \\ &= \int_0^\infty \int_U \frac{d}{du} \Big|_{u=0} \mathbb{E}_y[\varphi(X_u)]d\mathbb{P}_x(X_t = y)dt \\ &\stackrel{\text{(Markov)}}{=} \int_0^\infty \int_U \frac{d}{du} \Big|_{u=t} \mathbb{E}_x[\varphi(X_u)|X_t = y]d\mathbb{P}_x(X_t = y)dt \\ &= \int_0^\infty \frac{d}{du} \Big|_{u=t} \mathbb{E}_x[\varphi(X_u)] \\ &= \lim_{u \rightarrow \infty} \mathbb{E}_x[\varphi(X_u)] - \mathbb{E}_x[\varphi(x_0)] = 0 - \varphi(x) = -\varphi(x) \end{aligned}$$

On utilise ici de manière essentielle les hypothèses $V \geq \alpha$ et $\varphi \in C_c^\infty$:

- pour assurer la convergence de l'intégrale : on utilise que $L\varphi$ est borné par une constante et que $g(x, \cdot)$ est L^1 — car $\int_U g(x, y) 1 dy = \int_0^\infty R1(x) dt \leq \int_0^\infty e^{-\alpha t} dt < \infty$.
- pour intervertir dérivation et intégrale lors du déconditionnement : on utilise que $L\varphi$ est borné par une constante.
- pour établir que $\mathbb{E}_x[\varphi(X_u)] \xrightarrow{u \rightarrow \infty} 0$: on utilise que $\mathbb{E}_x[\varphi(X_u)] \leq e^{-\alpha u} \sup \varphi$.

□

Montrons maintenant (21). Comme on sait que v est continue sur $U \setminus K$, il n'y a qu'à montrer (21) au sens des distributions, i.e. prouver que pour toute fonction $\varphi \in C_c^\infty(U \setminus K)$:

$$\int_{\mathbb{R}^d} \left(-\frac{1}{2}\Delta + V \right) \varphi(x)v(x) dx = 0$$

Soit donc $\varphi \in C_c^\infty(U \setminus K)$, et supposons dans un premier temps V minoré par un $\alpha > 0$. Utilisant la symétrie de $g(\cdot, \cdot)$, (22) donne :

$$\int \left(-\frac{1}{2}\Delta + V \right) \varphi(x)g(x, y) dx = \varphi(y)$$

d'où en intégrant par rapport à $e(dy)$:

$$\int \left(-\frac{1}{2}\Delta + V \right) \varphi(x) \left(\int_K g(x, y)e(dy) \right) dx = \int \varphi(y)e(dy)$$

soit

$$\int_{\mathbb{R}^d} \left(-\frac{1}{2}\Delta + V \right) \varphi(x)v(x) dx = 0$$

puisque φ est nulle sur K .

Maintenant, si V est juste positive ou nulle, pour $\alpha > 0$ on considère la fonction v_α définie par :

$$v_\alpha(x) = \int_K g_{U, V+\alpha}(x, y)e_K(dy)$$

Il est clair, par définition de g et par convergence monotone, que

$$g_{U, V+\alpha}(x, y) \underset{\alpha \searrow 0}{\nearrow} g_{U, V}(x, y)$$

d'où, toujours par convergence monotone :

$$v_\alpha(x) \underset{\alpha \searrow 0}{\nearrow} v(x)$$

Remarque 27. Attention ! v_α n'est pas le potentiel d'équilibre associé à U et $V+\alpha$, puisque on a intégré par rapport à la mesure $e_K(dy)$, laquelle se rapporte au potentiel V .

Le même calcul que ci-dessus donne alors :

$$\int \left(-\frac{1}{2}\Delta + V \right) \varphi(x)v_\alpha(x)dx = -\alpha \int \varphi(x)v_\alpha(x)$$

Comme $v_\alpha(x)$ est dominé par $v(x)$, lui-même dominé par 1, le théorème de convergence dominée donne quand $\alpha \searrow 0$:

$$\int \left(-\frac{1}{2}\Delta + V \right) \varphi(x)v(x)dx = 0$$

ce qu'on voulait.

5.3 Les conditions aux bords

Montrons maintenant les conditions aux bords. Rappelons pour X un fermé, x un point régulier de X , on a pour tout $\varepsilon > 0$: $W_y(H_X \leq \varepsilon) \xrightarrow{y \rightarrow x} 1$.

Soit x un point régulier de cU ; soit $\eta > 0$ aussi petit qu'on veut. On a pour ε suffisamment petit, pour y voisin de x :

$$W_y(H_K > \varepsilon) \leq \eta$$

Choisissant un tel ε , attendu qu'on a au voisinage de x

$$W_y(T_U \leq \varepsilon) \geq 1 - \eta$$

on a finalement

$$W_y(T_U \leq H_K) \geq 1 - 2\eta$$

d'où *a fortiori*

$$\mathbb{P}_y(H_K < \infty) \leq 2\eta$$

La condition au bord de U s'ensuit.

Soit maintenant x' un point régulier de K . Comme l'heure de la mort du mouvement brownien est presque sûrement > 0 (cf. remarque 8), on a $v(x') = 1$. Observant que $v(z) = \int_K g(z, y)e_K(dy)$, avec $z \mapsto g(z, y)$ continue pour tout y , on a par le lemme de Fatou que v est semi-continue inférieurement. Comme par ailleurs la définition de v rend clair que $0 \leq v(\cdot) \leq 1$, cela impose à v d'être continue en x' □

6 Application : la saucisse de Wiener

6.1 Présentation du problème

☞ *On considère ici $d \geq 3$.*

L'idée de cette partie est d'évaluer la « longueur » d'un mouvement brownien. Même si cela ne signifie rien *stricto sensu*, on peut toujours évaluer la taille de la zone proche de la trajectoire du mobile, comme suit :

Définition 28. Soit $(Z_t)_{t \geq 0}$ un mouvement brownien standard de \mathbb{R}^d . Pour $a > 0$, la saucisse de Wiener de rayon a pour l'intervalle de temps $[t_1, t_2]$ est définie comme l'ensemble des points de \mathbb{R}^d situés à distance $\leq a$ de la trajectoire entre les temps t_1 et t_2 :

$$W_{t_1, t_2}^a = \{x \in \mathbb{R}^d; \exists s \in [t_1, t_2] \quad |x - Z_s| \leq a\}$$

$W_{0, t}^a$ sera noté W_t^a pour simplifier.

☞ **Dans cette partie, \mathbb{P} désigne la loi d'un mouvement brownien standard de \mathbb{R}^d .**

On note $|W_t^a|$ l'aire de W_t^a . Notre objectif est donc d'étudier le comportement de $|W_t^a|$ lorsque a tend vers 0.

6.2 Volume de la saucisse

On observe que $W_t^a = \{x \in \mathbb{R}^d; H_{\bar{B}(x, a)} \leq t\}$. D'où

$$\mathbb{E}[|W_t^a|] \stackrel{\text{(Fubini)}}{=} \int_{\mathbb{R}^d} \mathbb{P}(H_{\bar{B}(x, a)} \leq t) dx$$

Cela suggère d'utiliser les temps d'arrêt introduits dans les parties précédentes.

Commençons par

Lemme 29.

$$\forall x \in \mathbb{R}^d \quad \mathbb{P}\text{-p.s.} \quad L_{\bar{B}(x, a)} \neq 0$$

Démonstration. Soit $0 \notin \bar{B}(x, a)$ et alors c'est trivial par continuité du mouvement brownien. Si $0 \in \bar{B}(x, a)$, cela découle de ce que, pour un mouvement brownien standard, tous les points d'une boule fermée sont réguliers (p. ex. ils vérifient la condition de cône de Zarembka). \square

Intuitivement, lorsque a est petit, la transience du mouvement brownien assure que le mobile, lorsqu'il touchera $\bar{B}(x, a)$, la quittera presque aussitôt. Cela suggère d'écrire :

$$\mathbb{P}(H_{\bar{B}(x, a)} \leq t) = \mathbb{P}(0 < L_{\bar{B}(x, a)} \leq t) + \mathbb{P}(H_{\bar{B}(x, a)} \leq t \text{ et } L_{\bar{B}(x, a)} > t)$$

Calculons d'abord le premier terme. On montre que la mesure d'équilibre de la boule $\bar{B}(x, a)$ est

$$e(dy) = \frac{d-2}{2a} dS_{x, a}(y) \tag{23}$$

où $dS_{x, a}(\cdot)$ est la mesure de Lebesgue sur la sphère $\partial\bar{B}(x, a)$.

En utilisant le théorème 22, on obtient alors :

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathbb{P}(0 < L_{\bar{B}(x,a)} \leq t) dx &= \\
&= \int_{\mathbb{R}^d} \int_0^t \int_{\partial \bar{B}(x,a)} \frac{1}{(2\pi u)^{d/2}} e^{-|y|^2/2} \frac{d-2}{2a} dS_{x,a}(y) du dx \\
&= \int_{\mathbb{R}^d} \int_0^t \int_{\partial \bar{B}_{0,a}} \frac{1}{(2\pi u)^{d/2}} e^{-|x+z|^2/2} \frac{d-2}{2a} dS_{0,a}(z) du dx \\
&\stackrel{\text{(Fubini)}}{=} \int_0^t \int_{\partial \bar{B}_{0,a}} \frac{d-2}{2a} dS_{0,a}(z) du = \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} t
\end{aligned}$$

(on utilise ici la notation $x! = \Gamma(x+1)$ pour $x \in \mathbb{R}$).

On va maintenant majorer le second terme. Pour ce faire, on coupe ce dernier en deux : pour tout $\eta > 0$, on écrit

$$\begin{aligned}
\mathbb{P}(0 < H_{\bar{B}(x,a)} \leq t \text{ et } L_{\bar{B}(x,a)} > t) &\leq \\
&\mathbb{P}(t < L_{\bar{B}(x,a)} \leq t + \eta) + \mathbb{P}(H_{\bar{B}(x,a)} \leq t \text{ et } L_{\bar{B}(x,a)} > t + \eta)
\end{aligned}$$

Un calcul totalement indentique au précédent montre que le premier terme du membre de droite s'intègre en :

$$\int_{\mathbb{R}^d} \mathbb{P}(t < L_{\bar{B}(x,a)} \leq t + \eta) dx = \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} \eta$$

Pour le second terme, on a :

$$\begin{aligned}
\mathbb{P}(H_{\bar{B}(x,a)} \leq t \text{ et } L_{\bar{B}(x,a)} > t + \eta | \mathcal{G}_{H_{\bar{B}(x,a)}}) &\stackrel{\text{(Markov)}}{\leq} \\
&\mathbb{1}_{H_{\bar{B}(x,a)} \leq t} \mathbb{P}^{Z_{H_{\bar{B}(x,a)}}}(L_{\bar{B}(x,a)} > \eta) \\
&\leq \mathbb{1}_{H_{\bar{B}(x,a)} \leq t} \sup_{h \in \bar{B}(x,a)} \mathbb{P}_h(L_{\bar{B}(x,a)} > \eta) \\
&\stackrel{\text{(thm 22)}}{\leq} \mathbb{1}_{H_{\bar{B}(x,a)} \leq t} \int_{\eta}^{\infty} ds \int_{\bar{B}(x,a)} p(s, h, y) e_{\bar{B}(x,a)}(dy) \\
&\stackrel{\text{(23)}}{\leq} \mathbb{1}_{H_{\bar{B}(x,a)} \leq t} \int_{\eta}^{\infty} ds \int_{\partial \bar{B}(x,a)} p(s, h, y) \frac{d-2}{2a} dS_{x,a}(y)
\end{aligned}$$

D'où, en utilisant $p(s, \cdot, \cdot) \leq (2\pi s)^{-d/2}$, $S_{x,a}(\partial \bar{B}(x, a)) = \frac{2\pi^{d/2}}{(d/2-1)!} a^{d-1}$ et en déconditionnant :

$$\mathbb{P}(H_{\bar{B}(x,a)} \leq t \text{ et } L_{\bar{B}(x,a)} > t + \eta) \leq \frac{1}{(d/2-1)!(2\eta)^{d/2-1}} a^{d-2} \mathbb{P}(H_{\bar{B}(x,a)} \leq t) \quad (24)$$

On conviendra de noter $K_{d,\eta}$ la constante $\frac{1}{(d/2-1)!(2\eta)^{d/2-1}}$. Il nous suffira par la suite de nous rappeler que cette constante est finie.

Intégrant (24) par rapport à x , il vient enfin :

$$\int_{\mathbb{R}^d} \mathbb{P}(H_{\bar{B}(x,a)} \leq t \text{ et } L_{\bar{B}(x,a)} > t + \eta) dx \leq K_{d,\eta} \mathbb{E}[|W_t^a|] a^{d-2}$$

Nous sommes maintenant en état de conclure. Nous aurons besoin du lemme suivant :

Lemme 30.

$$\mathbb{E}[|W_t^a|] < \infty$$

Démonstration. On a en effet $W_t^a \subset \bar{B}(0, a + \sup_{0 \leq s \leq t} |Z_t|)$, d'où $|W_t^a| \leq k_d(a + \sup_{0 \leq s \leq t} |Z_t|)^d$, avec $k_d = \pi^{d/2}/(d/2)!$. Maintenant, on observe que $|Z_s| \leq \sum_{i=1}^d |Z_s^{(i)}|$, donc $\sup_{0 \leq s \leq t} |Z_t| \leq \sum_{i=1}^d \sup_{0 \leq s \leq t} |Z_s^{(i)}|$ et

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq s \leq t} |Z_t| \geq X\right) &\leq d \cdot \mathbb{P}\left(\sup_{0 \leq s \leq t} |Z_s^{(1)}| \geq \frac{X}{d}\right) \\ &\leq d \cdot \mathbb{P}\left(|Z_t^{(1)}| \geq \frac{X}{d}\right) \leq 2d \cdot e^{-X^2/2d^2t} \end{aligned}$$

où, dans la dernière inégalité, on a utilisé la relation bien connue :

Lemme 31. Pour ξ une variable aléatoire gaussienne centrée, $\mathbb{P}(\xi \geq X) \leq e^{-X^2/2}$.

On obtient alors

$$\mathbb{P}(|W_t^a| \geq X) \leq \mathbb{P}\left(\sup_{0 \leq s \leq t} |Z_t| \geq \left(\frac{X}{k_d}\right)^{1/d} - a\right) \leq 2d \cdot \exp\left\{-\frac{\left(\left(\frac{X}{k_d}\right)^{1/d} - a\right)^2}{2d^2t}\right\}$$

Cela assure à la loi de $|W_t^a|$ une décroissance suffisamment rapide pour que son espérance soit finie. \square

Remarque 32. Cette démonstration donne même que $|W_t^a|$ a tous ses moments finis, i.e. $\forall p \in [1, \infty[\quad \mathbb{E}[|W_t^a|]^p < \infty$.

Récapitulant les calculs des paragraphes précédents, il vient alors :

$$\frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} t \leq \mathbb{E}[|W_t^a|] \leq \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} (t + \eta) + K_{d,\eta} a^{d-2} \mathbb{E}[|W_t^a|]$$

La majoration de $\mathbb{E}[|W_t^a|]$ donne, en rassemblant $\mathbb{E}[|W_t^a|]$ dans un seul membre⁴, pour a assez petit :

$$\mathbb{E}[|W_t^a|] \leq \frac{2\pi^{d/2}}{(d/2-2)!} \cdot \frac{1}{1 - K_{\eta,d} a^{d-2}} (t + \eta) a^{d-2} \underset{a \rightarrow 0}{\sim} \frac{2\pi^{d/2}}{(d/2-2)!} (t + \eta) a^{d-2}$$

Comme cela est vrai pour η aussi petit qu'on veut, on a finalement :

⁴C'est là qu'on utilise le lemme 30.

Théorème 33.

$$\mathbb{E}[|W_t^a|] \underset{a \rightarrow 0}{\sim} \frac{2\pi^{d/2}}{(d/2 - 2)!} t a^{d-2}$$

6.3 Un résultat plus fort

Corollaire 34. Pour tout $p \in [1, \infty[$, on a :

$$\frac{|W_t^a|}{a^{d-2}} \underset{a \rightarrow 0}{\xrightarrow{L^p}} \frac{2\pi^{d/2}}{(d/2 - 2)!} t$$

Démonstration. L'astuce consiste à utiliser l'invariance par changement d'échelle du mouvement brownien pour appliquer un argument de sous-additivité. Soit $a > 0$ fixé. Par changement d'échelle, le théorème précédent donne que

$$\frac{\mathbb{E}[|W_t^a|]}{t} \underset{t \rightarrow \infty}{\rightarrow} \frac{2\pi^{d/2}}{(d/2 - 2)!} a^{d-2}$$

Maintenant, on observe qu'on a clairement :

- Pour $s \leq t$ $|W_{s,t}^a| \stackrel{\text{loi}}{=} |W_{t-s}^a|$;
- Pour $t_1 \leq t_2 \leq \dots \leq t_n$, $|W_{t_1,t_2}^a|, \dots, |W_{t_n,t_{n-1}}^a|$ sont indépendants ;
- $|W_{t_1,t_n}^a| \leq |W_{t_1,t_2}^a| + \dots + |W_{t_{n-1},t_n}^a|$.

Soit $u > 0$ une durée de référence ; si on note $X_u^{(i)}$ pour $|W_{(i-1)u,iu}^a|$, alors $X_u^{(1)}, X_u^{(2)}, \dots$ sont des variables aléatoires indépendantes et de même loi que $|W_u^a|$. On a alors :

$$\mathbb{P}\text{-p.s.} \quad \frac{|W_t^a|}{t} \leq \frac{1}{t} \sum_{i=1}^{\lceil \frac{t}{u} \rceil} X_u^{(i)}$$

avec

$$\frac{1}{t} \sum_{i=1}^{\lceil \frac{t}{u} \rceil} X_u^{(i)} \underset{t \rightarrow \infty}{\xrightarrow{\text{LGN}}} \frac{1}{u} \mathbb{E}[|W_u^a|]$$

la convergence ayant lieu dans L^1 et même dans L^p (cf. remarque 32)⁵. Par ailleurs, on vient de voir que

$$\frac{\mathbb{E}[|W_t^a|]}{t} \underset{t \rightarrow \infty}{\rightarrow} \frac{2\pi^{d/2}}{(d/2 - 2)!} a^{d-2}$$

⁵On en déduit en particulier que la suite $\left(\frac{|W_t^a|}{t}\right)_{t \geq 0}$ est bornée dans L^p pour tout $p \in [1, \infty[$.

d'où

$$\begin{aligned}
& \left\| \frac{|W_t^a|}{t} - \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} \right\|_{L^1} \leq \\
& \left\| \frac{1}{t} \sum_{i=1}^{\lfloor \frac{t}{u} \rfloor} X_u^{(i)} - \frac{|W_t^a|}{t} \right\|_{L^1} + \left\| \frac{1}{t} \sum_{i=1}^{\lfloor \frac{t}{u} \rfloor} X_u^{(i)} - \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} \right\|_{L^1} \\
& = \mathbb{E} \left[\frac{1}{t} \sum_{i=1}^{\lfloor \frac{t}{u} \rfloor} X_u^{(i)} \right] - \mathbb{E} \left[\frac{|W_t^a|}{t} \right] + \left\| \frac{1}{t} \sum_{i=1}^{\lfloor \frac{t}{u} \rfloor} X_u^{(i)} - \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} \right\|_{L^1} \\
& \xrightarrow{t \rightarrow \infty} \frac{\mathbb{E}[|W_u^a|]}{u} - \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} + \left| \frac{\mathbb{E}[|W_u^a|]}{u} - \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2} \right| \xrightarrow{u \rightarrow \infty} 0
\end{aligned}$$

Cela démontre que $\frac{|W_t^a|}{t} \xrightarrow[t \rightarrow \infty]{L^1} \frac{2\pi^{d/2}}{(d/2-2)!} a^{d-2}$. On en conclut la convergence L^p , via le lemme d'uniforme intégrabilité suivant (laissé au lecteur) :

Lemme 35. Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires, bornée dans L^p pour tout $p \in [1, \infty[$. On suppose que $X_n \xrightarrow[n \rightarrow \infty]{\text{proba}} Y$. Alors la convergence a aussi lieu dans les L^p .

□

Remarque 36. En fait, on a également la convergence presque sûre, mais c'est un peu plus délicat à démontrer ; il faut passer par un contrôle de la variance des $|W_t^a|$.

6.4 Dimension de Hausdorff du mouvement brownien

Les résultats précédents montrent donc que la taille de la saucisse de rayon a autour du mouvement brownien est de l'ordre de grandeur de a^{d-2} . Cela confirme l'idée que le mouvement brownien est de longueur infinie, puisque pour un chemin simple, mettons C^1 , de longueur γ , la taille de la saucisse quand $a \rightarrow 0$ est équivalente à $k_{d-1} \gamma a^{d-1}$, où k_{d-1} est la mesure de Lebesgue de la boule unité en dimension $d-1$, soit $\frac{\pi^{(d-1)/2}}{((d-1)/2)!}$. La croissance en a^{d-2} , par contre, est typique d'une surface, puisque, si on a une surface C^1 d'aire σ , le volume de la saucisse est équivalent à $k_{d-2} \sigma a^{d-2}$ quand $a \rightarrow 0$.

Pour $X \subset \mathbb{R}^d$ non vide, on appelle *dimension (fractale) de Hausdorff* de X le réel de $[0, d]$, s'il existe, défini par

$$\lim_{a \rightarrow 0} \log_a(|W_X^a|)$$

où W_X^a note la saucisse de rayon a autour de X . Admettant la remarque 36, on voit donc que la trajectoire du mouvement brownien pour $d \geq 3$ est de dimension 2 presque sûrement. Toujours par la remarque 36, on peut même définir l'aire (!) de la trajectoire entre 0 et t , égale à $(d-2)\pi$ presque sûrement.

Références

- [1] A-S. Znitman, *Brownian motion, obstacles and random media*, Springer (1998), pp. 51–67 & préc.

Document VI

Autour du mouvement brownien

Autour du mouvement brownien

Rémi Peyre

21 mars 2007

Résumé

Le document suivant constitue les notes d'une série d'exposés donnés du 7 au 21 mars 2007 dans le cadre du séminaire étudiant de l'ENS Lyon, destiné à des non-probabilistes. Le thème général est l'étude des processus aléatoires continus et s'organise en deux parties, respectivement appelées « Mouvement brownien dans \mathbb{R}^d » et « Calcul stochastique »

Le plan de la première partie est le suivant. Dans une première section, on introduit le mouvement brownien comme une sorte de marche aléatoire universelle, jusqu'à en arriver à une définition rigoureuse. Les propriétés élémentaires du mouvement brownien sont passées en revue dans une deuxième section. Enfin, la troisième section de cet exposé explique le lien entre mouvement brownien et laplacien, tant du point de vue du générateur que de celui du générateur dual. Pour le plan de cette partie, j'ai essayé de rester le plus proche possible de ma vision personnelle des choses, de sorte que j'hésiterais à désigner un ouvrage de référence. Je me suis en partie appuyé sur mes notes de cours de J-F. Le Gall, dont on trouvera une version informatisée en [1], mais je conseillerais plutôt au néophyte l'approche plus concrète de R. Bass [2].

Dans la deuxième partie, on parle de calcul stochastique : j'ai essayé d'introduire ce domaine de la façon la plus intuitive possible, mais sans trop faire l'impasse sur la rigueur pour autant. Pour les sources, je me suis appuyé sur mes notes manuscrites du cours de M2 d'Olivier Raimond, ainsi que sur [2] dont j'apprécie la légèreté du style, mais la référence incontestée du sujet est [4]. L'essentiel de cette partie est assez didactique : notion de martingale continue, variation quadratique, intégrales d'Itô et de Stratonovitch. Puis en guise d'application, nous montrons comment la formule d'Itô permet de montrer l'invariance conforme du mouvement brownien. Ainsi, on peut donner une construction explicite de la représentation conforme d'un ouvert non plein de \mathbb{R}^2 .

Table des matières

1	Mouvement brownien dans \mathbb{R}^d
---	--

1.1	Construction du mouvement brownien	122
1.1.1	Rappels : la loi des grands nombres, le théorème-limite central	122
1.1.2	Motivation	123
1.1.3	Une construction du mouvement brownien	124
1.2	propriétés du mouvement brownien	126
1.2.1	Visualisation graphique	126
1.2.2	Premières propriétés	129
1.2.3	Propriétés non triviales	130
1.3	Mouvement brownien et laplacien	133
1.3.1	Générateur du mouvement brownien	133
1.3.2	Résolution du problème de Dirichlet	134
1.3.3	Une application : transience du mouvement brownien en dimension 3.	135
1.3.4	Générateur dual	137
1.3.5	Application : conditions au bord de Von Neuman en dimension 1	138
2	Calcul stochastique	140
2.1	Concept de martingale	140
2.2	Intégration contre le mouvement brownien	141
2.2.1	Motivations	141
2.2.2	Intégrale d'Itô	143
2.2.3	Intégrale de Stratonovitch	144
2.3	Calcul stochastique général	146
2.3.1	Variation quadratique d'une martingale continue	146
2.3.2	Intégration contre une martingale continue	147
2.3.3	Le théorème de Lévy	147
2.3.4	Semimartingales	149
2.3.5	Formalisme différentiel stochastique	150
2.4	Invariance conforme du mouvement brownien	150

2.4.1	Le mouvement brownien est invariant conforme en dimension 2	151
2.4.2	Application : représentation conforme de Riemann	152

1 Mouvement brownien dans \mathbb{R}^d

☞ Dans toute cette partie, on travaille dans l'espace euclidien \mathbb{R}^d , où $d \geq 1$ est un entier fixé. Les vecteurs de \mathbb{R}^d sont notés avec une flèche, et les matrices de $\mathbb{R}^{d \times d}$ en gras.

1.1 Construction du mouvement brownien

1.1.1 Rappels : la loi des grands nombres, le théorème-limite central

Théorème 1.1 (loi des grands nombres). Soit \vec{X} une variable aléatoire à valeurs dans \mathbb{R}^d , de classe L^1 (c'est-à-dire que chacune des composantes de ce vecteur est L^1). On se donne $\vec{X}_1, \vec{X}_2, \dots$ des variables aléatoires indépendantes et de même loi que \vec{X} . Alors,

$$p.s. \quad \frac{1}{N} \sum_{i=1}^N \vec{X}_i \xrightarrow{N \rightarrow \infty} \mathbb{E}[\vec{X}], \quad (1)$$

où l'espérance du vecteur \vec{X} n'est rien d'autre que le vecteur dont chaque composante est l'espérance de la composante correspondante de \vec{X} .

Théorème 1.2 (théorème-limite central). Soit \vec{X} une variable aléatoire à valeurs dans \mathbb{R}^d , de classe L^2 , centrée, i.e. $\mathbb{E}[\vec{X}] = \vec{0}$. Soit \mathbf{C} la matrice de covariance de \vec{X} : notant $\vec{X} = (X^{(1)}, \dots, X^{(d)})$, \mathbf{C} est une matrice symétrique dont les coefficients sont définis par :

$$C_{ij} = \mathbb{E}[X^{(i)} X^{(j)}]. \quad (2)$$

On se donne $\vec{X}_1, \vec{X}_2, \dots$ des variables aléatoires indépendantes et de même loi que \vec{X} . Alors,

$$\frac{1}{N^{1/2}} \sum_{i=1}^N \vec{X}_i \xrightarrow[N \rightarrow \infty]{\text{loi}} \mathcal{N}(\mathbf{C}), \quad (3)$$

où $\mathcal{N}(\mathbf{C})$ est le vecteur gaussien de covariance \mathbf{C} , dont la définition est rappelée ci-dessous. (Rappelons également que la convergence en loi d'une suite de variables $(\vec{Y}_n)_{n \in \mathbb{N}^*}$ vers une variable \vec{Z} signifie que, pour $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction continue bornée quelconque, $\mathbb{E}[f(\vec{Y}_n)] \xrightarrow[N \rightarrow \infty]{} \mathbb{E}[f(\vec{Z})]$).

Définition 1.3. On appelle vecteur gaussien réduit tout vecteur aléatoire \vec{X} de \mathbb{R}^d dont la loi possède une densité par rapport à la mesure de Lebesgue donnée par :

$$\frac{d\mathbb{P}(\vec{X} = \vec{v})}{dv} = (2\pi)^{-d/2} e^{-|\vec{v}|^2/2}. \quad (4)$$

Pour $M \in \mathbb{R}^{d \times d}$ une matrice, le vecteur aléatoire $M\vec{v}$ a une loi qui ne dépend que de $M^t M$, laquelle matrice est d'ailleurs sa covariance ; $M\vec{v}$ (ou, plus précisément, sa loi) est appelé le vecteur gaussien de covariance $M^t M$.

1.1.2 Motivation

Comme il est bien connu, l'expression « mouvement brownien » vient du nom du biologiste Robert Brown qui, en 1827, observa au sein de grains de pollens les mouvements très irréguliers de poussières microscopiques [3]. Nous allons plutôt nous poser une question abstraite : « À quoi ressemble une marche au hasard dans \mathbb{R}^d , vue de loin ? ».

Soit donc \vec{X} un vecteur aléatoire de \mathbb{R}^d de classe L^2 , centré et de covariance C . Soient $\vec{X}_1, \vec{X}_2, \dots$ des copies indépendantes de \vec{X} . Alors on peut définir un processus à temps discret $(M_t)_{t \in \mathbb{N}}$ à valeurs dans \mathbb{R}^d (dont on choisit un point O comme origine) par :

$$\overrightarrow{OM}_t = \sum_{i=1}^t \vec{X}_i. \quad (5)$$

Maintenant, nous voulons étudier le comportement de ce processus « vu de loin ». Comme le théorème-limite central nous dit que M_t va s'éloigner de O d'une distance de l'ordre de $t^{1/2}$, il y a une renormalisation toute trouvée par laquelle nous pouvons définir un processus $(M_t^\lambda)_{t \in (1/\lambda)\mathbb{N}}$, pour $\lambda > 0$:

$$\overrightarrow{OM}_t^\lambda = \lambda^{-1/2} \overrightarrow{OM}_{\lambda t}. \quad (6)$$

Nous avons envie d'étudier le processus M^λ pour λ très grand, dans l'idée de voir apparaître à la limite la loi d'un processus à temps continu $(M_t^\infty)_{t \geq 0}$.

À quoi doit ressembler un tel processus ? D'une part, le théorème-limite central nous dit que la loi du vecteur $\overrightarrow{M_t M_{t'}}$, pour $t \leq t'$, doit converger vers $\mathcal{N}((t' - t)C)$. D'autre part, par passage à la limite on voit que la trajectoire du processus après l'instant t (en prenant M_t^∞ pour point de départ) sera indépendante de la trajectoire du processus avant cet instant.

Donc, un mouvement brownien $(B_t)_{t \geq 0}$, s'il est possible de le définir, devra vérifier le

Théorème 1.4. Pour $t_0 \leq t_1 \leq \dots \leq t_n$, la loi de $(\overrightarrow{B_{t_0} B_{t_1}}, \overrightarrow{B_{t_1} B_{t_2}}, \dots, \overrightarrow{B_{t_{n-1}} B_{t_n}})$ est le produit de vecteur gaussiens indépendants de lois respectives $\mathcal{N}((t_1 - t_0)C), \dots, \mathcal{N}((t_n - t_{n-1})C)$.

On appellera mouvement brownien *standard* le mouvement brownien dont la covariance est l'identité. Il nous suffit alors de donner une construction du mouvement brownien standard puisque, si $(B_t)_{t \geq 0}$ est un mouvement brownien standard et M une matrice de $\mathbb{R}^{d \times d}$, le processus $(B'_t)_{t \geq 0}$ défini par $\overrightarrow{OB'_t} = M \overrightarrow{OB_t}$ est un mouvement brownien de covariance $M^t M$ ⁽¹⁾.

1.1.3 Une construction du mouvement brownien

Soit $\mathbb{D} = \{x \in \mathbb{R}_+ ; \exists p \in \mathbb{N}, q \in \mathbb{N} \ x = \frac{p}{2^q}\}$ l'ensemble des nombres dyadiques positifs ; on note $\mathbb{D} = \bigsqcup_{i \geq 0} \mathbb{D}^{(i)}$, avec $\mathbb{D}^{(0)} = \mathbb{N}$, $\mathbb{D}^{(1)} = \frac{1}{2}\mathbb{N} \setminus \mathbb{N}$, $\mathbb{D}^{(2)} = \frac{1}{4}\mathbb{N} \setminus \frac{1}{2}\mathbb{N}$, etc. Il n'existe qu'une loi possible du mouvement brownien sur l'ensemble des points dyadiques, puisque nous connaissons la loi du mouvement brownien sur tous les ensembles finis⁽²⁾. Voici comment on peut construire effectivement une telle loi⁽³⁾ :

1^{re} étape : On prend des vecteurs aléatoires gaussiens réduits⁽⁴⁾ indépendants $\vec{v}_1, \vec{v}_2, \dots$, et pour $t \in \mathbb{N}$, on définit

$$\overrightarrow{OB_t} = \sum_{i=1}^t \vec{v}_i. \quad (7)$$

Il est évident que $(B_t)_{t \in \mathbb{N}}$ vérifie la loi du mouvement brownien.

2^e étape : Pour $t \in \mathbb{D}^{(1)}$, on donne à B_t la loi

$$B_t = \frac{B_{t-1/2} + B_{t+1/2}}{2} + \mathcal{N}\left(\frac{1}{4}\mathbf{I}_d\right), \quad (8)$$

et on fixe ainsi la loi de chaque $x \in \mathbb{D}^{(1)}$ de manière indépendante. On vérifie par le calcul que c'est bien ainsi qu'on récupère sur $\frac{1}{2}\mathbb{N}$ la loi demandée.

3^e étape : On recommence la seconde étape pour fixer la valeur de B_t pour $t \in \mathbb{D}^{(2)}$, puis $t \in \mathbb{D}^{(3)}$, et ainsi de suite. À l'étape i , la loi d'un point d'abscisse $t \in \mathbb{D}^{(i)}$ est fixée comme :

$$B_t = \frac{B_{t-1/2^i} + B_{t+1/2^i}}{2} + \mathcal{N}\left(\frac{1}{2^{i+1}}\mathbf{I}_d\right). \quad (9)$$

¹Cela permet bien de construire des mouvements browniens de n'importe quelle covariance, puisque toute matrice réelle symétrique C peut se mettre sous la forme $M^t M$.

²Si nous avons pris les dyadiques dans un premier temps, plutôt que \mathbb{R}_+ directement, c'est parce que la tribu sur \mathbb{R}_+ engendrée par les cylindres ne permet pas de mesurer certaines propriétés des trajectoires à temps continu, par exemple la continuité.

³La méthode que nous décrivons ici est particulièrement bien adaptée aux simulations car on peut augmenter progressivement sa résolution.

⁴i.e., de covariance \mathbf{I}_d .

Et maintenant que nous avons défini une fonction $B : \mathbb{D} \longrightarrow \mathbb{R}^d$, y a-t-il un moyen naturel d'étendre cette fonction à tout \mathbb{R}_+ ? Oui ! C'est le

Théorème 1.5. *Presque-sûrement, la fonction sur \mathbb{D} que nous venons de définir est continue. On peut donc la prolonger par continuité⁽⁵⁾ en une fonction $B : \mathbb{R}_+ \longrightarrow \mathbb{R}^d$, laquelle vérifie le théorème 1.4, à cause de la satisfaction de ce théorème par la fonction restreinte à \mathbb{D} et de la continuité de B .*

Démonstration. Nous allons démontrer que B est absolument continu sur $\mathbb{D} \cap [0, 1]$; par réunion dénombrable il s'ensuivra que B est continu sur \mathbb{D}_+ presque-sûrement. On pose :

$$\delta_n = \max_{x \in [0, 1] \cap \frac{1}{2^n} \mathbb{Z}} |\overrightarrow{B_x B_{x+1/2^n}}|. \quad (10)$$

Alors, pour démontrer l'absolue continuité de B sur $\mathbb{D} \cap [0, 1]$, il nous suffit de démontrer que $\sum_n \delta_n$ converge p.s.. Allons-y donc !

δ_n s'écrit comme le maximum de 2^n variables aléatoires indépendantes de même loi $\mathcal{N}(\frac{1}{2^n} \mathbf{I}_d)$. Or, on démontre aisément l'existence de constantes a et K telles que, si \vec{X} est un vecteur aléatoire gaussien réduit, alors, pour tout $x \geq 0$, $\mathbb{P}(|\vec{X}| \geq x) \leq K e^{-x^2/a}$. De sorte que

$$\forall x \geq 0 \quad \mathbb{P}(\delta_n \geq x) \leq 2^n K \exp\left(-\frac{2^n x^2}{a}\right). \quad (11)$$

En particulier, posons $x_n = 2^{-n/4}$, alors :

$$\mathbb{P}(\delta_n \geq x_n) \leq 2^n K \exp\left(-\frac{2^{n/2}}{a}\right), \quad (12)$$

qui est le terme général d'une série convergente. Par le lemme de Borel-Cantelli, il s'ensuit que p.s. $\delta_n \leq x_n$ à partir d'un certain rang ; comme $\sum_n x_n$ converge, $\sum_n \delta_n$ converge p.s.. \square

Remarque 1.6. La continuité du B n'avait rien d'évident *a priori* ! Imaginons par exemple que nous ayons considéré, au départ, une marche aléatoire dont les pas n'étaient pas L^2 . Alors, sous certaines hypothèses convenables, la marche s'éloigne de l'origine en t^α pour un $\alpha \in]\frac{1}{2}, \infty[$. On peut alors faire comme précédemment un travail de renormalisation pour arriver à un processus, qualifié de processus $\frac{1}{\alpha}$ -stable. Ce processus n'est *pas* continu et présente des sauts, comme on peut le constater sur la figure 1.1.3 : la meilleure régularité qu'on puisse obtenir est la régularité de classe càd-làg (continu à droite, limite à gauche).

⁵Dans la pratique, on ajoute comme hypothèse que le mouvement brownien est *toujours* continu, et pas seulement presque-sûrement, et ce afin de pouvoir dire qu'un mouvement brownien est une variable aléatoire à valeurs dans l'espace des fonctions continues de \mathbb{R}_+ dans \mathbb{R}^d .

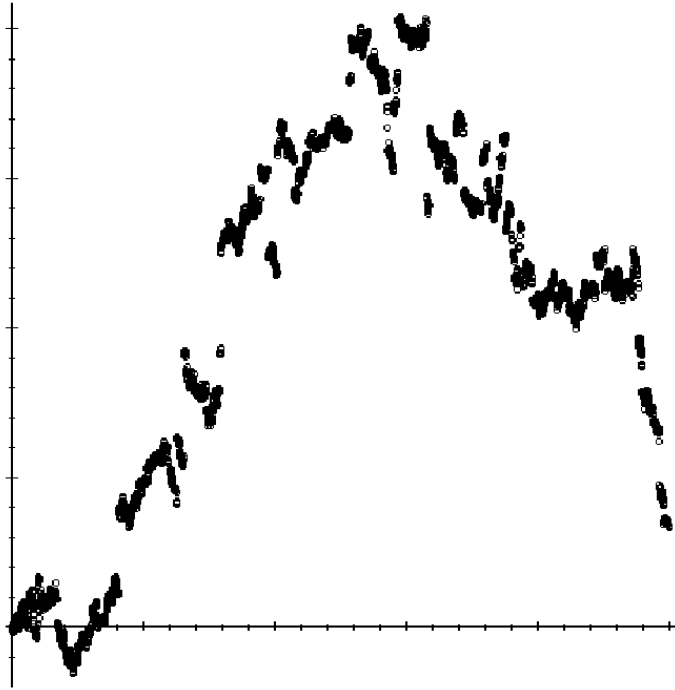


FIG. 1 – Une réalisation d'un processus $\frac{3}{2}$ -stable : ce n'est clairement pas un processus continu

1.2 propriétés du mouvement brownien

☞ *Dorénavant tous les mouvements browniens que nous considérerons seront standard ; les propriétés des mouvements browniens de covariance quelconque se déduisent facilement des propriétés du mouvement brownien standard énoncées ci-dessous.*

☞ *Dans la partie précédente, nous avons formalisé le mouvement brownien comme un processus à valeurs dans l'espace physique (i.e., B_t est un point), issu d'un point-origine O ; dorénavant pour alléger les notations nous noterons \vec{B}_t pour $\overrightarrow{OB_t}$, de sorte que $(\vec{B}_t)_{t \geq 0}$ sera un processus à valeurs vectorielles.*

1.2.1 Visualisation graphique

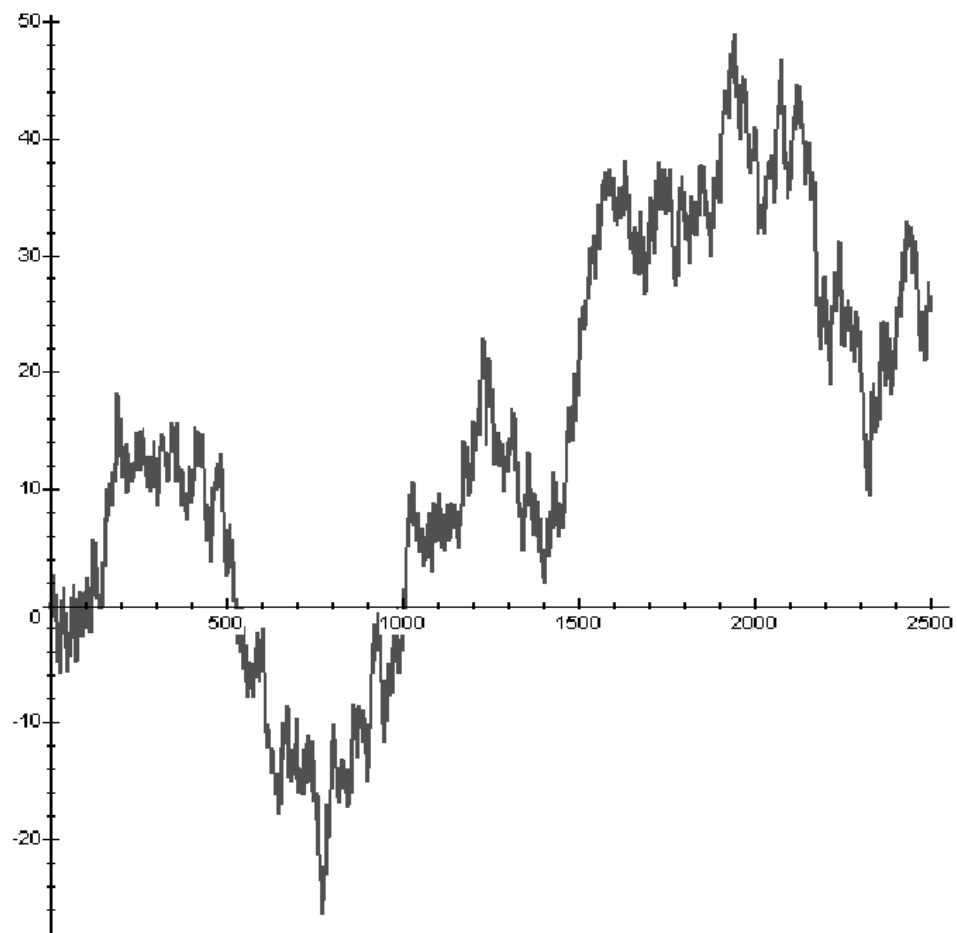


FIG. 2 – Graphe d'une réalisation du mouvement brownien en dimension 1

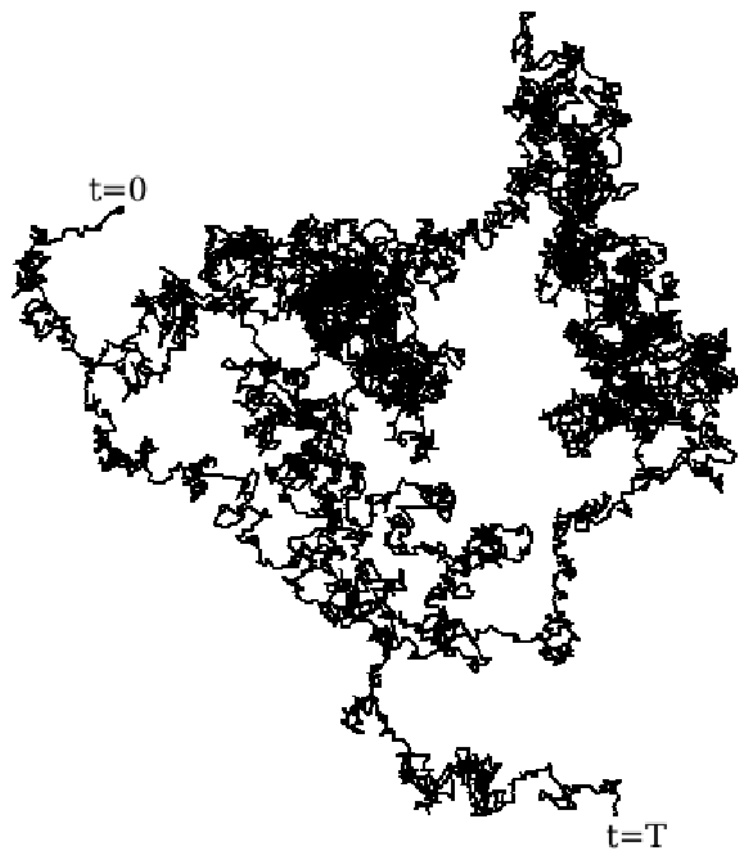


FIG. 3 – Trajectoire sur $[0, T]$ d'une réalisation du mouvement brownien en dimension 2

1.2.2 Premières propriétés

Proposition 1.7 (propriété de Markov faible). *Pour tout $t \geq 0$, on note $\mathcal{F}_t = \sigma(\vec{B}_u; u \leq t)$ la tribu engendrée par la trajectoire du mouvement brownien jusqu'au temps t . Alors le processus $(\vec{B}_{u+t} - \vec{B}_t)_{u \geq 0}$ est indépendant de \mathcal{F}_t , et sa loi est celle d'un mouvement brownien.*

Cette propriété est évidente à partir des définitions. Elle a pour corollaire essentiel la propriété de Markov forte, que nous énonçons ci-dessous sans démonstration. En fait, la démonstration de Markov fort à partir de Markov faible repose simplement sur un jeu d'écriture formel assez pénible, mais l'intuition permet de l'admettre sans difficulté.

Nous avons besoin d'introduire la

Définition 1.8. *Une variable aléatoire τ à valeurs dans $[0, \infty]$ est un temps d'arrêt si, pour tout $t \geq 0$, l'événement $\{\tau \leq t\}$ est \mathcal{F}_t -mesurable. Moralement, un temps d'arrêt correspond à une règle permettant de crier « stop ! » au cours de l'évolution du mouvement brownien pour un observateur qui découvre la trajectoire au fur et à mesure.*

Pour τ un temps d'arrêt, on définit la tribu du passé avant τ , \mathcal{F}_τ , qui est moralement l'ensemble des informations dont notre observateur dispose sur le mouvement brownien au moment où il crie stop. Techniquement, un événement A est dans \mathcal{F}_τ si et seulement si :

$$\forall t \geq 0 \quad (A \text{ et } \tau \leq t) \in \mathcal{F}_t. \quad (13)$$

Corollaire 1.9 (propriété de Markov forte). *Pour tout temps d'arrêt τ , le processus $(\vec{B}_{\tau+u} - \vec{B}_\tau)_{u \geq 0}$, s'il existe⁽⁶⁾, est indépendant de \mathcal{F}_τ et est aussi un mouvement brownien.*

Les propriétés de Markov expriment donc que la trajectoire future d'un mouvement brownien (en ayant remis l'origine à la position actuelle du mobile) ne dépend pas de son passé.

Quelques autres propriétés immédiates du mouvement brownien :

Proposition 1.10 (isotropie). *Si $(\vec{B}_t)_{t \geq 0}$ est un mouvement brownien et M une matrice de $O(d)$, alors $(M\vec{B}_t)_{t \geq 0}$ est aussi un mouvement brownien standard.*

Proposition 1.11 (invariance d'échelle). *Si $(\vec{B}_t)_{t \geq 0}$ est un mouvement brownien, alors, pour $\lambda \in]0, \infty[$, le processus $(\vec{B}_t^\lambda)_{t \geq 0}$ défini par*

$$\vec{B}_t^\lambda = \lambda^{-1/2} \vec{B}_{\lambda t} \quad (14)$$

est lui aussi un mouvement brownien standard.

⁶Ce processus n'existe en effet pas si $\tau = \infty$.

Une dernière propriété, évidente pour les lois finidimensionnelles de mouvement brownien, quoiqu'assez technique à démontrer sous sa forme continue :

Théorème 1.12 (théorème de support). *Notons \mathcal{C} l'espace des fonctions continues de \mathbb{R}_+ à valeurs dans \mathbb{R}^d issues de $\vec{0}$, que nous dotons de la topologie définie par la convergence sur tout compact, \mathcal{C} est alors un espace polonais (i.e., séparable et métrisable). Le mouvement brownien est alors une variable aléatoire à valeurs dans \mathcal{C} , et le support de cette variable aléatoire est \mathcal{C} tout entier.*

Remarque 1.13. En d'autres termes, le mouvement brownien a une probabilité non nulle d'être localement C^0 -proche de n'importe quelle fonction continue.

1.2.3 Propriétés non triviales

Lois de 0 – 1.

Les lois de 0 – 1 sont le point de départ de nombreuses démonstrations concernant les propriétés du mouvement brownien.

Définition 1.14. *On appelle tribu instantanée, notée \mathcal{F}_{0+} , la tribu des événements qu'on peut déterminer au bout d'un temps aussi court que l'on veut :*

$$\mathcal{F}_{0+} = \bigcap_{\varepsilon \searrow 0} \mathcal{F}_\varepsilon. \quad (15)$$

On appelle tribu asymptotique, notée \mathcal{F}^∞ ⁽⁷⁾, la tribu des événements qu'on peut déterminer à partir d'un moment aussi tardif qu'on veut :

$$\mathcal{F}^\infty = \bigcap_{t \nearrow \infty} \sigma(\vec{B}_u; u \geq t). \quad (16)$$

Théorème 1.15. *La tribu \mathcal{F}_{0+} est triviale, i.e. elle ne contient que des événements dont la probabilité est 0 ou 1.*

Démonstration. Soit un événement $A \in \mathcal{F}_{0+}$, et soit un événement $B \in \mathcal{F}^\infty$ ⁽⁸⁾. Nous allons montrer que A et B sont indépendants. Prenant $B = A$, il s'en suivra alors que A a pour probabilité 0 ou 1.

On observe d'abord la validité d'un lemme tout-à-fait général :

⁷Ne pas confondre avec $\mathcal{F}_\infty = \bigcup_{t \nearrow \infty} \mathcal{F}_t$ qui est la tribu complète de tous les événements mesurables par rapport à la trajectoire du mouvement brownien.

⁸Ne pas confondre \mathcal{F}_∞ et \mathcal{F}^∞ , cf. note 7.

Lemme 1.16. Soit A un événement de \mathcal{F}_∞ , soit $\varepsilon > 0$. Alors il existe un ensemble fini d'instants t_1, \dots, t_n et un événement $\tilde{A} \in \sigma(\vec{B}_{t_1}, \dots, \vec{B}_{t_n})$ tels que \tilde{A} soit « très proche » de A au sens où :

$$\mathbb{P}(A \Delta \tilde{A}) \leq \varepsilon. \quad (17)$$

On prend $\varepsilon > 0$ arbitrairement petit et \tilde{B} obtenu grâce au lemme 1.16 pour approximer B ; il nous suffit de montrer que \tilde{B} est indépendant de A . On a $\tilde{B} \in \sigma(\vec{B}_{t_1}, \dots, \vec{B}_{t_n})$, où on peut supposer que tous les t_i sont non nuls, car $\vec{B}_0 = \vec{0}$ p.s., donc la tribu $\sigma(\vec{B}_0)$ est triviale. Soit donc t^* le minimum des t_i . Comme $A \in \mathcal{F}_{0+} \subset \mathcal{F}_{t^*}$ et que, par la propriété de Markov faible, $(\vec{B}_t)_{t \geq t^*}$ ne dépend de \mathcal{F}_{t^*} qu'au travers de \vec{B}_{t^*} , il nous suffit de vérifier que \vec{B}_{t^*} est indépendant de A . Soit donc $X \subset \mathbb{R}^d$ mesurable, nous allons montrer que $\{\vec{B}_{t^*} \in X\}$ est indépendant de A .

Soit $\varepsilon \in]0, t^*[$. Pour $\vec{x} \in \mathbb{R}^d$, on pose :

$$\varphi(\varepsilon, \vec{x}) = \mathbb{P}(\vec{B}_{t^*} \in X | \vec{B}_\varepsilon = \vec{x}). \quad (18)$$

Comme on connaît explicitement les densités de transition du mouvement brownien⁽⁹⁾, il est aisé de vérifier que φ est continue en $(0, 0)$ par le théorème de convergence dominée. Comme A est dans \mathcal{F}_ε , la propriété de Markov faible au temps ε permet d'écrire :

$$\mathbb{P}(A \text{ et } \vec{B}_t \in X) = \mathbb{E}[\mathbb{1}_A \varphi(\varepsilon, \vec{B}_\varepsilon)] \xrightarrow[\varepsilon \rightarrow 0]{\text{CVD}} \mathbb{E}[\mathbb{1}_A \underbrace{\varphi(0, \vec{0})}_{\mathbb{P}(\vec{B}_{t^*} \in X)}] = \mathbb{P}(A) \mathbb{P}(\vec{B}_{t^*} \in X). \quad (19)$$

C'est l'indépendance recherchée. □

Remarque 1.17. Le théorème 1.15 signifie que le mouvement brownien n'est pas amené à « faire des choix » à un instant précis. C'est une propriété très agréable qui se retrouve chez bon nombre de processus, **contrairement à l'autre loi de 0 – 1**⁽¹⁰⁾ que nous allons maintenant énoncer.

Théorème 1.18. La tribu \mathcal{F}^∞ est triviale.

Démonstration. La démonstration la plus classique et la plus courte de ce théorème consiste à utiliser une transformation qui transforme \mathcal{F}_{0+} en \mathcal{F}^∞ . Je n'aime pas cette démonstration qui n'éclaire pas la généralité des phénomènes mis en jeu, aussi y ai-je préféré celle qui suit.

⁹Ici on considère le cas d'un mouvement brownien standard, ou plus généralement le cas où la matrice C est non dégénérée. Si C est dégénérée, on peut contourner la difficulté en se restreignant à un certain sous-espace de \mathbb{R}^d , dans lequel on a un mouvement brownien non dégénéré (de dimension inférieure à d).

¹⁰L'exemple le plus typique est celui du mouvement brownien dans l'espace hyperbolique, hélas nous n'aurons pas l'occasion de l'aborder dans cette série d'exposés.

Soit A un événement de \mathcal{F}^∞ . Pour $t \geq 0$, $\vec{x} \in \mathbb{R}^d$, on pose

$$\psi(t, \vec{x}) = \mathbb{P}(A | \vec{B}_t = \vec{x}). \quad (20)$$

Nous allons montrer que $\psi(t, \cdot)$ est uniforme ; en effet, un argument général nous permet d'affirmer que, p.s., $\psi(t, \vec{B}_t) \xrightarrow[t \rightarrow \infty]{} \mathbb{1}_A$, d'où, notant $\Psi(t)$ la valeur de $\psi(t, \vec{x})$ pour tout \vec{x} , il s'ensuit que la fonction *déterministe* $\Psi(t)$ converge p.s. vers $\mathbb{1}_A$ quand t tend vers l'infini, ce qui n'est possible que si A est de probabilité triviale.

Soient donc $t \geq 0$ un instant et \vec{x}, \vec{y} deux vecteurs distincts. Notons H l'hyperplan médiateur de $[\vec{x}, \vec{y}]$ et posons τ le temps d'atteinte de H par le mouvement brownien après le temps t :

$$\tau = \int \{u \geq t; \vec{B}_u \in H\}. \quad (21)$$

On vérifie facilement que τ est un temps d'arrêt, et, grâce aux propriétés du mouvement brownien en dimension 1, que :

Avis 1.19. τ est fini presque-sûrement.

Comme A ne dépend que du comportement du mobile après τ , la propriété de Markov forte nous donne :

$$\psi(t, \vec{x}) = \mathbb{E}[\psi(\tau, \vec{B}_\tau) | \vec{B}_t = \vec{x}]. \quad (22)$$

Mais, par symétrie du mouvement brownien, (τ, \vec{B}_τ) a la même distribution sous $\mathbb{P}(\cdot | \vec{B}_t = \vec{x})$ et sous $\mathbb{P}(\cdot | \vec{B}_t = \vec{y})$! Par conséquent, $\psi(t, \vec{x}) = \psi(t, \vec{y})$, CQFD. \square

Énonçons enfin des propriétés plus fines du mouvement brownien.

Théorème 1.20 (récurrence et transience). 1. Pour $d = 1$, presque-sûrement le mouvement brownien passe par tous les points de l'espace et y retourne toujours.

2. Pour $d = 2$, presque-sûrement le mouvement brownien visite tout ouvert de \mathbb{R}^2 et revient le visiter aussi tard qu'on veut. Par contre, presque-sûrement le mouvement brownien ne revient jamais en $\vec{0}$, et pour tout $\vec{x} \neq 0$, presque-sûrement le mouvement brownien ne passe jamais par \vec{x} .

3. Pour $d \geq 3$, presque-sûrement $|\vec{B}_t| \xrightarrow[t \rightarrow \infty]{} \infty$, et l'adhérence de la trajectoire du mouvement brownien est d'intérieur vide.

Théorème 1.21. Presque-sûrement, le mouvement brownien est localement α -höldérien pour tout $\alpha \in]0, \frac{1}{2}[$, et n'est pas localement α -höldérien pour tout $\alpha \in [\frac{1}{2}, 1]$.

Remarque 1.22. Le théorème 1.21 donc que la trajectoire du mouvement brownien est très « découpée ». L'étude des mouvements browniens amène d'ailleurs rapidement à évaluer des dimensions fractales, et c'est ainsi pour sa démonstration que l'enveloppe extérieure du mouvement brownien a pour dimension fractale $4/3$ en dimension 2, entre autres, que W. Werner a obtenu la médaille Fields en 2006.

1.3 Mouvement brownien et laplacien

Le mouvement brownien (standard) de \mathbb{R}^d est intimement lié au laplacien, et j'ose même dire qu'on ne comprend vraiment le laplacien qu'à l'aide du mouvement brownien. En fait, les liens entre mouvement brownien et laplacien, comme nous allons le voir, sont doubles : d'une part, l'opérateur $\frac{1}{2}\Delta$ est le *générateur* du mouvement brownien, en lien avec l'équation de Laplace, d'autre part c'est aussi son *générateur dual* par rapport à la mesure de Lebesgue, en lien avec l'équation de la chaleur.

1.3.1 Générateur du mouvement brownien

Définition 1.23 (générateur d'un processus markovien). Soit $(X_t)_{t \geq 0}$ un processus (au moins càd-làg) à valeurs dans un espace (polonais) \mathcal{E} , markovien, i.e. vérifiant les propriétés de Markov (1.7) et (1.9). Ce processus définit alors un semigroupe d'opérateurs $(P^t)_{t \geq 0}$ agissant sur $L^\infty(\mathcal{E})$ par :

$$P^t f(x) = \mathbb{E}[f(X_t) | X_0 = x]. \quad (23)$$

On appelle *générateur du processus markovien* X l'opérateur \mathcal{L} défini formellement par :

$$\mathcal{L}f(x) = \lim_{h \rightarrow 0} \frac{P^h f(x) - f(x)}{h}. \quad (24)$$

Moralement, le *générateur de f* dit de combien $f(X)$ a tendance à augmenter lorsqu'on se trouve au point x .

Exemple 1.24. Soit un processus markovien à valeurs dans un espace à deux points $\{a, b\}$: le processus saute de a vers b avec une loi exponentielle⁽¹¹⁾ de paramètre α , et de b vers a avec une loi exponentielle de paramètre β . Le *générateur* \mathcal{L} de ce processus associe à une fonction $f : \{a, b\} \rightarrow \mathbb{R}$ la fonction $\mathcal{L}f : \{a, b\} \rightarrow \mathbb{R}$ définie par :

$$\begin{cases} \mathcal{L}f(a) = \alpha(f(b) - f(a)); \\ \mathcal{L}f(b) = \beta(f(a) - f(b)). \end{cases} \quad (25)$$

¹¹Les lois de sauts sont nécessairement exponentielles à cause du caractère markovien du processus qui oblige les sauts à être sans mémoire.

La donnée d'un processus markovien, nous l'admettrons, est en fait équivalente à celle de son générateur : cette notion jette donc un pont essentiel entre probabilités et analyse.

Proposition 1.25. *Le générateur du mouvement brownien standard de \mathbb{R}^d est $\frac{1}{2}\Delta$.*

Démonstration. Soit $f \in C^3(\mathbb{R}^d)$, avec f''' bornée. Soit $\vec{\Xi}$ un vecteur aléatoire gaussien réduit. Alors la loi de \vec{B}_t sachant \vec{B}_0 , est celle de $\vec{B}_0 + t^{1/2}\vec{\Xi}$. Par conséquent,

$$\begin{aligned} \mathbb{E}[f(\vec{B}_t)|\vec{B}_0 = x] &= \mathbb{E}[f(\vec{B}_0) + t^{1/2}\vec{\Xi}] \\ &\stackrel{\text{Taylor}}{=} \mathbb{E}\left[f(\vec{B}_0) + t^{1/2}\vec{\nabla}_{\vec{B}_0}f \cdot \vec{\Xi} + \frac{1}{2}t\nabla^2_{\vec{B}_0}f(\vec{\Xi} \otimes \vec{\Xi}) + O(t^{3/2})|\vec{\Xi}|^3\right] \\ &= f(\vec{B}_0) + t^{1/2}\vec{\nabla}_{\vec{B}_0}f \cdot \underbrace{\mathbb{E}[\vec{\Xi}]}_{=\vec{0}} + \frac{1}{2}t\nabla^2_{\vec{B}_0}f \underbrace{\mathbb{E}[(\vec{\Xi} \otimes \vec{\Xi})]}_{=\mathbf{I}_d} + O(t^{3/2}) \\ &= f(\vec{B}_0) + t\left(\Delta_{\vec{B}_0}f + O(t^{1/2})\right), \quad (26) \end{aligned}$$

d'où le résultat. □

1.3.2 Résolution du problème de Dirichlet

Définition 1.26. *Pour \mathcal{L} le générateur d'un processus markovien, on appelle fonction \mathcal{L} -harmonique toute fonction f telle que $\mathcal{L}f = 0$.*

Une fonction \mathcal{L} -harmonique est donc une fonction vérifiant le

Lemme 1.27. *Si $(X_t)_{t \geq 0}$ est un processus markovien de générateur \mathcal{L} et f une fonction \mathcal{L} -harmonique bornée, alors on a pour tous $0 \leq t \leq t'$, pour tout $x \in E$:*

$$\mathbb{E}[f(X_{t'})|X_t = x] = f(x). \quad (27)$$

On dit que le processus aléatoire $f(X_t)$ est une *martingale*. Le lemme 1.27 est pratiquement évident à partir de la définition 1.23. Ce lemme a un corollaire puissant qui lui est ce que le corollaire 1.9 était à la proposition 1.7 :

Corollaire 1.28. *Si f est \mathcal{L} -harmonique et bornée, pour τ un temps d'arrêt fini p -s. :*

$$\mathbb{E}[f(X_\tau)|X_0 = x] = f(x). \quad (28)$$

Nous allons maintenant nous apercevoir que le mouvement brownien résout tout naturellement le problème de Dirichlet. Soit en effet U un ouvert borné de \mathbb{R}^d et \bar{f} une fonction définie sur ∂U et continue. Le problème de Dirichlet consiste à trouver une fonction f sur \bar{U} qui soit harmonique (au sens

usuel) sur U et qui coïncide avec \bar{f} sur ∂U tout en étant continue aux points de ∂U . L'analyse nous dit qu'une telle fonction f est unique. Supposons donc que nous connaissions f . Soit un mouvement brownien sur \mathbb{R}^d et définissons τ comme le *temps de sortie* de U :

$$\tau = \inf\{t \geq 0; X_t \notin U\}. \quad (29)$$

τ est le moment où le mouvement brownien quitte U et, par continuité du mouvement brownien, X_τ est donc un point de ∂U (quand le mouvement brownien part d'un point de U). Il est facile de vérifier que τ est un temps d'arrêt ; le corollaire 1.28 nous donne alors :

Proposition 1.29. *Pour tout $x \in U$, $f(x) = \mathbb{E}[f(X_\tau)]$.*

Nous sommes donc arrivés à exprimer f sur U en fonction de ses valeurs sur ∂U : le problème de Dirichlet est résolu de manière « explicite ».

Remarque 1.30. Réciproquement, on démontre simplement que, si U est un ouvert de \mathbb{R}^d et $(X_t)_{t \geq 0}$ un processus continu de générateur \mathcal{L} , pour \bar{f} une fonction définie sur ∂U , la fonction $\mathbb{E}[X_\tau | X_0 = \cdot]$ est \mathcal{L} -harmonique et vaut \bar{f} sur ∂U .

Remarque 1.31. Vous avez peut-être remarqué que j'ai passé sous silence la question de la continuité de f au bord de \bar{U} . Cette continuité paraît naturelle vu la façon dont f est définie, mais on peut se demander si elle est toujours vraie. En fait, on démontre que ce n'est le cas que si ∂U est suffisamment régulier. Plus précisément, la condition la plus usuelle qu'on utilise est la « condition de cône » de Zaremba : on veut qu'en tout point \bar{x} de ∂U , il existe un demi-cône d'intérieur non vide de sommet \bar{x} , qui soit entièrement à l'extérieur de \bar{U} . Si la condition de cône n'est pas respectée, la fonction $\mathbb{E}[X_\tau | X_0 = \cdot]$ peut ne pas être continue au bord. En fait, la condition nécessaire et suffisante pour assurer la continuité au bord, indépendamment de \bar{f} , est celle-ci : tout point \bar{x} de ∂U doit être *régulier*, ce qui signifie que, pour un mouvement brownien issu de \bar{x} :

$$\text{p.s. } \inf\{t > 0; X_t \notin U\} = 0. \quad (30)$$

En dimension 2, ainsi, un point isolé de ∂U ne sera pas régulier. Plus fort : en dimension 3, la pointe d'une « aiguille » d'équation en coordonnées cylindriques $r = e^{-1/z}$ n'est pas régulière non plus. On vérifie que, dans chacun de ces cas, la condition de cône n'est pas vérifiée (voir la figure 4).

1.3.3 Une application : transience du mouvement brownien en dimension 3.

Nous allons utiliser les résultats du § 1.3.2 pour démontrer la transience du mouvement brownien en dimension 3. Plus précisément, nous allons nous

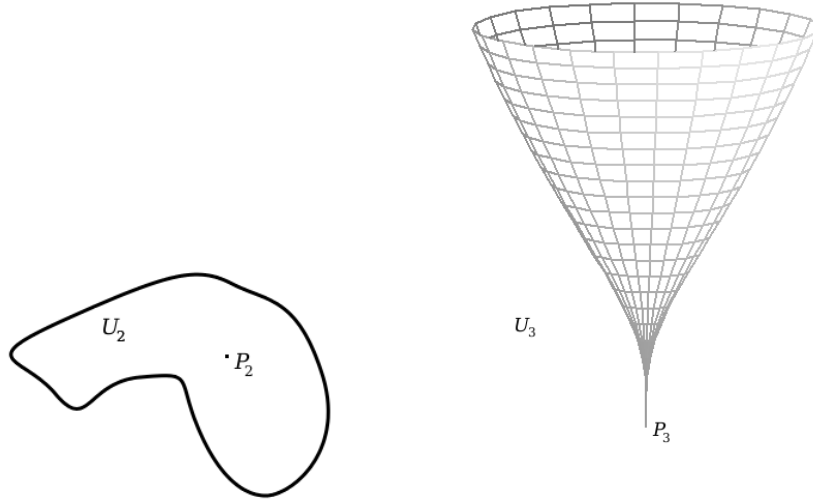


FIG. 4 – Les ouverts U_2 et U_3 ne sont respectivement pas réguliers en P_2 et P_3 .

concentrer sur la démonstration que, pour $\vec{x} \neq \vec{0}$, presque-sûrement le mouvement brownien n'a pas \vec{x} dans son adhérence, i.e. :

$$\exists \varepsilon > 0 \quad \forall t \geq 0 \quad \left| \vec{x} - \vec{B}_t \right| > \varepsilon. \quad (31)$$

Posons $R_0 = |\vec{x}|$. Pour $R > R_0$ et $\varepsilon \in]0, R_0[$, on va calculer la probabilité, notée $p(\varepsilon, R)$, que le mouvement brownien s'approche à une distance ε du point \vec{x} avant de s'en éloigner de la distance R . On note U l'ensemble des points situés à une distance de \vec{x} comprise entre ε et R (U a donc une forme de cerise dénoyautée). Nous définissons une fonction \bar{f} sur ∂U , qui vaut 1 sur $\partial B(\vec{x}, \varepsilon)$ — la surface du noyau de la cerise — et 0 sur $\partial B(\vec{x}, R)$ — la peau de la cerise. Notant τ le temps de sortie de U , c'est un simple jeu d'écritures que d'affirmer :

$$p(\varepsilon, R) = \mathbb{E} \left[\bar{f}(\vec{B}_\tau) \right]. \quad (32)$$

Mais, dans ce cas précis pour U , nous savons résoudre *exactement* le problème de Dirichlet. Plus précisément, la fonction suivante est harmonique et interpole \bar{f} , où on a noté ρ pour $|\vec{y} - \vec{x}|$:

$$f(\vec{y}) = \frac{\varepsilon R}{(R - \varepsilon)\rho} - \frac{\varepsilon}{R - \varepsilon}. \quad (33)$$

Comme la solution du problème de Dirichlet est unique, c'est donc que $f(\vec{y})$ coïncide avec $\mathbb{E} \left[\bar{f}(\vec{B}_\tau) \mid \vec{B}_0 = \vec{y} \right]$. En particulier :

$$p(\varepsilon, R) = \frac{\varepsilon(R - R_0)}{(R - \varepsilon)R_0}. \quad (34)$$

Supposons d'abord ε fixé. Notons $A_{\varepsilon,R}$ l'événement « le mouvement brownien s'approche à une distance ε du point \vec{x} avant de s'en éloigner de la distance R », de sorte que $p(\varepsilon, R) = \mathbb{P}(A_{\varepsilon,R})$. Il est clair que, pour $R \leq R'$, $A_{\varepsilon,R} \subset A_{\varepsilon,R'}$, (et, en effet, (34) montre que $p(\varepsilon, \cdot)$ est croissante). Par passage à la limite dénombrable⁽¹²⁾, il s'ensuit que l'événement $A_{\varepsilon,\infty}$, égal par définition à $\bigcup_{R>R_0}^{\nearrow} A_{\varepsilon,R}$, a pour probabilité

$$p(\varepsilon, \infty) = \lim_{R \nearrow \infty} p(\varepsilon, R) = \frac{\varepsilon}{R_0}. \quad (35)$$

Or, par continuité du mouvement brownien, l'événement $A_{\varepsilon,\infty}$ signifie que le mouvement brownien approche au moins une fois \vec{x} à une distance inférieure à ε . Pour ε suffisamment petit, cette probabilité peut être rendue aussi petite qu'on veut, d'où le résultat.

1.3.4 Générateur dual

Une chose agréable avec le mouvement brownien est que ses lois de transition sont à densité par rapport à la mesure de Lebesgue :

Définition 1.32. *On dit qu'un processus markovien $(X_t)_{t \geq 0}$ à valeurs dans \mathbb{R}^d est à densité par rapport à la mesure de Lebesgue dx s'il existe une fonction*

$$p : \mathbb{R}_+^* \times \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}_+ \\ (t, x, y) \mapsto p^t(x, y) \quad (36)$$

telle que les lois de transition aient pour densité $p^t(x, \cdot)$:

$$d\mathbb{P}(X_t = y | X_0 = x) = p^t(x, y) dy. \quad (37)$$

Dans ces conditions, on peut imaginer une assemblée de particules browniennes, dont le nombre est supposé presque infini : au cours de leur évolution, la densité locale de ces particules $\rho_t(\cdot)$ par rapport à la mesure de Lebesgue restera finie. On se demande alors quelle est la loi de l'évolution — déterministe, en vertu de la loi des grands nombres — de la fonction ρ_t . À cause du caractère markovien de l'évolution, il s'agira d'une équation différentielle linéaire du premier ordre en temps. Laquelle ? Soit f une fonction-test, l'évolution de ρ_t sera telle que :

$$\int_{\mathbb{R}^d} \rho_t(x) f(x) dx = \int_{\mathbb{R}^d} \rho_0(x) \mathbb{E}[f(X_t) | X_0 = x] dx \quad (38)$$

Dérivons par rapport au temps :

$$\int_{\mathbb{R}^d} \partial_t \rho(x) f(x) dx = \int_{\mathbb{R}^d} \rho(x) \mathcal{L}f(x) dx. \quad (39)$$

¹²D'où notre remarque sur la relation d'ordre : l'intersection que nous considérons est non dénombrable, mais se ramène à une intersection dénombrable puisque la cofinalité de l'ordre sur \mathbb{R}_+ est ω — pour le dire en termes pompeux.

Pour faire ressortir la fonction f dans l'intégrale au second membre, on utilise naturellement le concept de dualité dans l'espace de Hilbert $L^2(dx)$:

Définition 1.33. Pour \mathcal{L} un opérateur sur $L^2(dx)$, on note \mathcal{L}^* l'opérateur adjoint caractérisé par, pour toutes fonctions f et g :

$$\int_{\mathbb{R}^d} f(x)\mathcal{L}^*g(x)dx = \int_{\mathbb{R}^d} g(x)\mathcal{L}f(x)dx. \quad (40)$$

Il s'ensuit finalement que :

Proposition 1.34. La densité ρ de particules obéit à l'équation d'évolution :

$$\partial_t\rho = \mathcal{L}^*\rho. \quad (41)$$

Dans le cas du mouvement brownien, puisque $\mathcal{L} = \frac{1}{2}\Delta$, on a, d'après les formules d'intégration par parties, $\mathcal{L}^* = \frac{1}{2}\Delta$ (¹³). Ainsi, ρ évolue selon l'équation de la chaleur. C'était physiquement attendu, puisque l'équation de la chaleur régit les phénomènes de diffusion (de particules, de chaleur...), où l'on peut considérer que les objets diffusants suivent des marches aléatoires indépendantes, qui, vues de loin, sont browniennes.

1.3.5 Application : conditions au bord de Von Neuman en dimension

1

Nous allons voir dans ce paragraphe comment cette interprétation « physique » de l'équation de la chaleur peut donner naturellement la solution à certains problèmes non triviaux d'équations aux dérivées partielles.

Soit $[a, b]$ un segment non trivial de \mathbb{R} , $c \in]a, b[$. On s'intéresse au problème physique suivant : une assemblée de particules browniennes indépendantes, initialement concentrées au point c , diffuse dans le segment $[a, b]$ sans pouvoir en sortir.

Quelle est la traduction mathématique de ce problème ? Sur $]a, b[$, la densité ρ de l'assemblée de particules évoluera selon l'équation de la chaleur. Au bord, l'interdiction pour les particules de sortir se traduit par l'absence de flux de particules. On pourrait calculer facilement que celui-ci est proportionnel à $\partial_x\rho$, de sorte que nous sommes amenés à résoudre l'ÉDP :

$$\begin{cases} \rho(0, x) = \delta_c ; \\ \partial_t\rho = \frac{1}{2}\partial_x^2\rho \quad \text{sur }]a, b[; \\ \forall t \geq 0 \quad \partial_x\rho(t, a) = \partial_x\rho(t, b) = 0. \end{cases} \quad (42)$$

A priori, cette équation n'est pas facile. Mais tout devient plus simple si nous cherchons à suivre une particule à la trace. L'évolution de la particule,

¹³Pour moi, il ne faut rien voir de plus qu'une coïncidence amusante dans l'égalité entre \mathcal{L} et \mathcal{L}^* .

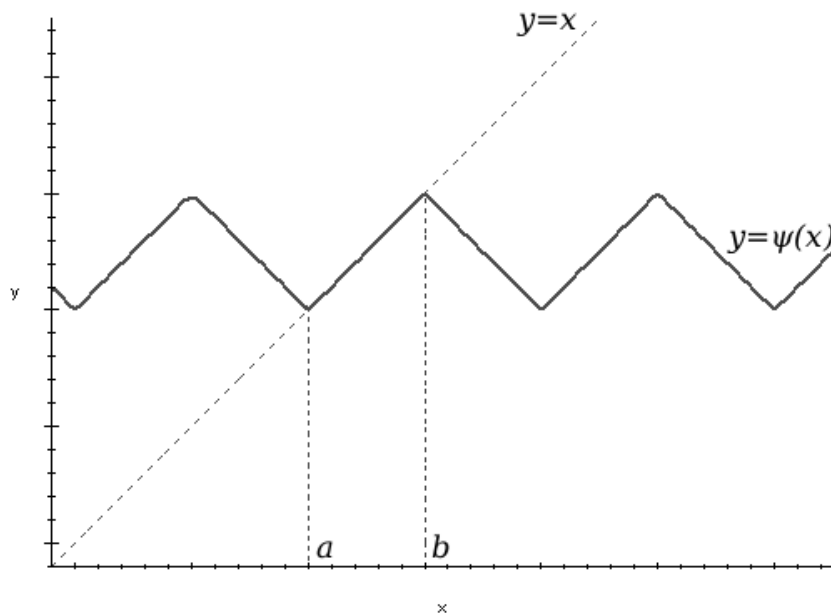


FIG. 5 – La fonction ψ

déjà, est brownienne sur $]a, b[$. Comment se comporte-t-elle au bord ? Eh bien, puisque le mouvement brownien est symétrique par retournement de l'axe des x (cf. proposition 1.10), on peut imaginer que tout se passe comme si on avait « replié » la droite réelle sur le segment $[a, b]$; concrètement :

Définition 1.35. On définit la fonction de pliage $\psi : \mathbb{R} \rightarrow [a, b]$ par $\psi(x) = x - 2n(b - a)$ si $x \in [a + 2n(b - a), b + 2n(b - a)]$ pour $n \in \mathbb{Z}$, et $\psi(x) = (b + a) - (x - (2n + 1)(b - a))$ si $x \in [a + (2n + 1)(b - a), b + (2n + 1)(b - a)]$. Voir aussi la figure 5.

Alors, si $(B_t)_{t \geq 0}$ est un mouvement brownien issu de c , la position de la particule à l'instant t sera donné par $\psi(B_t)$. Comme par ailleurs on connaît parfaitement la répartition statistique de B_t et la fonction ψ , on arrive à :

Théorème 1.36. L'(unique) solution de l'équation (42) est donnée par :

$$\rho(t, x) = \frac{1}{\sqrt{2\pi t}} \left\{ \sum_{n \in \mathbb{Z}} e^{-\frac{(x + 2n(b-a) - c)^2}{2t}} + \sum_{n \in \mathbb{Z}} e^{-\frac{((b+a) - (x + (2n+1)(b-a)) - c)^2}{2t}} \right\}. \quad (43)$$

Remarque 1.37. Une fois de plus, j'ai passé une escroquerie énorme sous le tapis. *In fine*, en effet, j'ai déclaré que la condition au bord revenait à réfléchir le mouvement brownien en a et en b . C'est vrai, mais en réalité dès que la dimension devient plus grande que 1, il peut exister d'autres façons de réfléchir le mouvement brownien, i.e. des réflexions *non orthogonales* ! Dans ces

situations, chaque fois qu'elle touche le bord la particule est (infinitésimalement) déviée dans une direction parallèle au bord, et cela conduit à une loi différente pour la trajectoire de la particule réfléchie⁽¹⁴⁾. Mais en dimension 1, il n'existe qu'une façon de réfléchir la particule au bord, et par conséquent celle que nous avons décrite est nécessairement la bonne.

2 Calcul stochastique

2.1 Concept de martingale

Pour parler de calcul stochastique, il est absolument nécessaire de connaître le concept de martingale, ce que cette partie se propose de faire.

☛ *Dans cet exposé, les martingales considérées seront toujours des martingales continues indexées par \mathbb{R}_+ .*

Pour parler de martingales, nous devons utiliser le vocabulaire des filtrations : lorsqu'on considère un processus aléatoire dépendant du temps (en l'occurrence, indexé par \mathbb{R}_+), un observateur qui ne connaît rien du processus au départ en acquiert l'information au fur et à mesure. D'où la

Définition 2.1. *On appelle filtration une suite croissante de tribus indexées par le temps $(\mathcal{F}_t)_{t \geq 0}$, telles que \mathcal{F}_t représente l'information acquise au temps t . Pour un processus $(X_t)_{t \geq 0}$, la filtration canonique associée à X est donnée par*

$$\mathcal{F}_t = \sigma(X_u; u \leq t). \quad (44)$$

Définition 2.2. *Une martingale est un processus $(M_t)_{t \geq 0}$ à valeurs dans \mathbb{R} qui n'a tendance, ni à croître, ni à décroître : pour tous $t \leq u$, on a :*

$$\mathbb{E}[M_u | \mathcal{F}_t] = M_t. \quad (45)$$

Par exemple, le mouvement brownien nous permet de définir de nombreuses martingales continues :

Proposition 2.3. *Soit $(B_t)_{t \geq 0}$ un mouvement brownien unidimensionnel, alors :*

1. $(B_t)_{t \geq 0}$ est une martingale ;
2. $(B_t^2 - t)_{t \geq 0}$ est une martingale (martingale quadratique) ;

¹⁴Que des déviations infinitésimales finissent par avoir un effet notable est, pour moi du moins, assez contre-intuitif. Pour mieux cerner ces réflexions non orthogonales, il nous faudrait parler de *temps locaux*, mais cela nous emmènerait trop loin...

3. Pour tout $a \in \mathbb{R}$, $\left(e^{aB_t - a^2 t/2}\right)_{t \geq 0}$ est une martingale (martingales exponentielles).

Démonstration. Pour $t \leq u$, on sait que la loi conditionnelle de B_u sachant B_t est $B_t + (u-t)^{1/2}\xi$, où ξ est gaussienne réduite. Par conséquent, le n° 1 signifie que $\mathbb{E}[B_u + (u-t)^{1/2}\xi] = B_t$, qui est vrai car $\mathbb{E}[\xi] = 0$, le n° 2 signifie que $\mathbb{E}\left[(B_u + (u-t)^{1/2}\xi)^2 - t\right] = B_t^2 - t$, qui découle de $\mathbb{E}[\xi] = 0$ et $\mathbb{E}[\xi^2] = 1$, et le n° 3 se ramène à vérifier la relation $\mathbb{E}\left[e^{a\xi}\right] = e^{a^2/2}$. \square

Un théorème très puissant sur les martingales est l'inégalité de Doob, qui permet de contrôler *globalement* une martingale à partir de sa valeur terminale :

Théorème 2.4 (Doob). Soit $(M_t)_{t \in [0, T]}$ une martingale continue, où T peut aussi bien être une constante qu'un temps d'arrêt (avec, dans ce dernier cas, une hypothèse d'uniforme intégrabilité cruciale). Notons $\|M\|_\infty = \sup_{t \in [0, T]} |M_t|$; alors, si M_T est une variable aléatoire de classe L^2 ⁽¹⁵⁾, $\|M\|_\infty$ l'est aussi, et

$$\mathbb{E}[\|M\|_\infty^2] \leq 4\mathbb{E}[|M_T|^2]. \quad (46)$$

2.2 Intégration contre le mouvement brownien

2.2.1 Motivations

Avant tout, rappelons le concept d'intégrale de Stieltjes. Pour μ une mesure *signée* sur \mathbb{R} , supposée ici sans atome, μ peut être décrite par une fonction $m : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \mu(-\infty, x]$. Dans ces conditions, pour $f : \mathbb{R} \rightarrow \mathbb{R}$ continue (ou, plus généralement, mesurable) et bornée, on a coutume de noter $\int f dm$ pour $\int f d\mu$, puisque $\mu([x, x+dx])$ est l'accroissement de m entre x et $x+dx$.

Il est assez courant en physique de voir des intégrales arriver directement sous la forme de Stieltjes (par exemple, le travail d'une force \vec{F} s'écrit $\int \vec{F} \cdot d\vec{x}$, où x est la position de la particule), et on est ainsi amené à se demander s'il est possible d'intégrer contre un mouvement brownien. Hélas...

Proposition 2.5. Le mouvement brownien n'est pas l'intégrale d'une mesure sur \mathbb{R} .

Démonstration. En effet, si m est l'intégrale d'une mesure signée μ , alors m

¹⁵En fait, il existe une inégalité de Doob L^p pour tout $p \in]1, \infty]$, mais ici le cas L^2 suffit à notre cause.

est à variations bornées, càd. que toutes les quantités

$$\sum_{i=1}^n |f(x_i) - f(x_{i-1})|, \quad x_0 < x_1 < \dots < x_n, n \in \mathbb{N}^* \quad (47)$$

sont bornées par une même valeur M — à savoir, la masse totale de $|d\mu|$ ⁽¹⁶⁾. Or, presque-sûrement le mouvement brownien n'est à variations bornées sur aucun segment $[a, b]$, comme nous l'allons montrer. Pour fixer les idées prenons $[a, b] = [0, 1]$, et notons V la variation totale du mouvement brownien sur ce segment. Par définition, pour $n \in \mathbb{N}^*$ arbitraire,

$$V \geq \sum_{i=1}^n \left| B_{\frac{i}{n}} - B_{\frac{i-1}{n}} \right|, \quad (48)$$

ou encore, puisque nous connaissons la loi de $(B_0, B_{1/n}, \dots, B_1)$:

$$V \geq \frac{1}{n^{1/2}} \sum_{i=1}^n |\xi_i|, \quad (49)$$

où les ξ_i sont des v.a. normales réduites indépendantes. Comme les $|\xi_i|$ ont une espérance E strictement positive, le loi des grands nombres affirme que, pour n suffisamment grand, $\sum_{i=1}^n |\xi_i|$ est plus grand que nE' , pour un $E' \in]0, E[$ arbitraire, avec une probabilité aussi grande que l'on veut. On en tire que V est plus grand que $E'n^{1/2}$ avec la même probabilité, donc, en faisant tendre n vers l'infini, V est infini presque-sûrement. \square

Question 2.6. *Peut-on quand même donner une définition sensée de $\int f dB$, où f est une fonction mesurable et $(B_t)_{t \geq 0}$ un mouvement brownien ? f doit pouvoir dépendre de la trajectoire du mouvement brownien, càd. que, si on note ω la réalisation du mouvement brownien, f peut être de la forme $f(t, \omega)$ ⁽¹⁷⁾.*

Il y a un cas où on n'hésite pas sur la définition à donner à l'intégrale : si, en effet, il existe presque-sûrement des instants $0 = t_0(\omega) < t_1(\omega) < \dots < t_{n_\omega-1}(\omega) < t_{n_\omega}(\omega) = t$ tels que, sur $[t_{i-1}(\omega), t_i(\omega)[$, f vaille $a_i(\omega)$, alors on veut poser :

$$\int_0^t f(t, \omega) dB_t = \sum_{i=1}^{n_\omega} a_i(\omega) (B_{t_i(\omega)} - B_{t_{i-1}(\omega)}). \quad (50)$$

Par conséquent, toute définition ultérieure de $\int f dB$ devra vérifier (50).

On aurait deux *desiderata* pour notre intégrale stochastique :

¹⁶Précisons qu'en fait, il suffit que μ soit une mesure signée *locale* pour pouvoir définir le concept d'intégrale de Stieltjes, auquel cas il ressort que f doit être à variations *localement* bornées.

¹⁷Nous ne rentrerons pas ici dans la question — pourtant importante ! — du type précis de fonctions f contre lesquelles l'intégration est possible. Mais il faut savoir qu'apparaît naturellement une condition dite de *mesurabilité progressive*, qui exprime en substance que $f(\cdot, t)$ est \mathcal{F}_t -mesurable, et que $f(\omega, \cdot)$ est mesurable presque-sûrement. D'autre part, les conditions d'intégrabilité à respecter par f ne sont pas les conditions L^1 habituelles, mais des conditions de type L^2 . En général, on demande à f d'être localement L^∞ presque-sûrement.

1. Si $f(t)$ est \mathcal{F}_t -mesurable, dans le mesure où dB_t est centré et indépendant de \mathcal{F}_t , on veut que le processus $\left(\int_0^t f(\omega, u)dB_u\right)_{t \geq 0}$ soit une martingale.
2. Par analogie avec le cas des intégrales de Stieltjes, si pour une certaine fonction g , f est de la forme $g'(B_t)$, on veut $\int_0^t g'(B_u)dB_u = g(B_t) - g(B_0)$.
Hélas...

Avis 2.7. *Les conditions 1 et 2 ci-dessus sont incompatibles.*

Démonstration. Si $f(\omega, t) = g'(B_t)$, pour que les conditions 1 et 2 soient vérifiées il faudrait que $(g(B_t))_{t \geq 0}$ soit une martingale, ce qui est clairement impossible si 0 est le minimum absolu strict de g . \square

Mais ne paniquons pas : en fait, les deux voies exposées ci-dessus vont simplement aboutir à deux définitions différentes de l'intégrale. La condition n° 1 aboutit à l'*intégrale d'Itô*, et la condition n° 2 à l'*intégrale de Stratonovitch*. La définition plus ou moins informelle que nous allons donner de ces deux intégrales aura au moins l'avantage d'éclairer le parallélisme direct entre les deux constructions, qui n'apparaît pas toujours dans les ouvrages de référence. Encore un mot : l'intégrale « par défaut » qu'utilisent les probabilistes est l'intégrale d'Itô. Mais il existe de nombreux cas où c'est l'intégrale de Stratonovitch qui apparaît pertinente ! Le lecteur est donc invité à garder en mémoire l'idée suivante :

☛ *Il existe deux façons de définir l'intégrale stochastique, et selon les contextes l'emploi de l'une ou l'autre est préférable.*

2.2.2 Intégrale d'Itô

Supposons donnée une fonction aléatoire $f : [0, T] \rightarrow \mathbb{R}$, continue p.s., avec $f(t)$ \mathcal{F}_t -mesurable (on dit sur ce point que f est *adaptée* à la filtration $(\mathcal{F}_t)_{t \geq 0}$). Par une manœuvre ordinaire de troncature, on peut supposer que p.s. les fonctions $f(t)$ ont un module de continuité commun h . On va simplement essayer d'écrire l'intégrale stochastique comme une limite de sommes de Riemann : soient $\sigma_1 \subset \sigma_2 \subset \dots$ des subdivisions de plus en plus fines de $[0, T]$, notées $\sigma_i = \{0 = t_0^i < t_1^i < \dots < t_{n_i}^i = T\}$, dont le pas $|\sigma_i| = \max_j(t_j^i - t_{j-1}^i)$ tend vers 0 ; on pose, si cette limite existe :

$$\int_0^T f(t)dB_t = \lim_{i \rightarrow \infty} \sum_{j=1}^{n_i} f(t_{j-1}^i)(B_{t_j^i} - B_{t_{j-1}^i}). \quad (51)$$

On peut définir de même $\int_0^u f(t)dB_t$ pour tout $u \in [0, t]$. Comme à chaque étape i , l'objet qu'on considère est une martingale, il en sera de même à la limite, conformément à notre souhait. Encore faut-il prouver que la limite existe...

Posons $S_i = \sum_{j=1}^{n_i} f(t_{j-1}^i)(B_{t_j^i} - B_{t_{j-1}^i})$; nous allons démontrer que, p.s., S_i converge au sens de la norme L^2 (il s'ensuit que, pour une sous-suite $(S_{i(i')})_{i'}$ bien choisie, la convergence est presque-sûre). Pour $i_1 < i_2$, notons

$$\{t_0^{i_2}, t_1^{i_2}, \dots, t_{n_{i_2}}^{i_2}\} = \{t_{0,0} = t_0^{i_1}, t_{0,1}, \dots, t_{0,m_0-1}, t_{0,m_0} = t_{1,0} = t_1^{i_1}, t_{1,1}, \dots, t_{n_{i_1}-1, m_{n_{i_1}}} = T\}. \quad (52)$$

On a

$$\begin{aligned} S_{i_2} - S_{i_1} &= \sum_{j=1}^{n_{i_1}} \sum_{k=1}^{m_{j-1}} f(t_{j-1,k-1}) \left[B_{t_{j-1,k}} - B_{t_{j-1,k-1}} \right] - \sum_{j=1}^{n_{i_1}} f(t_{j-1,0}) \left[B_{t_{j-1,m_{j-1}}} - B_{t_{j-1,0}} \right] \\ &= \sum_{j=1}^{n_{i_1}} \sum_{k=1}^{m_{j-1}} \left[f(t_{j-1,k-1}) - f(t_{j-1,0}) \right] \cdot \left[B_{t_{j-1,k}} - B_{t_{j-1,k-1}} \right], \quad (53) \end{aligned}$$

donc par indépendance des accroissements du mouvement brownien, $S_{i_2} - S_{i_1}$ est une variable aléatoire gaussienne dont la variance est majorée par $|\sigma_{i_1}|t$. Ainsi, la suite des variables aléatoires $(S_i)_i$ est de Cauchy, ce qu'on voulait.

Remarque 2.8. L'intégrale stochastique n'a pas seulement d'intérêt de par sa valeur *ponctuelle*, à savoir $\int_0^T f(t)dB_t$, c'est également un *processus* $u \mapsto \int_0^u f(u)dB_u$. Ici, nous avons démontré la convergence ponctuelle des sommes de Riemann, mais l'inégalité de Doob (théorème 2.4) nous affirme que la convergence des processus en résulte automatiquement.

Remarque 2.9. On voit donc apparaître l'intégrale d'Itô comme une limite L^2 . Il faudra donc garder à l'esprit que l'intégrale stochastique n'est définie que *modulo l'identité p.s.*

2.2.3 Intégrale de Stratonovitch

Ce qui « choque » dans l'approche de l'intégrale d'Itô comme limite de sommes de Riemann que nous venons de faire dans le § 2.2.2, c'est que la valeur de f que nous prenons pour faire l'intégration sur $[t_{i-1}, t_i]$ est $f(t_{i-1})$, alors que c'est un fait bien connu en calcul numérique que dans une telle situation, pour les intégrales ordinaires, l'écart entre la somme de Riemann et l'intégrale véritable est en $O(|\sigma_i|)$ mais que, si on avait évalué f en $\frac{t_{i-1}+t_i}{2}$, la convergence se ferait en $O(|\sigma_i|^2)$. Il semblerait donc plus juste de définir l'intégrale stochastique comme — si elle existe —

$$\lim_{i \rightarrow \infty} \sum_{j=1}^{n_i} f\left(\frac{t_{j-1}^i + t_j^i}{2}\right) (B_{t_j^i} - B_{t_{j-1}^i}). \quad (54)$$

Appelons *intégrale de Stratonovitch cette intégrale*, et notons-la $\int_0^t f \circ dB_t$ pour la différentiel de l'intégrale d'Itô. Précisons tout de suite que nous n'étudions pas ici la question de la convergence de l'expression (54)⁽¹⁸⁾; dans la pratique, les ouvrages préfèrent définir l'intégrale de Stratonovitch à partir de l'intégrale d'Itô et de la formule d'Itô.

Nous allons maintenant montrer que l'intégrale de Stratonovitch répond à la condition n° 2 énoncée au § 2.2.1. Soit f une fonction de classe \mathcal{C}^2 . Appelons α -intégrale stochastique, et notons $\int f' \circ_\alpha dB_t$ la limite, supposée exister,

$$\lim_{i \rightarrow \infty} \sum_{j=1}^{n_i} f'(B_{(1/2-\alpha)t_{j-1}^i + (1/2+\alpha)t_j^i}) (B_{t_j^i} - B_{t_{j-1}^i}), \quad (55)$$

et calculons $\int f' \circ_\alpha dB_t - (f(B_t) - f(B_0))$.

On essaye de passer à la limite pour calculer $\int f' \circ_\alpha dB_t - (f(B_t) - f(B_0))$; pour ce faire, on remarque que

$$\begin{aligned} & \sum_{j=1}^{n_i} f'(B_{(1/2-\alpha)t_{j-1}^i + (1/2+\alpha)t_j^i}) (B_{t_j^i} - B_{t_{j-1}^i}) - (f(B_t) - f(B_0)) \\ &= \sum_{j=1}^{n_i} \left(f'(B_{(1/2-\alpha)t_{j-1}^i + (1/2+\alpha)t_j^i}) (B_{t_j^i} - B_{t_{j-1}^i}) - (f(B_{t_j^i}) - f(B_{t_{j-1}^i})) \right) \\ & \stackrel{\text{(Taylor)}}{=} \sum_{j=1}^{n_i} \left(\alpha (B_{t_j^i} - B_{t_{j-1}^i})^2 f''\left(B_{\frac{t_{j-1}^i + t_j^i}{2}}\right) + o((B_{t_j^i} - B_{t_{j-1}^i})^2) \right). \quad (56) \end{aligned}$$

Admettons que le terme résiduel s'élimine effectivement à la limite. Comme les $(B_{t_j^i} - B_{t_{j-1}^i})^2$ sont indépendants, d'espérances respectives $t_j^i - t_{j-1}^i$, une version améliorée de la loi des grands nombres donne finalement que :

$$\int f' \circ_\alpha dB_t - (f(B_t) - f(B_0)) = \alpha \int f''(B_t) dt. \quad (57)$$

Dans le cas de l'intégrale de Stratonovitch, où $\alpha = 0$, cela montre que la formule d'intégration par parties est bien valide. Dans le cas de l'intégrale d'Itô, où $\alpha = -1/2$, ce résultat est appelé *formule d'Itô* :

Théorème 2.10 (Itô).

$$f(B_t) - f(B_0) = \int_0^t f'(B_u) dB_u + \frac{1}{2} \int_0^t f''(B_u) du. \quad (58)$$

¹⁸la condition naturelle pour intégrer f contre B serait que f soit une semimartingale (voir plus loin), ce qui est légèrement plus restrictif que pour l'intégrale d'Itô.

2.3 Calcul stochastique général

Dans cette partie, nous allons voir quel est le cadre le plus général contre lequel définir l'intégrale stochastique, ce qui nous amènera au concept de semimartingale, et à l'introduction du formalisme des différentielles stochastiques. Nous verrons également le théorème de Lévy qui nous dit qu'en fait, les martingales continues ressemblent beaucoup au mouvement brownien.

2.3.1 Variation quadratique d'une martingale continue

Dans la formule d'Itô, nous avons vu apparaître un terme supplémentaire par rapport à la formule d'intégration par parties usuelle. Une observation attentive révèle que l'existence de ce terme est due au fait que, pour une subdivision très fine $\{t_0, t_1, \dots, t_N\}$, la variation quadratique $\sum_{i=1}^N (B_{t_i} - B_{t_{i-1}})^2$ n'est pas « minuscule » (comme on s'y attendrait par analogie avec le cas des processus à variation finie), mais converge vers une certaine valeur, en l'occurrence la constante $t_N - t_0$.

Soit $(M_t)_{t \geq 0}$ une martingale continue. Alors, notant formellement $M_{t+dt} = M_t + dM_t$, on a $M_{t+dt}^2 - M_t^2 = M_t dM_t + (dM_t)^2$. Dans cette expression, la grandeur $M_t dM_t$ est centrée conditionnellement à \mathcal{F}_t par la propriété de martingale, et dM_t^2 est un « élément de variation quadratique » de M , de sorte qu'on a envie d'écrire $M_u^2 - M_0^2 = \int_0^u M_t dM_t + \int_0^u (dM_t)^2$, où $\int_0^u M_t dM_t$ est une martingale. Eh bien, c'est exactement cela !

Proposition 2.11. *Soit M une martingale continue, alors il existe un unique processus croissant $\langle M \rangle$ et une unique martingale \widetilde{M} tels que :*

$$\forall t \geq 0 \quad \langle M_t^2 - M_0^2 \rangle = \langle M \rangle_t + \widetilde{M}_t. \quad (59)$$

$\langle M \rangle_t$ est la limite dans L^2 , pour des subdivisions de pas tendant vers 0, de $\sum_{j=1}^{n_i} (M_{t_j^i} - M_{t_{j-1}^i})^2$.

Remarque 2.12. Ainsi, le mouvement brownien standard de \mathbb{R} a pour variation quadratique $\langle B \rangle_t = t$ p.s..

Démonstration. On vérifie à partir de la propriété de martingale de M que, pour $t' \geq t$, $\mathbb{E}[M_{t'}^2 | \mathcal{F}_t] \geq M_t^2$: on dit que M^2 est une *sous-martingale*⁽¹⁹⁾. Dans ce cas, la *décomposition de Doob-Meyer* affirme qu'il est possible de décomposer M^2 comme la somme d'une martingale et d'un processus croissant continu. (La décomposition de Doob-Meyer est triviale dans le cas où la surmartingale est indexée par \mathbb{N} , mais demande quelques outils techniques dans le cas général). En outre, cette décomposition est unique, car la différence

¹⁹Le signe de la nomenclature apparaît souvent un peu curieux au premier abord... On le comprend en disant qu'en moyenne, M est *en dessous* de la valeur moyenne qu'elle aura plus tard.

de deux processus croissants est un processus à variation bornée, alors que la différence de deux martingales continues est une martingale continue, laquelle ne peut être à variation bornée que si elle est nulle — un exercice très formateur... \square

Ce qui est important, c'est que, pour une martingale continue, la variation quadratique reste finie presque-sûrement. C'est ce qui permet, *in fine*, de construire le calcul stochastique contre une martingale continue exactement de la même façon que contre le mouvement brownien. Aussi nous contenterons-nous d'exposer directement les résultats-clefs du calcul stochastique général sans les démontrer.

2.3.2 Intégration contre une martingale continue

Soit $(M_t)_{t \geq 0}$ une martingale continue, $(\mathcal{F}_t)_{t \geq 0}$ sa filtration associée, $(H_t)_{t \geq 0}$ un processus adapté, localement borné p.s.. Alors l'intégrale stochastique d'Itô $\int H dM$ est une martingale caractérisée par :

Théorème 2.13. *Pour $(\sigma_i)_{i \geq 1}$ une suite croissante de subdivisions de $[0, t]$ de pas tendant vers 0, on a*

$$\sum_{j=1}^{n_i} H_{t_{j-1}^i} (M_{t_j^i} - M_{t_{j-1}^i}) \xrightarrow{i \rightarrow \infty} \int_0^t H_u dM_u. \quad (60)$$

Pour $F : \mathbb{R} \rightarrow \mathbb{R}$ une fonction \mathcal{C}^2 , on a :

Théorème 2.14 (formule d'Itô).

$$F(M_t) - F(M_0) = \int_0^t F'(M_u) dM_u + \frac{1}{2} \int_0^t F''(M_u) d\langle M \rangle_u. \quad (61)$$

Théorème 2.15. *La variation quadratique de l'intégrale d'Itô $\int_0^t H_u dM_u$, notée pour le coup $((H \cdot M)_t)_{t \geq 0}$ est donnée par :*

$$\langle (H \cdot M) \rangle_t = \int_0^t H_u^2 d\langle M \rangle_u \quad (20), \quad (62)$$

ce qui peut se condenser sous la forme $d\langle (H \cdot M) \rangle = H^2 d\langle M \rangle$.

2.3.3 Le théorème de Lévy

Le théorème de Lévy nous dit en substance que toutes les martingales continues ressemblent au mouvement brownien.

²⁰Noter qu'on voit ici apparaître la condition générale d'intégrabilité pour définir l'intégrale d'Itô : p.s. H doit être dans $L^2(d\langle M \rangle)$.

Théorème 2.16 (Lévy). *Si $(M_t)_{t \geq 0}$ est une martingale issue de 0 dont la variation quadratique est $d\langle M \rangle = dt$, alors c'est un mouvement brownien.*

Démonstration. Il faut essentiellement démontrer que, pour $t \geq 0$, $M_t \stackrel{\text{loi}}{\sim} \mathcal{N}(t)$. Or pour connaître une loi, il suffit d'en connaître la transformée de Fourier (la définition de la transformée de Fourier d'une loi est rappelée *infra*). Soit donc $a \in \mathbb{R}$ et notons $F_a(t) = \mathbb{E}[e^{iaM_t}]$. La formule d'Itô donne :

$$e^{iaM_t} - e^{iaM_0} = ia \int_0^t e^{iaM_u} dM_u - \frac{a^2}{2} \int_0^t e^{iaM_u} du, \quad (63)$$

d'où en passant à l'espérance :

$$F_a(t) - 1 = -\frac{a^2}{2} \int_0^t F_a(u) du. \quad (64)$$

□

Comme en outre, $F_a(0) = 1$, $F_a(\cdot)$ est donc solution de l'équation différentielle :

$$\begin{cases} F_a(0) = 1; \\ F_a' = -\frac{a^2}{2} F_a, \end{cases} \quad (65)$$

et donc finalement $F_a(t) = e^{-a^2 t/2}$: c'est bien la transformée de Fourier d'une variable $\mathcal{N}(t)$.

Définition 2.17. *On rappelle que, pour X une variable aléatoire à valeurs réelles, la transformée de Fourier de X est l'application*

$$\begin{aligned} \widehat{X} : \mathbb{R} &\longrightarrow \{z \in \mathbb{C} \mid |z| \leq 1\} \\ a &\longmapsto \mathbb{E}[e^{iaX}]. \end{aligned} \quad (66)$$

On démontre que la transformée de Fourier d'une variable aléatoire détermine complètement sa loi.

Corollaire 2.18. *Soit $(M_t)_{t \geq 0}$ une martingale continue. Alors il existe un mouvement brownien $(B_s)_{s \geq 0}$ tel qu'on puisse passer de B à M par reparamétrisation du temps :*

$$M_t = B_{\langle M \rangle_t}. \quad (67)$$

Démonstration. Soit $s : t \mapsto \langle M \rangle_t$; il nous suffit de remarquer que $(M_{s^{-1}(s)})_{s \geq 0}$ est une martingale de variation quadratique s ⁽²¹⁾. □

²¹En effet, de par sa définition (cf. § 2.3.1), la variation quadratique ne change pas quand on reparamétrise le temps.

2.3.4 Semimartingales

Pour parachever « notre » œuvre — la définition de l'intégrale stochastique — nous allons regrouper l'intégrale stochastique et l'intégrale de Stieltjes en une seule définition.

Définition 2.19. On appelle semimartingale un processus continu $(X_t)_{t \geq 0}$ qui se décompose comme la somme d'une martingale continue $(M_t)_{t \geq 0}$ et d'un processus continu à variations localement bornées $(A_t)_{t \geq 0}$. Cette décomposition est appelée décomposition de Doob-Meyer, et elle est unique car une martingale continue à variations bornées est nulle, comme nous l'avons déjà fait remarquer p. 147 à la fin de la preuve de la proposition 2.11.

Définition 2.20. Pour $X = M + A$ une semimartingale et sa décomposition de Doob-Meyer, H un processus adapté, on définit l'intégrale d'Itô comme

$$\int H dX = \int H dM + \int H dA. \quad (68)$$

Définition 2.21. La variation quadratique d'une semimartingale est la variation quadratique de sa partie martinagale ; elle s'interprète, de la même façon que pour une martingale, comme la somme des carrés des accroissements infinitésimaux.

La formule d'Itô est encore valable avec les semimartingales :

Théorème 2.22. Pour F de classe C^2 , $(X_t)_{t \geq 0}$ une semimartingale :

$$F(X_t) - F(X_0) = \int_0^t F'(X_u) dX_u + \frac{1}{2} \int_0^t F''(X_u) d\langle X \rangle_u. \quad (69)$$

Avant de voir ce que devient l'intégrale de Stratonovitch dans le cadre des semimartingales, nous avons besoin de définir la variation quadratique croisée :

Définition 2.23. Pour X, Y deux semimartingales, la variation quadratique croisée de X et Y , notée $\langle X, Y \rangle$ est, moralement, la somme des produits des accroissements infinitésimaux de X et Y . On a :

$$\langle X, Y \rangle = \frac{1}{2} (\langle X + Y \rangle - \langle X \rangle - \langle Y \rangle). \quad (70)$$

Définition 2.24. Pour X et Y deux semimartingales adaptées à la même filtration, l'intégrale de Stratonovitch $\int X \circ dY$ est définie par :

$$\int_0^t X_u \circ dY_u = \int_0^t X_u dY_u + \frac{1}{2} \langle X, Y \rangle_t. \quad (71)$$

Comme précédemment, l'intégrale de Stratonovitch vérifie la définition intuitive (54) et la formule d'intégration par parties. Plus généralement, toutes les

formules d'intégration par parties sont valides avec l'intégrale de Stratonovitch ; par exemple, pour deux semimartingales X et Y :

$$X_t Y_t - X_0 Y_0 = \int_0^t X_u dY_u + \int_0^t Y_u dX_u. \quad (72)$$

Remarque 2.25. Tout le formalisme du calcul stochastique que nous venons de définir se généralise au cas multidimensionnel : définition des intégrales d'Itô et de Stratonovitch, formule d'Itô et formule d'intégration par parties pour l'intégrale de Stratonovitch, théorème de Lévy, concept de semimartingale. À noter que, dans le cadre multidimensionnel, les variations quadratiques deviennent des matrices (symétriques, et positives dans le cas de variations non croisées) ; notamment la variation quadratique du mouvement brownien est tI .

2.3.5 Formalisme différentiel stochastique

Pour la commodité de la compréhension et la facilité des calculs, il est très pratique de manipuler les différentielles stochastiques comme des entités propres et non comme de simples notations. Limitons-nous au cas du calcul d'Itô. On retiendra en particulier que, pour $X = A + M$ et $Y = B + N$ deux semimartingales et leurs décompositions de Doob-Meyer, adaptées à la même filtration $(\mathcal{F}_t)_{t \geq 0}$, on peut faire les manipulations formelles :

- $\mathbb{E}[Y_t dM_t | \mathcal{F}_t] = 0$;
- $\mathbb{E}[Y_t dA_t | \mathcal{F}_t] = Y_t dA_t$;
- $dX dY = d\langle X, Y \rangle$.

Le dernier point est crucial entre tous : il exprime que les différentielles stochastiques, plus précisément leur partie martingale, **doivent être considérées comme des infiniment petits d'ordre 1/2**. En particulier, le produit de deux différentielles stochastiques ne peut pas être négligé et explique le terme résiduel qui apparaît dans la formule d'Itô ; en effet la formule d'Itô est simplement l'intégrale du développement de Taylor :

$$dF(X_t) = F' dX_t + \frac{1}{2} F''(X_t) (dX_t)^2 = F' dX_t + \frac{1}{2} F''(X_t) d\langle X \rangle_t. \quad (73)$$

2.4 Invariance conforme du mouvement brownien

☞ Dans cette partie, \mathbb{R}^2 muni de sa structure euclidienne est identifié à \mathbb{C} .

2.4.1 Le mouvement brownien est invariant conforme en dimension

2

Théorème 2.26. *Soit $\varphi : \mathbb{C} \rightarrow \mathbb{C}$ une application holomorphe, et $(X_t)_{t \geq 0}$ un mouvement brownien dans \mathbb{C} . Alors la trajectoire de $(\varphi(X_t))_{t \geq 0}$ est celle d'un mouvement brownien, au sens où il existe un mouvement brownien $(Y_t)_{t \geq 0}$ dans \mathbb{C} et une application aléatoire $s : \mathbb{R} \rightarrow \mathbb{R}$ continue et croissante p.s. tels qu'on ait :*

$$\varphi(X_t) = Y_{s(t)}. \quad (74)$$

Démonstration. Cette preuve sera pour nous l'occasion de montrer la puissance extraordinaire des notations du calcul stochastique : en effet, plutôt que de nous encombrer avec des considérations techniques, nous allons nous contenter de manipuler des différentielles stochastiques, *en gardant à l'esprit que ces notations dénuées de sens rigoureux ne font que traduire un formalisme plus lourd, mais qui est lui parfaitement exact.* Le calcul stochastique utilisé ici est celui d'Itô.

Considérons φ comme une application de \mathbb{R}^2 dans \mathbb{R}^2 ; φ a une dérivée première $D\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, ainsi qu'une dérivée seconde $D^2\varphi : \mathcal{T} \rightarrow \mathbb{R}^2$, où \mathcal{T} est l'ensemble des tenseurs de $\mathbb{R}^2 \otimes \mathbb{R}^2$. Comme à l'accoutumée, cette dérivée seconde est évaluée contre $\vec{e}_x \otimes \vec{e}_x + \vec{e}_y \otimes \vec{e}_y$ est appelée le *laplacien* de φ , que nous noterons $\vec{\Delta}\varphi$.

Alors on a, d'après la formule d'Itô :

$$\overrightarrow{d\varphi(X_t)} = D\varphi(X_t) \cdot dX_t + \frac{1}{2} D^2\varphi(X_t) \cdot d\langle X \rangle_t. \quad (75)$$

Mais, puisque $(X_t)_{t \geq 0}$ est un mouvement brownien standard du plan, on a $d\langle X \rangle_t = (\vec{e}_x \otimes \vec{e}_x + \vec{e}_y \otimes \vec{e}_y) dt$, d'où finalement

$$\overrightarrow{d\varphi(X_t)} = D\varphi(X_t) \cdot dX_t + \frac{1}{2} \vec{\Delta}\varphi(X_t) dt. \quad (76)$$

Or, c'est un résultat bien connu, comme φ est holomorphe on a $\vec{\Delta}\varphi \equiv 0$. En outre, φ est également conforme, i.e. $\forall X \in \mathbb{R}^2 \frac{1}{\det D\varphi(X)} D\varphi(X) \in O^+(\mathbb{R}^2)$, de sorte que $\frac{1}{\det D\varphi(X_t)} D\varphi(X_t) \cdot dX_t$ est l'accroissement d'un mouvement brownien (par isotropie du mouvement brownien standard). Ainsi, $(\varphi(X_t))_{t \geq 0}$ est une martingale (bidimensionnelle) de covariation quadratique

$$d\langle \varphi(X) \rangle_t = (\det D\varphi(X_t))^2 (\vec{e}_x \otimes \vec{e}_x + \vec{e}_y \otimes \vec{e}_y) dt. \quad (77)$$

Par conséquent, posant

$$s(t) = \int_0^t (\det D\varphi(X_s))^2 ds, \quad (78)$$

$(\varphi(X_{s^{-1}(s)}))_{s \geq 0}$ est une martingale de covariation quadratique $(\vec{e}_x \otimes \vec{e}_x + \vec{e}_y \otimes \vec{e}_y) ds$, c'est-à-dire un mouvement brownien par le théorème de Lévy. Appelons Y ce mouvement brownien, alors le théorème 2.26 est vérifié. \square

Remarque 2.27. On notera que la nullité de $\vec{\Delta}\varphi$ a été utilisée de façon fondamentale dans cette preuve. Ainsi, **en dimension différente de 2, le mouvement brownien n'est pas invariant conforme**, car le terme d'Itô qui intervient dans le calcul de $d\varphi(X_t)$ est non nul. Par exemple, en dimension 1, où toutes les applications sont conformes, le milieu d'un segment est le point tel que, partant de ce segment, un mouvement brownien atteindra chaque extrémité du segment la première avec probabilité 1/2. Donc, si le théorème 2.26 était vrai en dimension 1, cela signifierait que les applications conformes préservent le milieu, ce qui est faux.

2.4.2 Application : représentation conforme de Riemann

Théorème 2.28 (Riemann). *Tous les ouverts simplement connexes et non pleins de \mathbb{C} sont conformément équivalents. Le bord orienté d'un de ces ouverts dans $\widehat{\mathbb{C}}$ est donc homéomorphe à \mathbb{S}^1 ; si nous fixons trois points distincts de ce bord⁽²²⁾ avec un certain ordre cyclique, alors le biholomorphisme entre deux tels ouverts munis de ces trois points est unique.*

Notre objectif ici va être d'explicitier en termes probabilistes le biholomorphisme entre n'importe quel ouvert U du type décrit ci-dessus muni de trois points A_-, A_+, A_∞ sur son bord, et un ouvert particulier U^0 , en l'occurrence le demi-plan supérieur doté des trois points $-1, 1, \infty$.

On rappelle que le mouvement brownien de dimension 2 est récurrent, et visite donc presque-sûrement le voisinage de tout point. Comme en plus, partant d'un certain point, il peut suivre localement n'importe quelle trajectoire à une déformation \mathcal{C}^0 près (théorème de support), il s'ensuit que toute courbe $c : [0, 1] \rightarrow \mathbb{C}$ sera presque-sûrement \mathcal{C}^0 -approchée par le mouvement brownien au moins une fois, et donc, finalement, que

Avis 2.29. *Presque-sûrement le mouvement finit par quitter U .*

On peut décomposer le bord orienté de U en trois intervalles $I_1 = [A_\infty, A_-]$, $I_2 = [A_-, A_+]$, $I_3 = [A_+, A_\infty]$. Notons $p_i(x)$ la probabilité pour le mouvement brownien partant de $x \in U$ de quitter U par l'intervalle I_i .

Comme la trajectoire du mouvement brownien, à reparamétrisation du temps près, est invariante conforme, on a :

Proposition 2.30. *Le triplet $(p_1(\cdot), p_2(\cdot), p_3(\cdot))$ est invariant conforme.*

Maintenant nous allons démontrer que, pour l'ouvert $U_0 = \{z \in \mathbb{C} \mid \Re(z) > 0\}$ avec $A_-^0 = -1, A_+^0 = 1, A_\infty^0 = \infty$, l'application (p_1, p_2, p_3) réalise une bijection

²²Je passe sous silence le fait que certains bords trop « laids » peuvent dégénérer par l'application de notre biholomorphisme ; ici on supposera que ce n'est jamais le cas, par exemple parce que le bord de notre ouvert est une courbe \mathcal{C}^1 dans $\widehat{\mathbb{S}}$.

de U_0 sur le triangle $T = \{(p_1, p_2, p_3) ; \forall i p_i > 0 \text{ et } \sum_i p_i = 1\}$, dont nous expliciterons la bijection inverse, de sorte que pour un point $x \in U$, on pourra par le calcul des $p_i(x)$ en déduire l'image de x par le biholomorphisme recherché, à savoir l'image réciproque de $(p_1(x), p_2(x), p_3(x))$ par $(p_1, p_2, p_3) : U^0 \longrightarrow T$.

Soit $x \in U_0$. Pour calculer $p_1(x)$, resp $p_3(x)$, il suffit par invariance par translation se savoir calculer $p_-(x+1)$, resp. $p_+(x-1)$, où on pose :

$$\tau = \inf\{t \geq 0; \Im(X_t) = 0\}, \quad (79)$$

puis

$$p_\varepsilon(x) = \mathbb{P}_x(\text{sgn } \Re(X_\tau) = \varepsilon) \quad (23). \quad (80)$$

Calculons donc la fonction $p_-(\cdot)$. On remarque tout de suite :

Avis 2.31. $p_-(z)$ ne dépend que de l'argument de z .

Démonstration. C'est dû à l'invariance de la trajectoire du mouvement brownien par homothétie. \square

D'autre part,

Avis 2.32. $p_-(\cdot)$ est harmonique.

Par l'avis 2.31, $p_-(\cdot)$ est de la forme $p_-(\theta)$ en coordonnées polaires, donc son laplacien est $\frac{1}{r^2}p''(\theta)$. D'après l'avis 2.32, il en résulte que $p''(\theta) = 0$. Comme par ailleurs il est clair que $p_-(0) = 0$, $p_-(\pi) = 1$, on a finalement $p_-(\theta) = \frac{1}{\pi}\theta$, soit :

$$p_-(z) = \frac{\arg(z)}{\pi}. \quad (81)$$

Repassons en coordonnées réelles. Soit $z = (x, y) \in \mathbb{R}^2$, on suppose connus $p_1(z), p_2(z), p_3(z)$. D'après la formule (81) et la valeur de $p_1(z)$, z est sur la demi-droite d'équation

$$(x, y) = (-1, 0) + \lambda(\cos(\pi p_1), \sin(\pi p_1)), \quad \lambda > 0. \quad (82)$$

De même, d'après la valeur de p_3 :

$$(x, y) \in \{(1, 0) + \lambda(\cos(\pi(1-p_3)), \sin(\pi(1-p_3)))\}. \quad (83)$$

Finalement on obtient, après simplification :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sin(\pi p_2)} \begin{pmatrix} \sin(p_3 - p_1) \\ 2 \sin p_1 \sin p_3 \end{pmatrix}. \quad (84)$$

²³On a $p_+ + p_- \equiv 1$ car, p.s., $X_\tau \neq 0$.

Références

- [1] <http://www.dma.ens.fr/~legall/IPPA2.pdf>, pp. 219 à 246, 2006.
- [2] R. F. Bass, *Probabilistic Techniques in Analysis*, Springer-Verlag, 1995.
- [3] http://fr.wikipedia.org/wiki/Robert_Brown_%28botaniste%29
- [4] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 3e édition, Springer, 1999.

Document VII

A probabilistic approach to Carne's bound

A probabilistic approach to Carne's bound

Rémi Peyre

29 November 2006

Abstract

Carne's bound is a sharp bound controlling the transition probabilities for a discrete reversible Markov chain (§ 1). Its ordinary proof (§ 2) uses spectral techniques which look as efficient as miraculous. We devised a new proof using reversibility to compare a "drift" for ways "out" and "back" (§ 3). Actually, our result lacks a flight factor which we use a conditioning technique to get again (§ 5). Before that, we show how our proof is more "supple" than Carne's one and may generalize (§ 4.2).

1 Introduction

1.1 The Markov chain

Let V be a finite or countable set of points. Let us consider an irreducible Markov chain $(X_t)_{t \in \mathbb{N}}$ on V , whose transition kernel is $(p(x, y))_{x, y \in V}$, and whose law is denoted by \mathbb{P}_x when starting at x . That chain is supposed to be *reversible*, i.e. we suppose that there exists a measure μ on V with $\forall x \in V \ 0 < \mu(x) < \infty$ such that

$$\forall x, y \in V \quad \mu(x)p(x, y) = \mu(y)p(y, x). \quad (1.1)$$

By irreducibility, μ is then uniquely determined up to a multiplicative factor⁽¹⁾.

Then one may associate to the kernel a (non-oriented) graph (V, E) with vertices set V by defining the set of edges through

$$\{x, y\} \in E \Leftrightarrow p(x, y) \neq 0^{(2)}. \quad (1.2)$$

⁽¹⁾Note though that $\mu(V)$ may as well be finite as infinite, depending on the recurrence properties of the chain \mathbb{P} .

⁽²⁾*A priori* that definition should determine an *oriented* graph, but actually $p(x, y) \neq 0 \Leftrightarrow p(y, x) \neq 0$ by reversibility.

As usual, we shall write $z \sim z'$ to mean that $\{z, z'\} \in V^{(3)}$. The graph distance, denoted by d , will stand for the length of the shortest path(s) in E joining two points. Speaking in terms of probability, one has:

$$d(x, y) = \inf \{t \in \mathbb{N}; p^t(x, y) \neq 0\}, \quad (1.3)$$

where p^t denotes the t -th convolution power of the kernel p .

This paper aims at explaining by probabilistic arguments a bound due to Carne to sharply bound $p^t(x, y)$ above when $d(x, y) \gtrsim \sqrt{t}$. Indeed to the best of our knowledge, all the methods used so far to get that kind of bounds indeed used spectral analysis techniques [1, 2]. We shall also show how our probabilistic approach allows us to generalize Carne–Varopoulos type bounds for more “flexible” distances than the graph distance.

1.2 Historical review

In 1985, N. Th. Varopoulos [1] was the first to give a concentration result bounding $p^t(x, y)$ above for a reversible Markov chain, whose leading term was $\exp\left(-\frac{d(x, y)^2}{Ct}\right)$, $C > 0$ being an explicit constant depending on the transition kernel p . His method introduced a time-continuous Markov process on the cabled graph associated to (V, E) , and studied the spectral properties of that process in an L^2 space. Moreover that proof required extra assumptions about the transition kernel.

The same year, T. K. Carne [2], by a simpler spectral method, got a finer result under the general assumptions stated in § 1.1:

Theoreme 1.1 (Carne 1985). *Suppose the hypotheses of § 1 are satisfied. Denote by P the $L^2(\mu)$ -operator associated to the transition kernel p and let $|P|$ stand for its norm, which is always ≤ 1 (see a more precise definition in § 2.1). Then:*

$$p^t(x, y) \leq 2 \left(\frac{\mu(y)}{\mu(x)}\right)^{1/2} |P|^t \exp\left(\frac{-d(x, y)^2}{2t}\right). \quad (1.4)$$

My work was motivated by two goals: first, find a less “miraculous” proof of theorem 1.1 than the original proof of Carne, then, adapt Carne–Varopoulos type bounds so that they would apply to distances which depend continuously on the transition kernel (see § 4.2).

2 Carne’s proof

We give here the proof of [2] as it was exposed in [3].

⁽³⁾Look out for the fact that \sim is *not* an equivalence relation.

2.1 Norm of the transition kernel

Let us first give a precise definition of $|P|$:

Definition 2.1. Let μ be an equilibrium measure for the Markov chain. Then \mathbb{P} induces an operator P on $L^2(\mu)$ through:

$$Pf(x) = \mathbb{E}_x[f(X_1)] = \sum_{y \sim x} p(x, y)f(y), \quad (2.1)$$

Then we define $|P|$ as the operator norm of P in $L^2(\mu)$.

Remark 2.2. P is self-adjoint by reversibility of μ , and $|P| \leq 1$ by Jensen's inequality.

A more intrinsic definition of $|P|$ is given by the classical

Lemma 2.3. Let any $x \in V$, then we have:

$$|P| = \limsup_{t \rightarrow \infty} (p^t(x, x))^{1/t} = \sup_{t \geq 1} (p^t(x, x))^{1/t}. \quad (2.2)$$

Proof. See [3], chap. 5-2. □

2.2 Chebychev's polynomials

Since $P^n f(x) = \mathbb{E}_x[f(X_n)]$, one can write:

$$p^t(x, y) = \left\langle \frac{\delta_x}{\mu(x)}, P^t \delta_y \right\rangle_{L^2(\mu)} = \frac{|P|^t}{\mu(x)} \left\langle \delta_x, \left(\frac{P}{|P|} \right)^t \delta_y \right\rangle_{L^2(\mu)}. \quad (2.3)$$

The trick then consists in decomposing the polynomial Z^t in the basis of Chebychev's polynomials. The following results are classical:

Lemma 2.4. For any $k \in \mathbb{Z}$, there exists a unique polynomial $Q_k(Z)$ satisfying

$$\forall \theta \in \mathbb{C} \quad Q_k(\cos \theta) = \cos(k\theta), \quad (2.4)$$

called the k -th (first type) Chebychev polynomial. It satisfies:

1. $\deg Q_k = |k|$;
2. $|x| \leq 1 \Rightarrow |Q(x)| \leq 1$;
3. $\forall t \in \mathbb{N} \quad Z^t = \frac{1}{2^t} \sum_{k \in \mathbb{Z}} \binom{t}{(t+k)/2} Q_k(Z)$, where by convention $\binom{t}{p} = 0$ whenever $p \notin \{0, 1, \dots, t\}$.

By the point 3 of Lemma 2.4, formula (2.3) gives:

$$p^t(x, y) = \frac{|P|^t}{2^t \mu(x)} \sum_{k \in \mathbb{Z}} \binom{t}{(t+k)/2} \left\langle \delta_x, Q_k \left(\frac{P}{|P|} \right) \delta_y \right\rangle_{L^2(\mu)}. \quad (2.5)$$

The linear operator $\frac{P}{|P|}$ on $L^2(\mu)$ is self-adjoint and its norm is 1 by construction; so it decomposes onto a countable orthonormal basis of eigenvectors as:

$$\frac{P}{|P|} \left(\sum_{\lambda \in \text{Spec}(P/|P|)} a_\lambda \mathbf{v}_\lambda \right) = \sum_{\lambda \in \text{Spec}(P/|P|)} \lambda a_\lambda \mathbf{v}_\lambda, \quad (2.6)$$

where \mathbf{v}_λ is the eigenvector associated to the eigenvalue λ , the eigenvalues being counted with multiplicity. By definition of $|P|$ we have $\text{Spec} \left(\frac{P}{|P|} \right) \subset [-1, 1]$. So

$$Q_k \left(\frac{P}{|P|} \right) \left(\sum_{\lambda \in \text{Spec}(P/|P|)} a_\lambda \mathbf{v}_\lambda \right) = \sum_{\lambda \in \text{Spec}(P/|P|)} Q_k(\lambda) a_\lambda \mathbf{v}_\lambda, \quad (2.7)$$

where the $Q_k(\lambda)$ are all of absolute value less than or equal to 1 by lemma 2.4, point 2. Then, the operator norm of $Q_k \left(\frac{P}{|P|} \right)$ on $L^2(\mu)$ is less than or equal to 1, which enables us to write:

$$\left\langle \delta_x, Q_k \left(\frac{P}{|P|} \right) \delta_y \right\rangle_{L^2(\mu)} \leq \|\delta_x\|_{L^2(\mu)} \cdot \|\delta_y\|_{L^2(\mu)} = \sqrt{\mu(x)\mu(y)}. \quad (2.8)$$

Now, we notice that for $|k| < d(x, y)$, $Q_k \left(\frac{P}{|P|} \right) \delta_y$ is a linear combination of the $P^u \delta_y$, $0 \leq u < d(x, y)$, by point 1 of Lemma 2.4; then $Q_k \left(\frac{P}{|P|} \right) \delta_y$ is a function supported by the $z \in V$ satisfying $d(z, y) < d(x, y)$, in particular $\left\langle \delta_x, Q_k \left(\frac{P}{|P|} \right) \delta_y \right\rangle_{L^2(\mu)} = 0$. It finally gives:

$$\begin{aligned} p^t(x, y) &= \frac{|P|^t}{2^t \mu(x)} \sum_{|k| \geq d(x, y)} \binom{t}{(t+k)/2} \left\langle \delta_x, Q_k \left(\frac{P}{|P|} \right) \delta_y \right\rangle_{L^2(\mu)} \\ &\stackrel{(2.8)}{\leq} \frac{|P|^t}{2^t \mu(x)} \sqrt{\mu(x)\mu(y)} \sum_{|k| \geq d(x, y)} \binom{t}{(t+k)/2} \\ &\leq 2|P|^t \left(\frac{\mu(y)}{\mu(x)} \right)^{1/2} \frac{1}{2^t} \sum_{k \geq d(x, y)} \binom{t}{(t+k)/2} \end{aligned} \quad (2.9)$$

(where the last inequality is an equality as soon as $d(x, y) > 0$).

To conclude, it just remains to prove the relation:

$$\frac{1}{2^t} \sum_{k \geq d(x, y)} \binom{t}{(t+k)/2} \leq \exp(-d(x, y)^2/2t). \quad (2.10)$$

To do that, we notice that, if X is a random variable equidistributed on $\{-1, 1\}$, then, denoting by X^{*t} the t -th convolution power of X , X^{*t} 's law is $\frac{1}{2^t} \sum_{k \in \mathbb{Z}} \binom{t}{(t+k)/2} \delta_k$, and so

$$\frac{1}{2^t} \sum_{k \geq d(x,y)} \binom{t}{(t+k)/2} = \mathbb{P}(X^{*t} \geq d(x,y)). \quad (2.11)$$

Now, we check by direct computation that for all $\lambda > 0$, $\mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2/2}$, hence $\mathbb{E}[e^{\lambda X^{*t}}] \leq e^{t\lambda^2/2}$, and then by Chebychev's inequality:

$$\mathbb{P}(X^{*t} \geq d(x,y)) = \mathbb{P}\left(e^{\lambda X^{*t}} \geq e^{\lambda d(x,y)}\right) \leq \frac{e^{t\lambda^2/2}}{e^{\lambda d(x,y)}}, \quad (2.12)$$

hence we get (2.10) by taking $\lambda = d(x,y)/t$, which ends the proof.

3 The Gaussian factor

In this section, I will present a new, probabilistic proof of Carne's bound. In a first time, only the Gaussian part of the bound will be considered. The fundamental estimate is the

Theoreme 3.1. *Let \mathbb{P} be a Markov chain as described in § 1.1; let $t \geq 2$; let $x \neq y \in V$; then we have:*

$$p^t(x,y) \leq \left(\frac{\mu(y)}{\mu(x)}\right)^{1/2} \exp\left(-\frac{(d(x,y)-1)^2}{2(t-1)}\right). \quad (3.1)$$

The reader who wants a more pleasant formula can easily check the

Corollary 3.2. *For $t \geq 1$ and $x, y \in V$:*

$$p^t(x,y) \leq \sqrt{e} \left(\frac{\mu(y)}{\mu(x)}\right)^{1/2} \exp\left(-\frac{d(x,y)^2}{2t}\right). \quad (3.2)$$

Remark 3.3. The first factor in the bound (3.2) is slightly better than that of (1.4), but actually one could replace the 2 by a \sqrt{e} in the proof of § 2 by refining majoration (2.10).

Proof. Denote $d = d(x,y)$. The first thing to notice is that, by reversibility of the chain, one has

$$p^t(y,x) = \frac{\mu(x)}{\mu(y)} p^t(x,y); \quad (3.3)$$

so to prove (3.1) it suffices to show:

$$p^t(x,y) p^t(y,x) \leq \exp\left(-\frac{(d-1)^2}{t-1}\right). \quad (3.4)$$

Now, rather than reasoning on the graph, which is a “complicated” object, we shall introduce a function $\xi : V \rightarrow \mathbb{R}$ which measures how much the particle X is closer to x or to y . ξ must satisfy:

Hypothesis 3.4.

- $\xi(x) = 0; \xi(y) = d;$
- ξ is 1-Lipschitz, i.e. for $z \sim v$ we have $|\xi(v) - \xi(z)| \leq 1.$

Such a map ξ always exists since, for instance, the map $d(x, \cdot)$ always satisfies assumption 3.4. Each point of V tends to make ξ increase or decrease, depending on the values of the transition kernel. Let us denote by $m(z)$ the expected value for the variation of ξ after the particle having visited z , i.e.

$$m(z) = \mathbb{E}_z [\xi(X_1)] - \xi(z). \quad (3.5)$$

Now, let $(M_u)_{u \geq 1}$ be the process defined by:

$$M_u = \xi(X_u) - \xi(X_1) - \sum_{s=1}^{u-1} m(X_s) = \sum_{s=1}^{u-1} (\xi(X_{s+1}) - \xi(X_s) - m(X_s)) ; \quad (3.6)$$

obviously M is a martingale starting at 0. Let us look at the chain starting at x . On the event $\{X_t = y\}$, one trivially has $\xi(X_t) - \xi(X_1) \geq d - 1$, hence

$$\mathbb{E}_x [M_t | X_t = y] \geq d - 1 - \mathbb{E}_x \left[\sum_{u=1}^{t-1} m(X_u) \middle| X_t = y \right]. \quad (3.7)$$

One may carry out the same reasoning starting at y , which gives:

$$\mathbb{E}_y [M_t | X_t = x] \leq -(d - 1) - \mathbb{E}_y \left[\sum_{u=1}^{t-1} m(X_u) \middle| X_t = x \right]. \quad (3.8)$$

What can we see? If the terms $\mathbb{E}_x [\sum_{u=1}^{t-1} m(X_u) | X_t = y]$, resp. $\mathbb{E}_y [\sum_{u=1}^{t-1} m(X_u) | X_t = x]$ were not present in (3.7) and (3.8), these formulae would reduce to $\mathbb{E}_x [M_t | X_t = y] \geq d - 1$, resp. $\mathbb{E}_y [M_t | X_t = x] \leq -(d - 1)$, so that we would observe a large deviation phenomenon on martingales, which would permit us to control respectively $p^t(x, y)$ and $p^t(y, x)$. Unfortunately, that phenomenon seems to be wiped out because of the terms $m(X_s)$. The cruxpoint then consists in noticing that, by reversibility, these $m(X_s)$ are the same for the “way out” as for the “way back”; subsequently, if the $m(X_s)$ tend to make the right hand side of (3.7) diminish (which would *damp* the large deviation phenomenon), then they tend to make the right hand side of (3.8) increase, which this time translates into a *strengthening* of the large deviation phenomenon! So, $p^t(x, y)$ and $p^t(y, x)$ cannot be large simultaneously, which will permit us to get (3.4) and conclude.

So, we consider X^x, X^y two independent chainswuth respective laws \mathbb{P}_x and \mathbb{P}_y ; let $\mathbb{P}_{x \otimes y}$ be their joint law. The respective realizations of $(M_u)_{u \geq 1}$ for the paths starting at x and at y are denoted by $(M_u^x)_{u \geq 1}$ and $(M_u^y)_{u \geq 1}$. We write then, by reversibility:

$$\forall u \in \{1, \dots, t-1\} \quad \mathbb{E}_x[m(X_u)|X_t = y] = \mathbb{E}_y[m(X_{t-u})|X_t = x], \quad (3.9)$$

hence by combining (3.7) and (3.8):

$$\mathbb{E}_{x \otimes y}[M_t^x - M_t^y | X_t^x = y \text{ et } X_t^y = x] \geq 2(d-1). \quad (3.10)$$

It remains to control the deviations of $M_t^x - M_t^y$. We remark that this random variable may be interpreted as the final value of a $2(t-1)$ steps martingale, whose steps satisfy the assumptions of the following lemma:

Lemma 3.5. *Let $(\mathcal{F}_t)_{t \in \mathbb{N}}$ be a filtration; let $(X_t)_{t \geq 1}$ be an adapted real-valued process with $\mathbb{E}[X_{t+1}|\mathcal{F}_t] = 0$ ⁽⁴⁾. We suppose that, for all $t \in \mathbb{N}$, $\text{Law}(X_{t+1}|\mathcal{F}_t)$ is supported by an interval of length 2 almost surely. Then, letting $u \geq 0$ be a fixed time, we have for all $\lambda \in \mathbb{R}$:*

$$\mathbb{E} \left[\exp \left(\lambda \sum_{t=1}^u X_t \right) \right] \leq \exp \left(u \frac{\lambda^2}{2} \right). \quad (3.11)$$

Proof. The proof relies on Hoeffding's inequality, whose statement is recalled here below:

Lemma 3.6 (Hoeffding). *Let X be a centered real-valued random variable, supported by an interval of length 2, then*

$$\forall \lambda \in \mathbb{R} \quad \mathbb{E} [e^{\lambda X}] \leq e^{\lambda^2/2}. \quad (3.12)$$

That point being taken for granted, we prove lemma 3.5 by induction on u :

- For $u = 0$ the result is trivial.
- Let $u \geq 1$; suppose the result to be true for $u-1$. Let $\lambda \in \mathbb{R}$; we write:

$$\begin{aligned} \mathbb{E} \left[\exp \left(\lambda \sum_{t=1}^u X_t \right) \right] &= \mathbb{E} \left[\exp \left(\lambda \sum_{t=1}^{u-1} X_t \right) \mathbb{E} [e^{\lambda X_u} | \mathcal{F}_{u-1}] \right] \\ &\leq \underbrace{\mathbb{E} \left[\exp \left(\lambda \sum_{t=1}^{u-1} X_t \right) \right]}_{\leq e^{(u-1)\lambda^2/2} \text{ by induction}} \cdot \underbrace{\|\mathbb{E} [e^{\lambda X_u} | \mathcal{F}_{u-1}]\|_\infty}_{\leq e^{\lambda^2/2} \text{ by (3.12)}} \leq e^{u\lambda^2/2}. \end{aligned} \quad (3.13)$$

□

⁽⁴⁾In other words, the X_t 's are the increments of a martingale.

To conclude, it just remains to us to prove the following measure concentration lemma:

Lemma 3.7. *Let X be a centered real-valued random variable satisfying for some $k > 0$:*

$$\forall \lambda \in \mathbb{R} \quad \mathbb{E}[e^{\lambda X}] \leq e^{k\lambda^2/2}. \quad (3.14)$$

We suppose that \mathcal{A} is an event such that

$$\mathbb{E}[X|\mathcal{A}] \geq C \quad (3.15)$$

for some $C \geq 0$, then

$$\mathbb{P}(\mathcal{A}) \leq \exp\left(-\frac{C^2}{2k}\right). \quad (3.16)$$

Proof. To lighten notations, let us denote $p = \mathbb{P}(\mathcal{A})$. Let us fix $\lambda > 0$, then we have:

$$\mathbb{E}[e^{\lambda X}|\mathcal{A}] = \frac{\mathbb{E}[\mathbb{1}_{\mathcal{A}}e^{\lambda X}]}{\mathbb{P}(\mathcal{A})} \stackrel{(3.14)}{\leq} \frac{1}{p}e^{k\lambda^2/2}. \quad (3.17)$$

It follows, by Jensen's inequality, that

$$\mathbb{E}[X|\mathcal{A}] \leq \frac{1}{\lambda} \ln\left(\frac{1}{p}e^{k\lambda^2/2}\right), \quad (3.18)$$

hence by assumption (3.15):

$$\frac{1}{\lambda} \ln\left(\frac{1}{p}e^{k\lambda^2/2}\right) \geq C, \quad (3.19)$$

whence finally:

$$p \leq e^{-C\lambda + k\lambda^2/2}. \quad (3.20)$$

Then it suffices to take $\lambda = C/k$ to get the announced result. \square

Now we conclude the proof of Theorem 3.1. Lemma 3.5 permits us to control the Laplace transform of $M_t^x - M_t^y$ under $\mathbb{P}_{x \otimes y}$:

$$\forall \lambda \in \mathbb{R} \quad \mathbb{E}_{x \otimes y} \left[e^{\lambda(M_t^x - M_t^y)} \right] \leq e^{(t-1)\lambda^2}, \quad (3.21)$$

and the formula (3.10) then gives, *via* Lemma 3.7:

$$\mathbb{P}_{x \otimes y}(X_t^x = y \text{ et } X_t^y = x) \leq \exp\left(-\frac{(d-1)^2}{t-1}\right), \quad (3.22)$$

i.e. (3.4). \square

4 Generalization to a larger class of distances

4.1 Statement of the generalized theorem

Now we will show that the reasoning made above can in fact adapt to deal with a whole class of distances. So let us consider a new distance on V , which we will also call d —to avoid confusions, the graph distance that we had defined by (1.3) will be denoted d^G from now on. We have:

Theorem 4.1. *Suppose that d is built so that, if $\xi : V \rightarrow \mathbb{R}$ is any 1-Lipschitz function with respect to d , one has for all $x \in V$:*

- $$|\mathbb{E}_x [\xi(X_1)] - \xi(x)| \leq B; \quad (4.1)$$

- $$\forall \lambda \geq 0 \quad \mathbb{E}_x [e^{\lambda(\xi(X_1) - \mathbb{E}_x[\xi(X_1)])}] \leq e^{A\lambda^2}, \quad (4.2)$$

for some constants A and B independent of ξ . Then, for all $x, y \in V$, one has:

$$\begin{aligned} p^t(x, y) &\leq \left(\frac{\mu(y)}{\mu(x)} \right)^{1/2} \exp \left(-\frac{(d(x, y) - B)_+^2}{4At} \right) \\ &\leq e^{B/2A} \left(\frac{\mu(y)}{\mu(x)} \right)^{1/2} \exp \left(-\frac{d(x, y)^2}{4At} \right). \end{aligned} \quad (4.3)$$

Remark 4.2. Let us immediately point out that, in the case when the distance is d^G , it is impossible to get anything better than $A = 1/2$ and $B = 1$. Subsequently, the result will be worsened down by a \sqrt{e} factor compared to (3.2)—which remains negligible compared to the exponential part of the bound—: as we will see it later, it is due to the difference in treating the first steps, since the specific argument that we have used for d^G in the proof of Theorem 3.1 may not generalize.

Proof. We follow again the proof of theorem 3.1: denote $d = d(x, y)$, let ξ satisfy assumption 3.4, define m by (3.5) and let $\mathbb{P}_{x \otimes y}$ be the joint law of two independent chains of respective laws \mathbb{P}_x and \mathbb{P}_y ; we want to bound $\mathbb{P}_{x \otimes y}[X_t^x = y \text{ et } X_t^y = x]$ above to conclude by formula (3.3).

The first difference lies in the definition of M : now, the martingale starts at time 0. So it is defined by:

$$M_u = \xi(X_u) - \xi(X_0) - \sum_{s=0}^{u-1} m(X_s) = \sum_{s=0}^{u-1} (\xi(X_{s+1}) - \xi(X_s) - m(X_s)). \quad (4.4)$$

We get then:

$$\begin{cases} \mathbb{E}_x [M_t | X_t = y] = d - \mathbb{E}_x \left[\sum_{u=0}^{t-1} m(X_u) \middle| X_t = y \right], \\ \mathbb{E}_y [M_t | X_t = x] = -d - \mathbb{E}_y \left[\sum_{u=0}^{t-1} m(X_u) \middle| X_t = x \right] \end{cases} \quad (4.5)$$

When we want to combine these two formulae as we did in (3.10), we observe that all the terms $\mathbb{E}[m(X_u)]$ will cancel pair by pair, *excepted* the terms corresponding to the first steps, i.e. to $u = 0$ in the two respective formulae. Fortunately, we know perfectly what these terms are, since under \mathbb{P}_x , we have $X_0 = x$ a.s. (by definition!), resp. $X_0 = y$ a.s. under \mathbb{P}_y . Then we get:

$$\mathbb{E}_{x \otimes y} [M_t^x - M_t^y | X_t^x = y \text{ et } X_t^y = x] = 2d - m(x) + m(y) \geq 2d - 2B. \quad (4.6)$$

Taking into account assumption (4.2) —which plays here the role played before by Hoeffding’s inequality—, we copy off the proof of lemma 3.5 to get:

$$\forall \lambda \geq 0 \quad \mathbb{E}_{x \otimes y} \left[e^{\lambda(M_t^x + M_t^y)} \right] \leq e^{2At\lambda^2}, \quad (4.7)$$

and it just remains to conclude by lemma 3.7⁽⁵⁾. □

4.2 More flexible distances

Now we will show how one may build distances satisfying theorem 4.1, such that the metric structure of V continuously depends on the transition kernel. The method developed here below is certainly neither the best nor the most elegant, but it has for advantage to be of relative pedagogical simplicity.

We keep on the principle of putting a length to each edge, but that time all the edges will not have the same size: indeed we will put a larger length to the edges that are the most difficult to visit, in order to ensure that the metric structure of the graph will not be too much disrupted when we add a very “unlikely” edge.

Let $\alpha > 0$ be an arbitrary parameter. To each couple $(x, y) \in V \times V$, we associate a length $\ell(x, y)$ such that:

$$\forall a \geq 0 \quad \mathbb{P}_x (\ell(x, X_1) \geq 1 + a) \leq e^{-a^2/\alpha}, \quad (4.8)$$

and we define the length of the edge $[xy]$ by $|[xy]| = \min\{\ell(x, y), \ell(y, x)\}$. As soon, for all 1-Lipschitz function ξ on V , we have:

Claim 4.3.

$$\forall x \in V \quad \forall a \geq 0 \quad \mathbb{P}_x (|\xi(X_1) - \xi(x)| \geq 1 + a) \leq e^{-a^2/\alpha}. \quad (4.9)$$

⁽⁵⁾If it occurs that $B \geq d$, lemma 3.7 cannot be applied and we may only bound over $p^t(x, y)p^t(y, x)$ by 1; it is the cause that explains the positive part appearing in (4.3).

Now we give a formula for $\ell(x, y)$ satisfying (4.8). First we define what here we will call the β -entropy of a probability law:

Definition 4.4. Let $\beta \in]0, 1]$; let p be a probability measure on a discrete space \mathcal{X} . We call β -entropy of p the (possibly infinite) number:

$$H_\beta(p) = \sum_{x \in \mathcal{X}} p(x)^{1-\beta} \quad (4.10)$$

A transition kernel p on V being given, we will also denote, for $x \in V$, $H_\beta(x) = H_\beta(p(x, \cdot))$.

The β -entropy permits us to control the probability that the observed event is rare:

Claim 4.5. Let $\beta \in]0, 1]$; let \mathbb{P} be a probability law on a discrete state space \mathcal{X} . We suppose $H_\beta(p) < \infty$. Then, for all $\varpi \in]0, 1]$, one has:

$$\mathbb{P}(p(x) \leq \varpi) \leq H_\beta(p) \varpi^\beta. \quad (4.11)$$

Proof. Use the identity $H_\beta(p) = \mathbb{E}[p(x)^{-\beta}]$ and the fact that the map $\varpi \mapsto \varpi^{-\beta}$ is decreasing, then apply Markov's inequality. \square

So, one may choose the following expression for $\ell(x, y)$ to satisfy (4.8), where we set that, for $a < 0$, $a^{1/2} = 0$:

$$\ell(x, y) = 1 + \sqrt{\alpha} \left(\beta \ln(p(x, y)^{-1}) - \ln H_\beta(x) \right)^{1/2}. \quad (4.12)$$

Now, we want to show that (4.9) permits us to get (4.1) and (4.2) indeed. Let us begin with an easy observation:

Claim 4.6. Let ξ be a 1-Lipchitz function on V for a distance d built as above. Then there exists a random variable Y whose repartition map satisfies $\forall a \geq 0 \mathbb{P}(Y \geq 1 + a) = e^{-a^2/\alpha}$, i.e. Y has a law with density

$$d\mathbb{P}(Y = y) = \mathbb{1}_{y \geq 1} \frac{2(y-1)}{\alpha} e^{-(y-1)^2/\alpha} dy, \quad (4.13)$$

such that one has:

$$\mathbb{P}_x\text{-p.s.} \quad |\xi(X_1) - \xi(x)| \leq Y. \quad (4.14)$$

So, we easily find that we can take $B = \mathbb{E}[Y] = 1 + \frac{\sqrt{\pi}}{2}\alpha$ into (4.1).

To get a formula for A in (4.2), things are a bit more complicated. The tool which we will use is the

⁽⁶⁾For $\beta = 1$ we set by continuity $H_\beta(p) = \#\{x \in \mathcal{X}; p(x) > 0\}$.

Lemma 4.7. *Let Y be a positive random variable whose Laplace transform $\widehat{Y}(\lambda) = \mathbb{E}[e^{\lambda Y}]$ is supposed to be finite for all $\lambda \geq 0$, and let us denote $\bar{Y} = \mathbb{E}[Y]$. Let X be a real-valued random variable satisfying $|X| \leq Y$ a.s.; let us denote $\bar{X} = \mathbb{E}[X]$ and $\widetilde{X} = X - \mathbb{E}[X]$. Then the Laplace transform of \widetilde{X} satisfies:*

$$\forall \lambda \in \mathbb{R} \quad \widehat{\widetilde{X}}(\lambda) \leq e^{|\lambda|\bar{Y}} \widehat{Y}(|\lambda|) - 2|\lambda|\bar{Y}. \quad (4.15)$$

Remark 4.8. The bound (4.15) is of quite poor quality close to 0, as it may be particularly striking in the case $Y \equiv 1$, where we get the bound $e^{2|\lambda|} - 2|\lambda|$, while we know (Hoeffding's lemma, cf. lemma 3.6) that $e^{\lambda^2/2}$ would work. In fact, we can compute that in a neighborhood of 0, the right hand side of (4.15) takes the form:

$$1 + (\mathbb{E}[Y^2] + 3\mathbb{E}[Y]) \frac{\lambda^2}{2} + o(\lambda^2), \quad (4.16)$$

while a variance calculation proves that one has in fact:

$$\widehat{\widetilde{X}}(\lambda) \leq 1 + \mathbb{E}[Y^2] \frac{\lambda^2}{2} + o(\lambda^2). \quad (4.17)$$

Proof. We may restrict ourselves to the case $\lambda \geq 0$, the case $\lambda \leq 0$ being then treated by using the result for $-X$.

We write that

$$\widehat{\widetilde{X}}(\lambda) = \mathbb{E} \left[e^{\lambda \widetilde{X}} - \lambda \widetilde{X} \right]. \quad (4.18)$$

But, since $|X| \leq Y$, we have $|\bar{X}| \leq \bar{Y}$, hence $\widetilde{X} \in [-Y - \bar{Y}, Y + \bar{Y}]$. And since, on an interval of the form $[-a, a]$, the map $x \mapsto e^{\lambda x} - \lambda x$ takes its maximum at a , it follows that

$$\widehat{\widetilde{X}}(\lambda) \leq \mathbb{E} \left[e^{\lambda(Y + \bar{Y})} - \lambda(Y + \bar{Y}) \right] = e^{\lambda \bar{Y}} \widehat{Y}(\lambda) - 2\lambda \bar{Y}. \quad (4.19)$$

□

Now we have the following control on the Laplace transform of the random variable Y defined by (4.13):

Proposition 4.9. *If Y is a random variable whose law is given by (4.13), then for $\lambda \geq 0$ one has:*

$$\mathbb{E} \left[e^{\lambda Y} \right] \leq \left(1 + \frac{\sqrt{\pi}\alpha}{2}\lambda + \frac{\alpha^2}{4}\lambda^2 \right) e^{\lambda + \alpha^2 \lambda^2 / 4}. \quad (4.20)$$

Moreover, that bound is sharp close to 0, by which we mean that, if we rewrite (4.20) under the form $\mathbb{E}[e^{\lambda Y}] \leq f(\lambda)$, we have $\mathbb{E}[Y] = \frac{d}{d\lambda} \Big|_{\lambda=0} \mathbb{E}[e^{\lambda Y}] = f'(0) = 1 + \frac{\sqrt{\pi}}{2}\alpha$.

Proof. We begin with noticing that we can write $Y = 1 + \alpha Z$, where Z is a random variable with law:

$$d\mathbb{P}(Z = z) = \mathbb{1}_{z \geq 0} 2z e^{-z^2} dz. \quad (4.21)$$

Then it suffices to prove that one has

$$\forall \lambda \geq 0 \quad \mathbb{E} [e^{\lambda Z}] \leq \left(1 + \frac{\sqrt{\pi}}{2} \lambda + \frac{\lambda^2}{4}\right) e^{\lambda^2/4}. \quad (4.22)$$

To do that, we write:

$$\mathbb{E} [e^{\lambda Z}] = e^{\lambda^2/4} \int_0^\infty 2z e^{-(z-\lambda/2)^2} dz \stackrel{t=z-\lambda/2}{=} e^{\lambda^2/4} \underbrace{\int_{-\lambda/2}^\infty (2t + \lambda) e^{-t^2} dt}_{I(\lambda)}, \quad (4.23)$$

where $I(0) = 0$ and, by the theorem of derivation under the integral:

$$I'(\lambda) = \int_{-\lambda/2}^\infty e^{-t^2} dt \leq \frac{\sqrt{\pi}}{2} + \frac{\lambda}{2}, \quad (4.24)$$

whence $I(\lambda) \leq 1 + \frac{\sqrt{\pi}}{2} \lambda + \frac{\lambda^2}{4}$, and (4.20). \square

From that we deduce the existence of a suitable value for A :

Proposition 4.10. *For all $\alpha > 0$, there exists a constant $A(\alpha) < \infty$ such that, if d satisfies condition (4.9) for the value α , then (4.2) is satisfied for the value $A(\alpha)$.*

Proof. For Y with law (4.13), denoting $\bar{Y} = \mathbb{E}[Y] = 1 + \frac{\sqrt{\pi}}{2} \alpha$, proposition 4.9 shows that $\ln(e^{\lambda \bar{Y}} \mathbb{E}[e^{\lambda Y}] - 2\lambda \bar{Y})/\lambda^2$ is bounded for $\lambda \xrightarrow{\geq} 0$ and $\lambda \rightarrow \infty$. By continuity, this function is thus bounded on the whole half-line $[0, +\infty)$. Lemma 4.7 then gives us the existence of A . \square

Remark 4.11. We have not found any simple expression bounding over $A(\alpha)$, but, for a given value of α , it is easy to compute numerically the maximum of the map $\lambda \mapsto \ln(e^{\lambda \bar{Y}} \mathbb{E}[e^{\lambda Y}] - 2\lambda \bar{Y})/\lambda^2$, which gives the optimal value of A according to the the frame of our proof.

4.3 A concrete example

We are going to illustrate the preceding subsection by showing how our generalization permits us to give some results in cases when the usual Carne bound is irrelevant.

Here we consider V to be the set of vertices of an infinite 3-tree. The tree distance is denoted by d^A , which will not be the same as the graph distance

d^G . Let L be an integer devised to become large and $\varepsilon > 0$ a real number devised to become small. We consider as the Markov chain on V the process which, from the point x , jumps on each neighbor of x with probability $(1-\varepsilon)/3$, and, with probability ε , chooses uniformly an arrival point in the (closed) ball centered on x with radius L . As one may check immediately, that chain is reversible and its invariant measure is the counting measure. Let us sum up: our process looks much like the simple random walk on a 3-tree, but sometimes it may occur an “accident” which makes the mobile jump by roughly L units. We would like to say that, even if L is large, it suffices for ε to be small enough to get an exponential bound where L does not appear.

If we naively apply formula (1.4) to this transition kernel, we will not get anything interesting: indeed, as small as ε might be, the graph distance is the same:

$$\forall x, y \in G \quad d^G(x, y) = \left\lceil \frac{d^A(x, y)}{L} \right\rceil. \quad (4.25)$$

Then, bounding below d^G by d^A/L , and merely bounding $|P|$ by $1^{(7)}$, the formula that we get by Carne’s bound is:

$$p^t(x, y) \leq 2 \exp\left(-\frac{d^A(x, y)}{2tL^2}\right). \quad (4.26)$$

Concretely, if $L = 17$ and $\varepsilon = 1/2^{2^{30}}$, it will give a bound which is strongly worse than in the case $\varepsilon = 0$, while it is obvious that the influence of large jumps should be nearly zero: the bound (4.26) thus must be improvable!

So we will apply the techniques exposed in §4.2. Here we have chosen arbitrarily $\alpha = 1$ and $\beta = 1/2$. We suppose that L is large enough; in fact our computations will be valid as soon as $L \geq 2$. Let us denote by $N = 3 \cdot 2^L - 2$ the cardinality of a ball of radius L . We have:

$$H_{1/2}(x) = 3 \left(\frac{1-\varepsilon}{3} + \frac{\varepsilon}{N} \right)^{1/2} + (N-3) \left(\frac{\varepsilon}{N} \right)^{1/2} \leq \sqrt{3} + \sqrt{N}\varepsilon^{1/2}, \quad (4.27)$$

hence $H_{1/2}(x) \leq 2$ for $\varepsilon \leq 1/(42 \cdot 2^L)$. So, if we bound over $H_{1/2}(x)$ by 2, we get that for $2 \leq d^A(x, y) \leq L$, one has:

$$\ell(x, y) \geq 1 + (\ln(N/\varepsilon)/2 - \ln 2)^{1/2}, \quad (4.28)$$

in particular $\ell(x, y) \geq 1 + \sqrt{\ln(\varepsilon^{-1})}/2$.

Our observation is that, if ε is small enough, d coincides with d^A : indeed one has $\ell(x, y) = 1$ for $x \sim y$, $\ell(x, y) = \infty$ if $d^A(x, y) > L$, and $\ell(x, y) \geq L$ for $2 \leq d^A(x, y) \leq L$, as soon as:

$$\varepsilon \leq e^{-2L^2} \quad (4.29)$$

⁽⁷⁾Actually, here we could do better, but in this subsection we are only interested in the Gaussian part of the bound.

—which is however a quite strong condition. (Note that for L large enough, condition (4.29) implies that $\varepsilon \leq 1/(42 \cdot 2^L)$ indeed.)

Numerical computations for $\alpha = 1$ give $A = 8.097, \dots$, resp. $B = 1.886, \dots$; so we have for $\varepsilon \leq e^{-2L^2}$, worsening slightly numerical values:

$$p^t(x, y) \leq \frac{9}{8} \exp\left(-\frac{d^A(x, y)^2}{33t}\right). \quad (4.30)$$

Although that bound undoubtedly improves Carne’s bound (4.26) in “extreme” cases like that mentioned above, and even if it is certainly possible to get some better results by a more subtle choice of α and β , that bound is rather disappointing. . . Anyway, the method described in the § 4.2 remains of theoretical interest; moreover there may be better ways than it to use theorem 4.1.

5 The flight factor

5.1 Foreword

The Gaussian bound (3.2) has got a disadvantage with regard to Carne’s bound (1.4): in the case when $|P| < 1$, it does not show the exponential decreasing of $p^t(x, y)$ with respect to t . In fact, Lemma 2.3 implies $p^t(x, y)p^t(y, x) \leq p^{2t}(x, x) \leq |P|^{2t}$, whence

$$p^t(x, y) \leq \left(\frac{\mu(y)}{\mu(x)}\right)^{1/2} |P|^t, \quad (5.1)$$

but that is not enough to get again (1.4). That section precisely aims at doing this. Here we will exclusively focus on the case when $d = d^G$, cf. remark 5.8 *infra*.

Let $x, y \in V$. For u a time devised to go to infinity, denote by

$$\mathcal{R}_u = \{\exists s \geq u; X_s = x\} \quad (5.2)$$

the event which tells that the particle comes back at x at least once after time u . The strategy of our proof then consists in looking at our Markov chain *conditioned to the event* \mathcal{R}_u . Why this? Well, the fact that $|P| < 1$ expresses a possibility for the particle to “flee to infinity”, which is responsible for the exponential decay with respect to t of the quantity $p^t(x, y)p^t(y, x)$ introduced in (3.4), which measures the probability for the particle, starting at x , of going to y at time t and then coming back to x at time $2t$. Conditioning with respect to \mathcal{R}_u then aims, in a way, at preventing the particle from going to infinity, which will give us a Markov chain for which $|P| = 1$, where the bound (3.2) will be relevant. Then it will remain to show that this conditioning selects

sufficiently well the cases when the particle makes a return trip to get back a factor $|P|^t$ in (3.2).

Our proof shall use a kind of density argument: in a first time, we will have to add some more assumptions on our Markov chain to carry out our reasoning, then in a second time we will prove that we can get rid of these extra assumptions by slightly perturbing the original Markov chain.

5.2 Proof under extra assumptions

We will use the following notation:

Definition 5.1. We denote by τ_x the hitting time of x by a walk on V , i.e. $\tau_x = \inf\{t \geq 0; X_t = x\}$. For all $z \in V$, we denote:

$$R(z) = \mathbb{E}_z [\mathbb{1}_{\tau_x < \infty} |P|^{-\tau_x}]. \quad (5.3)$$

In this first part of the proof, we add to the assumptions of § 1.1 the following conditions:

Hypothesis 5.2.

1. V is finite.
2. There exists a cemetery point $\partial \in V$ such that $p(\partial, \cdot) = \delta_\partial$. We will denote \tilde{V} for $V \setminus \{\partial\}$ ⁽⁸⁾.
3. The chain \mathbb{P} is aperiodic on \tilde{V} .

Assumption 5.2 permits to obtain sharp results about the recurrence behaviour of the chain:

Lemma 5.3. Suppose we have an irreducible and reversible Markov chain satisfying Assumption 5.2. Then:

1. There exist two constants $0 < c_1 \leq c_2 < \infty$ such that

$$\forall t \geq 0 \quad c_1 |P|^t \leq \mathbb{P}_x(\mathcal{R}_T) \leq c_2 |P|^t. \quad (5.4)$$

2. $\mathbb{P}_x(\mathcal{R}_{t+1})/\mathbb{P}_x(\mathcal{R}_t) \xrightarrow[t \rightarrow \infty]{} |P|$.

3. For all $z \in \tilde{V}$, $R(z) < \infty$.

⁽⁸⁾In this context, the chain will just be required to be irreducible and reversible on \tilde{V} ; moreover the definition of $|P|$ will be that given by the formula (2.2) of lemma 2.3, for arbitrary $x \in \tilde{V}$.

The proof of this lemma, which is rather technical, is postponed to the Appendix.

Now we are armed to prove Carne's bound for a Markov chain satisfying Assumption 5.2. Let us fix $t \geq 0$; we have the key proposition:

Proposition 5.4. *The law of $(X_s)_{0 \leq s \leq 2t}$ under $\mathbb{P}_x(\cdot | \mathcal{R}_u)$ converges when $u \rightarrow \infty$ (for the total variation norm on $V^{\{0, \dots, 2t\}}$) to the law \mathbb{P}'_x of the Markov chain described here below:*

Definition 5.5. \mathbb{P}' is the Markov chain on \tilde{V} with transition probabilities:

$$\forall z, v \in \tilde{V} \quad p'(z, v) = \frac{R(v)p(z, v)}{\sum_{w \sim z} R(w)p(z, w)} \quad (9). \quad (5.5)$$

Proof. It is true in a general framework that $\mathbb{P}(\cdot | \mathcal{R}_u)$ is a time-inhomogeneous Markov chain with:

$$\mathbb{P}(X_{s+1} = v | X_s = z \text{ and } \mathcal{R}_u) = \frac{S_{s+1, u}(v)p(z, v)}{\sum_{w \sim z} S_{s+1, u}(w)p(z, w)}, \quad (5.6)$$

where we let $S_{s+1, u}(z) = \mathbb{P}(\mathcal{R}_u | X_{s+1} = z)$. Our attack will consist in proving that, for all $s \in \{0, \dots, 2t - 1\}$, for all $z \in V$, we have $S_{s+1, u}(z) / \mathbb{P}_x(\mathcal{R}_{u-(s+1)}) \xrightarrow{u \rightarrow \infty} R(z)$.

To begin with, let us notice that $S_{s+1, u}(z) = \mathbb{P}_z(\mathcal{R}_{u-(s+1)})$, which we shall denote by $S_{u-(s+1)}(z)$; what we want to prove then can be written:

$$\forall z \in \tilde{V} \quad \frac{S_u(z)}{\mathbb{P}_x(\mathcal{R}_u)} \xrightarrow{u \rightarrow \infty} R(z). \quad (5.7)$$

The idea consists in splitting the probability space according to the value of τ_x , thanks to the strong Markov property: we write

$$S_u(z) = \sum_{s=0}^u \mathbb{P}_z(\tau_x = s) \mathbb{P}_x(\mathcal{R}_{u-s}) + \sum_{s \geq u+1} \mathbb{P}_z(\tau_x = s), \quad (5.8)$$

whence

$$\frac{S_u(z)}{\mathbb{P}_x(\mathcal{R}_u)} = \sum_{s=0}^u \mathbb{P}_z(\tau_x = s) \frac{\mathbb{P}_x(\mathcal{R}_{u-s})}{\mathbb{P}_x(\mathcal{R}_u)} + \sum_{s \geq u+1} \frac{\mathbb{P}_z(\tau_x = s)}{\mathbb{P}_x(\mathcal{R}_u)}. \quad (5.9)$$

Let us fix an arbitrarily small $\varepsilon > 0$. Since $\sum_s \mathbb{P}_z(\tau_x = s) |P|^{-s}$ converges (cf. Lemma 5.3-3), we may introduce a time u_0 for which one has $\sum_{s > u_0} \mathbb{P}_z(\tau_x =$

⁽⁹⁾That defines a Markov chain on \tilde{V} indeed because $R(\partial) = 0$.

$s) \leq \varepsilon$. Then, for $u \geq u_0$ one has on the one hand:

$$\sum_{s=u_0+1}^u \mathbb{P}_z(\tau_x = s) \frac{\mathbb{P}_x(\mathcal{R}_{u-s})}{\mathbb{P}_x(\mathcal{R}_u)} + \sum_{s>u} \frac{\mathbb{P}_z(\tau_x = s)}{\mathbb{P}_x(\mathcal{R}_u)} \stackrel{\text{lemma 5.3-1}}{\leq} \frac{c_2 \vee 1}{c_1} \sum_{s>u_0} \mathbb{P}_z(\tau_x = s) |P|^{-s} = \frac{c_2 \vee 1}{c_1} \varepsilon, \quad (5.10)$$

on the other hand:

$$\sum_{s=0}^{u_0} \mathbb{P}_z(\tau_x = s) \frac{\mathbb{P}_x(\mathcal{R}_{u-s})}{\mathbb{P}_x(\mathcal{R}_u)} \stackrel{\text{lemma 5.3-2}}{\underset{u \rightarrow \infty}{\rightarrow}} \sum_{s=0}^{u_0} \mathbb{P}_z(\tau_x = s) |P|^{-s}. \quad (5.11)$$

It follows that

$$\limsup_{u \rightarrow \infty} \left| R(z) - \frac{S_u(v)}{\mathbb{P}_x(\mathcal{R}_u)} \right| \leq \left(1 + \frac{c_2 \vee 1}{c_1} \right) \varepsilon, \quad (5.12)$$

hence (5.7) by letting $\varepsilon \rightarrow 0$. \square

Now we want to look at the chain \mathbb{P}' . First, \mathbb{P}' is clearly irreducible. Then, one has:

Proposition 5.6. *The chain \mathbb{P}' is reversible, and its invariant measure is:*

$$\forall z \in \tilde{V} \quad \mu'(z) = \begin{cases} R(z)^2 \mu(z) & \text{if } z \neq x; \\ R(x) R^+(x) \mu(x) & \text{if } z = x, \end{cases} \quad (10) \quad (5.13)$$

where $R^+(x)$ is defined by:

Definition 5.7. *We denote by τ_x^+ the return time to x , i.e.:*

$$\tau_x^+ = \inf\{s \geq 1; X_s = x\}. \quad (5.14)$$

Then $R^+(x)$ is defined by:

$$R^+(x) = \mathbb{E}_x \left[\mathbb{1}_{\tau_x^+ < \infty} |P|^{-\tau_x^+} \right]. \quad (5.15)$$

Proof. Let $z, v \in \tilde{V}$ with $z \neq x$. We can lighten the expression of $p'(z, v)$, since one has by Markov's property:

$$R(z) = \sum_{w \sim z} p(z, w) \mathbb{E}_w \left[\mathbb{1}_{\tau_x < \infty} |P|^{-(\tau_x+1)} \right] = |P|^{-1} \sum_{w \sim z} p(z, w) R(w), \quad (5.16)$$

thus (5.5) can be rewritten as:

$$p'(z, v) = \frac{p(z, v) R(v)}{|P| R(z)}. \quad (5.17)$$

⁽¹⁰⁾The careful reader may have noticed that $R(x) = 1$; we let that factor appear for ease of understanding.

In the case when $z = x$, the same argument leads to

$$p^*(x, v) = \frac{p(x, v)R(v)}{|P|R^+(x)}. \quad (5.18)$$

So, it just remains to use (5.17), (5.18) and the reversibility of μ under \mathbb{P} to get the reversibility of μ' under \mathbb{P}' . \square

Now we are completely ready to end our proof. We observe that, by Markov's property, $p^t(x, y)p^t(y, x) = \mathbb{P}_x(X_t = y \text{ and } X_{2t} = x)$, which we will denote by $\mathbb{P}_x(\mathcal{A})$. For $u \geq 2t$, Bayes' formula gives:

$$\frac{\mathbb{P}_x(\mathcal{A})}{\mathbb{P}_x(\mathcal{A}|\mathcal{R}_u)} = \frac{\mathbb{P}_x(\mathcal{R}_u)}{\mathbb{P}_x(\mathcal{R}_u|\mathcal{A})} \stackrel{\text{Markov}}{=} \frac{\mathbb{P}_x(\mathcal{R}_u)}{\mathbb{P}_x(\mathcal{R}_{u-2t})}, \quad (5.19)$$

hence, letting u go to infinity:

$$\frac{\mathbb{P}_x(\mathcal{A})}{\mathbb{P}'_x(\mathcal{A})} = |P|^{2t}. \quad (5.20)$$

Now, the Markov chain \mathbb{P}' satisfies the assumptions of theorem 3.1, hence $\mathbb{P}'_x(\mathcal{A}) \leq e \cdot \exp(-d(x, y)^2/t)$, so $\mathbb{P}_x(\mathcal{A}) \leq e|P|^{2t} \exp(-d(x, y)^2/t)$, and finally we get the desired formula:

$$p^t(x, y) \leq \sqrt{e} \left(\frac{\mu(x)}{\mu(y)} \right)^{1/2} |P|^t \exp\left(-\frac{d(x, y)^2}{2t}\right). \quad (5.21)$$

Remark 5.8. You may notice that the reasoning carried out above cannot apply to other distances than d^G : indeed, the distance which appears in bound (5.21) in fact comes as the distance *associated to the process* \mathbb{P}' . When one works with the graph distance, that distance is actually the same for \mathbb{P} and for \mathbb{P}' , but this is no more true if d depends in a more subtle way on the transition kernel.

5.3 The density argument

Now, we want to get rid of Assumption 5.2. We will proceed in two steps: first we will just relax assumption 5.2-3, then we will deal with the general case.

Relaxing the aperiodicity condition

We consider a finite set V with a transition law $(p(x, y))_{x, y \in V}$, such that there exists a cemetery point $\partial \in V$ satisfying assumption 5.2-2. We suppose that the Markov chain defined by p is irreducible and reversible on $\tilde{V} = V \setminus \{\partial\}$, with a reversible measure μ . We denote by n the cardinality of \tilde{V} , and by M the matrix $((p(y, x)))_{x, y \in \tilde{V}}$. We had decided, cf. footnote (8), to define $|P|$ through formula (2.2). A more algebraic characterization is given by:

Lemma 5.9. $|P|$ is the spectral radius of M .

Proof. See [3], chap. 5-2. □

Remark 5.10. You may remark that lemma 5.9 holds even for a non-reversible chain, though we will not use that fact here.

For $\varepsilon \in [0, 1[$, let p_ε be the transition kernel defined by:

$$\forall x, y \in V \quad p_\varepsilon(x, y) = \begin{cases} p(x, x) + \varepsilon(1 - p(x, x)) & \text{if } y = x; \\ (1 - \varepsilon)p(x, y) & \text{if } y \neq x. \end{cases} \quad (5.22)$$

The Markov chain \mathbb{P}^ε generated by p_ε is then an irreducible reversible chain whose graph and whose reversible measure are the same as for $p = p_0$, and which satisfies *completely* assumption 5.2 as soon as $\varepsilon > 0$. Thus, we have for $x, y \in \tilde{V}$, $t > 0$ and $\varepsilon > 0$:

$$p_\varepsilon^t(x, y) \leq \sqrt{\varepsilon} \left(\frac{\mu(y)}{\mu(x)} \right)^{1/2} |P_\varepsilon|^t \exp\left(-\frac{d(x, y)^2}{2t}\right). \quad (5.23)$$

To conclude, we just have to notice that $p_\varepsilon^t(x, y)$, resp. P_ε , are functions of ε continuous at 0. Indeed, the finite-sized matrix M_ε varies continuously with ε , thus its spectral radius $|P_\varepsilon|$ also varies continuously, as well as $p_\varepsilon^t(x, y)$ which is the coefficient number (y, x) of M_ε^t .

Infinite graphs

Now we turn to the general case, i.e. we consider a chain that merely satisfies the assumptions of § 1.1. Let us give a mark $\nu(z) > 0$ to each vertice z of V , in such a way that for all $\varepsilon > 0$, $\#\{z \in V; \nu(z) > \varepsilon\}$ is finite.

Let us fix $x, y \in V$, and let us take $\varepsilon > 0$ arbitrarily small (we shall always suppose $\varepsilon < \nu(x), \nu(y)$ to avoid certain problems). We define a finite set V_ε equipped with a transition kernel $(p_\varepsilon(z, v))_{z, v \in V_\varepsilon}$ through the following way:

Definition 5.11. V_ε is obtained by identifying all the points with ν -mass less than ε to a cemetery point ∂ :

$$V_\varepsilon = \{z \in V; \nu(z) \geq \varepsilon\} \cup \{\partial\}. \quad (5.24)$$

From now on we will denote the points of V in the same way as their images on V_ε . p_ε then is the kernel p projected on V_ε , with demanding moreover that ∂ is a cemetery point:

$$\forall z, v \in V_\varepsilon \quad p_\varepsilon(z, v) = \begin{cases} p(z, v) & \text{if } z, v \in \tilde{V}_\varepsilon; \\ 0 & \text{if } z = \partial \text{ and } v \in \tilde{V}_\varepsilon; \\ 1 & \text{if } z = v = \partial; \\ \sum_{\nu(w) < \varepsilon} p(z, w) & \text{if } z \in \tilde{V}_\varepsilon \text{ and } w = \partial. \end{cases} \quad (5.25)$$

Then the chain \mathbb{P}^ε satisfies points 1 and 2 of assumption 5.2, and it is reversible with measure $\mu|_{\tilde{V}_\varepsilon}$. This chain is possibly not irreducible, but we can suppose that it is the case by keeping only the irreducible component of \tilde{V}_ε containing x .

So the relation (5.21) is satisfied for V_ε equipped with \mathbb{P}^ε ; it just remains us to prove that $p_\varepsilon^t(x, y) \xrightarrow{\varepsilon \rightarrow 0} p^t(x, y)$, resp. $|P_\varepsilon| \xrightarrow{\varepsilon \rightarrow 0} |P|$.

Let us deal immediately with the operator norm. The very construction of p_ε ensures that for all $z, v \in \tilde{V}_\varepsilon$, we have $p_\varepsilon^t(z, v) \leq p^t(z, v)$. Taking $z = v = x$, the characterization (2.2) of $|P|$ immediately gives that $|P_\varepsilon| \leq |P|$, which is enough for us (but convergence when $\varepsilon \rightarrow 0$ is also true).

Now, we observe that the law of the t first steps of the chain generated by p_ε tends to the law of the initial chain in the sense of total variation. Indeed, given the way how V_ε and p_ε are respectively got from V and p , we have a canonical map which associates a walk on V_ε to a walk on V , so that the law \mathbb{P}_x maps into the law \mathbb{P}_x^ε . That map is defined as follows: the points of the walk on V are sent onto their projections on V_ε until the image walk hits ∂ , and from that time on the walk stays at ∂ . In particular, if a realization of the original chain stays in \tilde{V}_ε up to time t , its image by our map is kept safe on $\{0, \dots, t\}$, and so:

$$\begin{aligned} \left\| \mathbb{P}_x^\varepsilon|_{V_\varepsilon^{\{0, \dots, t\}}} - \mathbb{P}_x|_{V^{\{0, \dots, t\}}} \right\|_{\text{TV}} &\leq \mathbb{P}_x \left(\exists u \in \{0, \dots, t\} X_u \notin \tilde{V}_\varepsilon \right) \\ &\leq \sum_{u=0}^t \mathbb{P}_x (\nu(X_u) < \varepsilon) \xrightarrow[\varepsilon \rightarrow 0]{\text{DCV}} 0. \end{aligned} \quad (5.26)$$

In particular, $p_\varepsilon^t(x, y) \xrightarrow{\varepsilon \rightarrow 0} p^t(x, y)$, and so (5.21) is satisfied for V equipped with p , QED.

A Finite sub-Markov chains

This appendix aims at proving lemma 5.3. Let us recall that in that lemma, we consider a sub-Markov chain on a finite graph \tilde{V} , given by a kernel p , which is irreducible and aperiodic (the fact that the chain is reversible is not used in the proof of lemma 5.3). Denote by n the cardinality of \tilde{V} .

Then the study of the chain may be expressed into matricial terms. We introduce the matrix:

$$M = \left((p(v, z))_{z, v \in \tilde{V}} \right); \quad (\text{A.1})$$

then the aperiodicity condition translates into the existence of a time t_0 such that, for all $t \geq t_0$, M^{t_0} has strictly positive coefficients⁽¹¹⁾. On the other hand,

⁽¹¹⁾Actually one can prove that $t_0 = n^2$ necessarily fits.

lemma 5.9 *supra* permits us to consider $|P|$ as the spectral radius of M .

So we have in hand all the assumptions to apply the strongest form of Perron–Frobenius theorem, whose general statement and proof the reader can find in [4], chap. 5:

Proposition A.1 (Perron–Frobenius). *$|P|$ is a simple eigenvalue of M , and all the other eigenvalues of M have an absolute value strictly less than $|P|$. Moreover, the eigenvector v associated to the eigenvalue $|P|$ has all its entries strictly positive.*

Now, let us begin with proving point 1 of lemma 5.3, and in a first time let us prove the second inequality. Markov’s strong property gives us the over-multiplicativity relation:

$$\forall t, u \geq 0 \quad \mathbb{P}_x(\mathcal{R}_{t+u}) \geq \mathbb{P}_x(\mathcal{R}_t)\mathbb{P}_x(\mathcal{R}_u). \quad (\text{A.2})$$

We deduce that, for all $t \geq 1$, one has $\mathbb{P}_x(\mathcal{R}_t)^{1/t} \leq \limsup_{u \rightarrow \infty} \mathbb{P}_x(\mathcal{R}_u)^{1/u}$. Moreover, if $|P| < 1$, we have for all $t \geq 1$:

$$\mathbb{P}(\mathcal{R}_t) \leq \sum_{u \geq t} p^u(x, x) \stackrel{(2.2)}{\leq} \sum_{u \geq t} |P|^u = \frac{|P|^t}{1 - |P|}, \quad (\text{A.3})$$

hence $\limsup_{u \rightarrow \infty} \mathbb{P}_x(\mathcal{R}_u)^{1/u} \leq |P|$, that last relation being also trivially true in the case when $|P| = 1$. Finally, the second inequality of (5.4) is satisfied for $t \geq 1$ with $c_2 = 1$, the case $t = 0$ being trivial.

For the lower bound, we will only show that there exists a constant $c_3 > 0$ such that one has, for t large enough:

$$p^t(x, x) \geq c_3 |P|^t, \quad (\text{A.4})$$

the first inequality of (5.4) then will follow for t large enough, and the case when t is small then will be dealt with by finiteness, thanks to noticing that irreducibility ensures that $\mathbb{P}_x(\mathcal{R}_t) > 0$ for all $t \in \mathbb{N}$. To prove (A.4), let us consider the eigenvector $v = (v_i)_{i \in \tilde{V}}$ associated to the eigenvalue $|P|$. We shall keep in mind that, by proposition A.1, $v_x > 0$. Denote $\bar{v} = \max_i v_i$; the relation $M^t v = |P|^t v$ then gives for all $t \geq 0$:

$$|P|^t v_x = \sum_{z \in \tilde{V}} p^t(z, x) v_z, \quad (\text{A.5})$$

hence

$$\max_{z \in \tilde{V}} p^t(z, x) \geq \frac{v_x}{n\bar{v}} |P|^t. \quad (\text{A.6})$$

Let now t_0 be such that one has $\forall z, v \in \tilde{V} \quad p^{t_0}(z, v) > 0$ (such a t_0 exists by aperiodicity, as we noticed it *supra*), and let us denote $\eta = \min_{z, v \in \tilde{V}} p^{t_0}(z, v) >$

0. For $t \geq t_0$, by (A.6) we can fix z_1 such that $p^{t-t_0}(z_1, x) \geq v_x |P|^{t-t_0}/n\bar{v}$. It follows that we have:

$$p^t(x, x) \geq \frac{\eta v_x}{n\bar{v}} |P|^{t-t_0}, \quad (\text{A.7})$$

hence (A.4) with $c_3 = \eta v_x/n\bar{v}|P|^{t_0}$.

Now, let us look at the fine behaviour of the sequence $\mathbb{P}_x(\mathcal{R}_t)$ when $t \rightarrow \infty$. Markov's property gives:

$$\mathbb{P}_x(\mathcal{R}_t) = \sum_{z \in \tilde{V}} p^t(x, z) \mathbb{P}_z(\mathcal{R}_0), \quad (\text{A.8})$$

subsequently to prove point 2 of lemma 5.3, we just have to show that we have $p^{t+1}(x, z)/p^t(x, z) \xrightarrow[t \rightarrow \infty]{} |P|$ for all $z \in \tilde{V}$. More precisely, we will show that there exists a constant $c_4(z) > 0$ such that $p^t(x, z)/|P|^t \xrightarrow[t \rightarrow \infty]{} c_4(z)$.

$p^t(x, z)$ can be rewritten in matricial terms as ${}^T\delta_z M^t \delta_x$. Now theorem A.1 gives us that, denoting by \dot{M} the matrix of the projection on $\mathbb{R}v$ relatively to the sum of the characteristic spaces relative to the eigenvalues of M other than $|P|$:

$$\frac{1}{|P|^t} M^t \xrightarrow[t \rightarrow \infty]{} \dot{M}. \quad (\text{A.9})$$

Subsequently, $p^t(x, z)/|P|^t$ tends to the value $c_4(z) = {}^T\delta_y \dot{M} \delta_x$ when t tends to infinity. The non-zerosness of $c_4(z)$ then is a consequence of point 1 which we have proved a few lines above. Indeed, taking again the notations t_0 and η used above, we have by Markov's property:

$$\forall t \geq t_0 \quad p^t(x, z) \geq p^{t-t_0}(x, x) p^{t_0}(x, z) \geq c_2 |P|^{t-t_0} \eta, \quad (\text{A.10})$$

hence $c_4 \geq c_2 > 0$.

Point 3 may be the most subtle. In fact we will prove that for all z in \tilde{V} , $\mathbb{P}_z(\tau_x = t)$ decreases exponentially *with an exponential decay factor strictly less than $|P|$* .

More precisely, we will estimate the decay factor of $\mathbb{P}_z(\tau_x \geq t \text{ and } \tau_\partial \geq t) = \mathbb{P}_z((\forall u \in \{0, \dots, t-1\})(X_u \neq x, \partial))$. In other words, we have to look at the decay speed of the sub-Markov chain associated to the transition kernel p , *but restricted to $\tilde{V} \setminus \{x\}$* . In matricial words, it is the spectral radius of the matrix M^* , which is the matrix M where the x -th line (and if you want also the x -th column) has been replaced by zeroes. Let us denote by $|P^*|$ its spectral radius. The weak form of Perron-Frobenius theorem (cf. [4]) claims that there exists a $|P^*|$ -eigenvector v^* with positive or zero entries for M^* . Each entry of M^* is less than or equal to the corresponding entry of M , and moreover $M^* \neq M$; since M is the matrix of an irreducible aperiodic chain, it follows that, for t sufficiently large, each entry of $(M^*)^t$ is *strictly* less than the corresponding entry of M^t . So, for t sufficiently large:

$$|P^*|^t v^* = (M^*)^t v^* < M^t v^*, \quad (\text{A.11})$$

which means that each entry of $|P^*|^t \mathbf{v}^*$ is strictly less than the corresponding entry of $M^t \mathbf{v}^*$. Now let us reason by contradiction by supposing that $|P^*| \geq |P|$, then (A.11) shows that we can find $t_1 > 0$ and $\rho_1 > |P|$ such that $M^{t_1} \mathbf{v}^* \geq \rho_1^{t_1} \mathbf{v}^*$, hence by iterating:

$$\limsup_{t \rightarrow \infty} |M^t \mathbf{v}^*|^{1/t} \geq \rho_1 > |P|. \quad (\text{A.12})$$

But that is absurd, since the spectral radius of M is actually $|P|$. This implies that $|P^*| < |P|$, what we wanted.

Further readings and acknowledgements

The method developed in this paper, I mean, the use of the martingales introduced in §3, can actually be seen as a discrete variant for the forward/backward martingale decomposition technique of Lyons and Weian [5], as was pointed out to me by Laurent Saloff-Costes. In fact, paper [6] gives a bound for continuous diffusions whose spirit is quite close to that of (3.1).

The present work was launched by informal discussions with my colleagues Y. Ollivier and V. Beffara, who usefully stimulated me into developing my ideas and helped me when necessary. I also thank my Ph. D. advisor C. Villani for having encouraged me and dealt with publication procedures.

References

- [1] N. Th. Varopoulos. Long range estimates for Markov chains. *Bul. Sci. Math. (2)*, 109:225–252, 1985.
- [2] T. K. Carne. A transmutation formula for Markov chains. *Bul. Sci. Math. (2)*, 109:399–405, 1985.
- [3] R. Lyons. Probabilities on trees and networks.
<http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html>.
- [4] D. Serre. *Matrices: Theory and Applications*. Springer, 2002.
- [5] T. J. Lyons and Z. Weian. A crossing estimate for the canonical process on a Dirichlet space and a tightness result. In *Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987)*, pages 249–271, 1988. *Astérisque* # 157–158.
- [6] J. Lunt, T. J. Lyons, and T. S. Zhang. Integrability of functionals of Dirichlet processes, probabilistic representations of semigroups, and estimates of heat kernels. *J. Funct. Anal.*, 153(2):320–342, 1998.

Document VIII

Un nouveau type de décorrélations exponentielles pour les mesures de Gibbs

Un nouveau type de décroissance exponentielle pour les mesures de Gibbs

Rémi Peyre

27 septembre 2007

1 Introduction

Un résultat générale de physique statistique en situation non-critique est l'existence de « décorrélations exponentielles ». L'exemple le plus typique est le modèle d'Ising sur-critique : pour β un paramètre sur-critique, il existe une constante $\gamma > 0$ pour laquelle, pour tous sites i et j , on a :

$$|\mathbb{E}[\sigma_i \sigma_j]| \leq e^{-\gamma|j-i|}, \quad (1)$$

où $|j - i|$ note la distance entre les sites i et j (dans le cas du modèle d'Ising, la distance naturelle est la distance L^1).

Ce genre d'estimée est ce que j'appellerai une relation de décorrélation exponentielle *compacte*, au sens où elle nous dit comment se décorrèlent deux sites, mais où elle ne nous permet pas de savoir comment se décorrèlent des ensembles de points arbitrairement grands. Il existe également une autre sorte d'estimée : si A et B sont des ensembles finis de points situés à distance $\geq d$ l'un de l'autre, on a :

$$\|\mathcal{L}oi(\sigma_A, \sigma_B) - \mathcal{L}oi(\sigma_A) \otimes \mathcal{L}oi(\sigma_B)\|_{VT} \leq \#A \#B e^{-\gamma d}. \quad (2)$$

Cela permet de contrôler des ensembles assez grand dès lors qu'ils sont suffisamment éloignés l'un de l'autre, mais ce n'est pas encore une estimée compacte.

En fait, utiliser la variation totale est voué à l'échec pour obtenir des estimées compactes, dans la mesure où on a l'avis suivant :

Avis 1.1. *Si A et B sont deux hyperplans parallèles, quelle que soit la distance entre eux, on a :*

$$\|\mathcal{L}oi(\sigma_A, \sigma_B) - \mathcal{L}oi(\sigma_A) \otimes \mathcal{L}oi(\sigma_B)\|_{VT} = 1. \quad (3)$$

Nous allons donc introduire une nouvelle façon de mesurer la décorrélation entre deux ensembles, peut-être un peu moins puissante, mais plus souple et qui nous permettra d'obtenir des estimées non compactes. Nous montrerons que pour une classe très vaste de modèles thermodynamiques (incluant en particulier le modèle d'Ising), des décorrélations exponentielles non compactes se produisent dès lors que la température est suffisamment élevée, ce qui montre la pertinence de ce concept.

Notations

- Si $f \in L^1$ et si \mathcal{F} est une tribu, on notera volontiers $f^{\mathcal{F}}$ pour $\mathbb{E}[f|\mathcal{F}]$, en particulier quand on se réfère à l'espérance conditionnelle en tant que variable aléatoire.
- La variance d'une fonction $f \in L^2$ sera notée $\text{Var}[f]$.

2 Niveau de dépendance

Définition

Définition 2.1. Soit $(\Omega, \mathcal{B}, \mathbb{P})$ un espace de probabilités ; soient \mathcal{F}, \mathcal{G} des sous-tribus de \mathcal{B} non-triviales, i.e. contenant des événements de probabilité différente de 0 et 1. On définit :

$$\{\mathcal{F} : \mathcal{G}\} = \sup_{\substack{f \text{ } \mathcal{F}\text{-mesurable} \\ g \text{ } \mathcal{G}\text{-mesurable} \\ \mathbb{E}[f]=0, \mathbb{E}[f^2]=1 \\ \mathbb{E}[g]=0, \mathbb{E}[g^2]=1}} \mathbb{E}[fg]. \quad (4)$$

Nous appellerons $\{\mathcal{F} : \mathcal{G}\}$ le niveau de dépendance des tribus \mathcal{F} et \mathcal{G} . Pour $\varepsilon \in [0, 1]$, on dit que \mathcal{F} et \mathcal{G} sont ε -indépendantes si $\{\mathcal{F} : \mathcal{G}\} \leq \varepsilon$, resp. ε -dépendantes si $\{\mathcal{F} : \mathcal{G}\} \geq \varepsilon$. On parle de même du niveau de dépendances de deux variables aléatoires X et Y , défini par :

$$\{X : Y\} = \{\sigma(X), \sigma(Y)\}. \quad (5)$$

Remarque 2.2. Dans le cas où une des deux tribus au moins est triviale, on posera $\{\mathcal{F} : \mathcal{G}\} = 0$.

Propriétés élémentaires

Proposition 2.3. Pour toutes $f, g \in L^2(\mathbb{P})$ respectivement \mathcal{F} et \mathcal{G} -mesurables, on a :

$$|\mathbb{E}[fg] - \mathbb{E}[f]\mathbb{E}[g]| \leq \{f : g\} \text{Var}^{1/2}(f) \text{Var}^{1/2}(g). \quad (6)$$

En outre, $\{f : g\}$ est la plus petite constante vérifiant (6).

Proposition 2.4. Soient \mathcal{F} et \mathcal{G} deux sous-tribus de \mathcal{B} . Alors :

1. $\{\mathcal{F} : \mathcal{G}\} \in [0, 1]$.
2. $\{\mathcal{F} : \mathcal{G}\} = 0 \Leftrightarrow \mathcal{F}$ et \mathcal{G} sont indépendantes.
3. $\{\mathcal{G} : \mathcal{F}\} = \{\mathcal{F} : \mathcal{G}\}$.
4. Pour $\mathcal{F}' \subset \mathcal{F}$, $\{\mathcal{F}' : \mathcal{G}\} \leq \{\mathcal{F} : \mathcal{G}\}$.
5. Si \mathcal{F} est non-triviale, alors $\{\mathcal{F} : \mathcal{F}\} = 1$.

Proposition 2.5. Pour X_1, X_2, \dots et Y_1, Y_2, \dots deux suites de variables aléatoires, le niveau de covariance entre (X_1, X_2, \dots) et (Y_1, Y_2, \dots) est le supremum des niveaux de covariance entre des parties finies de ces suites :

$$\{(X_i)_{i \in \mathbb{N}} : (Y_j)_{j \in \mathbb{N}}\} = \sup_{\substack{I \subset \mathbb{N} \text{ fini} \\ J \subset \mathbb{N} \text{ fini}}} \{(X_i)_{i \in I} : (Y_j)_{j \in J}\}. \quad (7)$$

Proposition 2.6. Si $X \rightarrow Y \rightarrow Z$ est une chaîne de Markov, alors $\{X : (Y, Z)\} = \{X : Y\}$.

Proposition 2.7. Soient \mathcal{F} et \mathcal{G} deux sous-tribus de \mathcal{B} , avec \mathcal{F} non triviale. On peut définir une « projection » de $L^2(\mathcal{G})$ dans $L^2(\mathcal{F})$ par :

$$\begin{aligned} \pi : L^2(\mathcal{G}) &\longrightarrow L^2(\mathcal{F}) \\ g &\longmapsto \mathbb{E}[g|\mathcal{F}] \end{aligned} \quad (8)$$

Il est clair que cette projection est linéaire, et contractante. Sur les fonctions constantes, π est l'identité ce qui permet de la quotienter en une application

$$\bar{\pi} : L^2(\mathcal{G})/\mathbb{R} \longrightarrow L^2(\mathcal{F})/\mathbb{R}. \quad (9)$$

Alors $\{\mathcal{F} : \mathcal{G}\}$ est la norme opérateur de $\bar{\pi}$.

Corollaire 2.8. Si $X \rightarrow Y \rightarrow Z$ est une chaîne de Markov, alors :

$$\{X : Z\} \leq \{X : Y\}\{Y : Z\}. \quad (10)$$

Remarque 2.9. L'inégalité peut être stricte. Prendre par exemple $X = (a_1, b_1)$, $Y = (a_2, b_1)$ et $Z = (a_2, b_2)$ où a_1, a_2, b_1 et b_2 sont quatre pile ou face indépendants ; alors $\{X : Y\}, \{Y : Z\} = 1$ et pourtant $\{X : Z\} = 0$.

Corollaire 2.10. Soit $\dots \rightarrow X_{n-1} \rightarrow X_n \rightarrow X_{n+1} \rightarrow \dots$ une chaîne de Markov stationnaire et réversible. Alors, pour tout $n \geq 0$:

$$\{X_0 : X_n\} = \{X_0 : X_1\}^n. \quad (11)$$

Remarque 2.11. Évidemment, le sens \leq de (11) était déjà donné par le corollaire 2.8 ; c'est le sens \geq qui est en fait intéressant.

Démonstration. L'application $\bar{\pi}$ est alors autoadjointe, et donc $\|\bar{\pi}^n\| = \|\bar{\pi}\|^n$. □

Remarque 2.12. On peut généraliser ce résultat en supposant que la chaîne de Markov est *quasi-réversible*, avec la

Définition 2.13. Un noyau de transition d'un espace mesurable (Ω, \mathcal{F}) dans lui-même, préservant une mesure de probabilité \mathbb{P} , est dit *quasi-réversible* pour \mathbb{P} quand il commute avec son noyau de transition adjoint.

Exemple 2.14. 1. Une bijection bimesurable préservant la mesure \mathbb{P} est un noyau markovien quasi-réversible.

2. La composée de deux noyaux quasi-markoviens réversibles qui commutent est quasi-réversible.

3. Une marche aléatoire i.i.d. sur un groupe compact abélien est quasi-réversible.

3 Lemme de tensorisation

Lemme 3.1. Soient X et $(Y_i)_{i \in I}$ des variables aléatoires, vérifiant pour une certaine famille de $(\varepsilon_i)_{i \in I}$:

$$\forall i \in I \quad \{X : Y_i\} \leq \varepsilon_i. \quad (12)$$

On suppose en outre que la propriété (12) est vraie même conditionnellement à la valeur de certains $Y_{j \neq i}$, i.e. conditionnellement aux tribus $\sigma(Y_j, j \in J)$ pour $J \subset I \setminus \{i\}$. Alors :

$$\{X : \vec{Y}\} \leq \left(\sum_{i \in I} \varepsilon_i^2 \right)^{1/2}. \quad (13)$$

Démonstration. Soit f une fonction X -mesurable L^2 et g une fonction \vec{Y} -mesurable L^2 , avec g centrée. On suppose I muni d'un ordre total, et on note :

$$\mathcal{F}_i = \sigma(Y_j, j \leq i), \quad \mathcal{F}_{i-} = \sigma(Y_j, j < i), \quad (14)$$

puis

$$g_i = g^{\mathcal{F}_i} - g^{\mathcal{F}_{i-}}. \quad (15)$$

Alors $g = \sum_{i \in I} g_i$ où g_i est \mathcal{F}_i -mesurable et centrée par rapport à \mathcal{F}_{i-} , de sorte que :

$$\text{Var}[g] = \sum_{i \in I} \text{Var}[g_i]. \quad (16)$$

Majorons $\mathbb{E}[f g_i]$: conditionnellement à \mathcal{F}_{i-} , g_i et f , qui sont respectivement Y_i - et X -mesurables, vérifient :

$$\mathbb{E}[f g_i | \mathcal{F}_{i-}] \leq \varepsilon_i \mathbb{E}^{1/2}[f^2 | \mathcal{F}_{i-}] \text{Var}^{1/2}[g_i | \mathcal{F}_{i-}], \quad (17)$$

d'où en déconditionnant et en utilisant l'inégalité de Cauchy-Schwartz :

$$\mathbb{E}[fg_i] \leq \mathbb{E}^{1/2}[f^2] \mathbf{Var}^{1/2}[g_i]. \quad (18)$$

Sommant, il en résulte :

$$\mathbb{E}[fg] \leq \|f\|_{L^2} \left(\sum_{i \in I} \varepsilon_i \mathbf{Var}^{1/2}[g_i] \right), \quad (19)$$

et le résultat en découle par l'inégalité de Cauchy-Schwartz. \square

Lemme 3.2. Soient $(X_i)_{i \in \mathbb{Z}}$ et $(Y_j)_{j \in \mathbb{Z}}$ des variables aléatoires, vérifiant pour une certaine famille de $(\varepsilon_i)_{i \in \mathbb{Z}}$:

$$\forall i \in I, j \in J \quad \{X_i : Y_j\} \leq \varepsilon_{j-i}. \quad (20)$$

On suppose en outre que la propriété (20) est vraie même conditionnellement à la valeur de certains $X_{i' \neq i}$ et $Y_{j' \neq j}$. Alors :

$$\{\vec{X} : \vec{Y}\} \leq \sum_{z \in \mathbb{Z}} \left(\varepsilon_z + \left(\sum_{y \leq z-1} \varepsilon_y^2 \right)^{1/2} \left(\sum_{y \geq z+1} \varepsilon_y^2 \right)^{1/2} \right). \quad (21)$$

Démonstration. Soient f et g resp. \vec{X} - et \vec{Y} -mesurables, centrées. On décompose comme précédemment f en $\sum_{i \in \mathbb{Z}} f_i$, resp. $g = \sum_{i \in \mathbb{Z}} g_j$.

Il va nous falloir majorer $\mathbb{E}[f_i g_j]$. En conditionnant par rapport à $\mathcal{F}_{i-} \vee \mathcal{G}_{j-}$, on obtient :

$$\begin{aligned} & \mathbb{E}[f_i g_j | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}] \leq \\ & \mathbb{E}[f_i | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}] \mathbb{E}[g_j | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}] + \varepsilon_{j-i} \mathbf{Var}^{1/2}[f_i | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}] \mathbf{Var}^{1/2}[g_j | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}]. \end{aligned} \quad (22)$$

\square

Après déconditionnement, le second terme donne par l'inégalité de Cauchy-Schwartz :

$$\varepsilon_{j-i} \mathbf{Var}^{1/2}[f_i] \mathbf{Var}^{1/2}[g_j]. \quad (23)$$

Si nous sommons à $j - i$ constant, cela va nous donner (après un coup de Cauchy-Schwartz) du

$$\varepsilon_{j-i} \mathbf{Var}^{1/2}[f] \mathbf{Var}^{1/2}[g]. \quad (24)$$

On a donc le premier terme du membre de droite de (21).

Pour majorer le premier terme, il va falloir majorer $\mathbb{E}[f_i | \mathcal{F}_{i-} \vee \mathcal{G}_{j-}]$, resp. son analogue pour g . Conditionnellement à \mathcal{F}_{i-} , f_i est centrée et ne dépend que de x_i . Or, toujours conditionnellement à \mathcal{F}_{i-} , x_i est α_{j-i} -indépendante de \mathcal{G}_{j-} , avec :

$$\alpha_{j-i} = \left(\sum_{z \leq j-i-1} \varepsilon_z^2 \right)^{1/2}, \quad (25)$$

en vertu du lemme 3.1. Il s'ensuit, d'après la proposition 2.7, que :

$$\mathbb{E}[(f_i^{\mathcal{F}_{i-}} - \mathbb{V}g_{j-})^2 | \mathcal{F}_{i-}] \leq \alpha_{j-i}^2 \mathbb{E}[\text{Var}[f_i | \mathcal{F}_{i-}] | \mathcal{F}_{i-}], \quad (26)$$

ce qui se déconditionne en :

$$\mathbb{E}[(f_i^{\mathcal{F}_{i-}} - \mathbb{V}g_{j-})^2] \leq \alpha_{j-i}^2 \text{Var}[f_i]. \quad (27)$$

Ainsi, l'intégrale du premier terme du membre de droite de (22) se majore en $\|(f_i^{\mathcal{F}_{i-}} - \mathbb{V}g_{j-})^2\|_{L^2} \|(g_j^{\mathcal{F}_{i-}} - \mathbb{V}g_{j-})^2\|_{L^2}$, qui à son tour se majore d'après (27) en $\alpha_{i-j}\beta_{j-i} \text{Var}[f_i] \text{Var}[g_j]$, où β_{j-i} est défini comme on pense. Ensuite, on somme à $j-i$ fixé en appliquant l'inégalité de Cauchy-Schwartz pour obtenir le second terme du membre de droite de (21), ce qui conclut la preuve du lemme.

4 Le résultat

On considère un système de physique statistique constitué d'un réseau d'atomes \vec{i} sur \mathbb{Z}^d , où tous les points sont identiques et chaque atome peut prendre un état $\sigma_{\vec{i}}$ dans un ensemble \mathcal{S} fini ; l'énergie d'un atome dans l'état s est notée $h_0(s)$. En outre, chaque atome \vec{i} interagit avec les $\vec{j} \neq i$ tels que $\|\vec{j} - \vec{i}\|_\infty \leq 1$, selon une énergie d'interaction donnée par $h_{\vec{j}-\vec{i}}(\sigma_{\vec{i}}, \sigma_{\vec{j}})$ ⁽¹⁾. On peut alors parler des mesures de Gibbs associées à ce système pour la température $\theta = \beta^{-1}$. Le résultat suivant est alors un classique :

Théorème 4.1 (source à citer). *Si θ est suffisamment élevée, la mesure de Gibbs est unique ; on la note alors \mathbb{P}_θ . En outre, si θ est suffisamment élevée il y a décroissance exponentielle du système au sens suivant : pour toute partie $Z \subset \mathbb{Z}^d$, pour tous $x, y \notin Z$, on a presque-sûrement que*

$$\frac{d\text{Loi}(y|x, Z)}{d\text{Loi}(y|Z)} \geq 1 - e^{-\lambda(\theta)\|y-x\|_\infty} \quad (28)$$

pour un $\lambda(\theta) > 0$. De plus, $\lambda(\theta)$ tend vers $+\infty$ quand $\theta \rightarrow +\infty$.

Notre résultat, lui, est le suivant :

Théorème 4.2. *Si θ est suffisamment élevée, il existe une constante $\lambda'(\theta) > 0$ telle que, pour tous sous-ensembles $A, B \subset \mathbb{Z}^d$ disjoints, notant σ_A l'état des sites de A et σ_B l'état des sites de B , on ait :*

$$\{\sigma_A : \sigma_B\} \leq e^{-\lambda'(\theta)}. \quad (29)$$

Ce résultat reste par ailleurs vrai même si on a conditionné par rapport à une partie $C \subset \mathbb{Z}^d$ disjointe de A et de B . En outre, $\lambda'(\theta) \xrightarrow{\theta \rightarrow \infty} \infty$.

¹On rajoute l'hypothèse de symétrie $h_{-\vec{a}}(\sigma_1, \sigma_2) = h_{\vec{a}}(\sigma_2, \sigma_1)$.

En utilisant une chaîne de Markov entre A et B , la proposition 2.8 nous donne le corollaire suivant :

Corollaire 4.3. *Pour toutes parties $A, B \subset \mathbb{Z}^d$, on a, même sous conditionnement éventuel par rapport à C :*

$$\{\sigma_A : \sigma_B\} \leq e^{-\lambda'(\theta)d(A,B)}, \quad (30)$$

où $d(A, B)$ est la distance entre les ensembles A et B ⁽²⁾.

5 La démonstration

La démonstration de notre théorème 4.2 utilise en fait comme seule hypothèse le résultat du théorème 4.1 et la structure géométrique de \mathbb{Z}^d . Dans un premier temps, nous allons étudier le niveau de dépendance de deux sous-espaces affines parallèles, puis nous nous servirons de ces premiers résultats pour obtenir notre preuve par une récurrence sur la dimension du plus petit sous-espace affine contenant A et B .

☛ *Dans tout ce qui va suivre, on travaille toujours sur le modèle de physique statistique dans \mathbb{Z}^d , même quand les corrélations qu'on étudie n'en concernent qu'un sous-espace. En outre, dans tous les résultats qu'on va mettre, il est sous-entendu que ces résultats restent vrais quand on conditionne par rapport à l'état d'une certaine partie de \mathbb{Z}^d (même si elle n'est pas contenue dans le sous-espace d'intérêt).*

Définition 5.1. *On appellera r -sous-espace de \mathbb{Z}^d un sous-espace affine de \mathbb{Z}^d (vu comme \mathbb{Z} -module affine) admettant une base constituée de r vecteurs élémentaires de \mathbb{Z}^d . On parlera resp. de point, droite, plan pour $r = 0, 1, 2$ respectivement.*

On a la proposition intermédiaire suivante :

Proposition 5.2. *Il existe des constantes $\lambda^{(r)}(\theta)$, $r \in \{0, \dots, d-1\}$, avec $\lambda^{(r)}(\theta) \xrightarrow{\theta \rightarrow \infty} \infty$ pour tout r , telles que, si A et B sont des r -sous-espaces parallèles à distance Δ l'un de l'autre, on ait :*

$$\{\sigma_A : \sigma_B\} \leq e^{-\lambda^{(r)}(\theta)\Delta}. \quad (31)$$

Démonstration. La preuve se fait par récurrence sur r . Pour $r = 0$, c'est une conséquence directe du théorème 4.1, via le lemme suivant :

Lemme 5.3. *Soient X et Y deux variables aléatoires telles que presque-sûrement, on ait $d\text{Loi}(X, Y) / (d\text{Loi}(X) \otimes d\text{Loi}(Y)) \geq 1 - \varepsilon$, alors $\{X : Y\} \leq \varepsilon$.*

²I.e. $d(A, B) = \inf_{\substack{\vec{a} \in A \\ \vec{b} \in B}} \|\vec{b} - \vec{a}\|_\infty$.

Nous allons expliciter comment on passe de $r = 0$ à $r = 1$: considérons deux droites parallèles A et B , situées à distance Δ l'une de l'autre. On numérote les points de A par $\dots, a_{z-1}, a_z, a_{z+1}, \dots$ en prenant soin de les numérotter dans l'ordre de la droite. Alors on en déduit une numérotation similaire pour la droite B , où chaque point $b \in B$ reçoit le numéro du point de A le plus proche de lui en distance L^1 ; on observe qu'alors les points a_i et b_j sont situés à distance $\Delta \vee |j - i|$ l'un de l'autre. On est alors en situation d'utiliser le lemme 3.2.

Pour obtenir exactement le résultat que nous souhaitons, et en vue de refaire plusieurs fois la même chose, nous allons utiliser le lemme technique suivant :

Lemme 5.4. *Soit $\lambda > 0$; pour $\Delta \in \mathbb{N} \setminus \{0\}$, $z \in \mathbb{Z}$, on pose $\varepsilon_z^{(\Delta)} = e^{-(\Delta \vee |z|)\lambda}$. Alors, si λ est suffisamment grand, il existe un $\mu > 0$ tel que :*

$$\sum_{z \in \mathbb{Z}} \left(\varepsilon_z^{(\Delta)} + \left(\sum_{y \leq z-1} \varepsilon_y^{(\Delta)^2} \right)^{1/2} \left(\sum_{y \geq z+1} \varepsilon_y^{(\Delta)^2} \right)^{1/2} \right) \leq e^{-\Delta \mu(\lambda)}, \quad (32)$$

avec $\mu(\lambda) \rightarrow \infty$ quand $\lambda \rightarrow \infty$.

Nous obtenons ainsi que deux droites situées à distance Δ l'une de l'autre sont $\exp(-\mu(\lambda(\theta))\Delta)$ -indépendantes, ce qui est également trivialement vrai pour $\Delta = 0$. Pour passer de $r = 1$ à $r = 2$, on va diviser les plans parallèles en droites parallèles, qu'on numérote de la même façon que précédemment. Comme on sait contrôler l'indépendance de deux droites parallèles par hypothèse de récurrence, on peut appliquer à nouveau le lemme 3.2, et ainsi de suite. On obtient finalement une expression de $\lambda^{(r)}(\theta)$ comme :

$$\lambda^{(r)}(\theta) = \mu^{or}(\lambda(\theta)). \quad (33)$$

□

Nous pouvons maintenant démontrer le théorème à proprement parler. En fait, on va montrer par récurrence que :

Proposition 5.5. *Si A et B sont des sous-ensembles disjoints d'un même r -sous-espace de \mathbb{Z}^d ($r \in \{1, \dots, d\}$), on a $\{\sigma_A : \sigma_B\} \leq C^{(r)}$ pour une certaine constante $C^{(r)} < 1$ qui tend vers 0 quand θ tend vers l'infini.*

Démonstration. L'initialisation de la récurrence se fait dans le cas $r = 0$, qui est tellement trivial qu'il est vide. Nous allons montrer à titre d'exemple comment passer de $r = 2$ à $r = 3$, sachant que le principe est toujours exactement le même.

On considère un 3-sous-espace de \mathbb{Z}^d , qu'on découpe en plans parallèles $\dots, \mathcal{P}_z, \mathcal{P}_{z+1}, \dots$, indexés par \mathbb{Z} dans l'ordre. On va alors utiliser le lemme 3.2, où X_i est l'état de sites de $A \cap \mathcal{P}_i$, resp. Y_i celui des sites de $B \cap \mathcal{P}_i$. On se

retrouve alors en situation d'utiliser le lemme 3.2, avec $\varepsilon_0 = C^{(2)}$ et, pour $|z| \geq 1$, $\varepsilon_z = e^{-\lambda^{(2)}|z|}$ d'après la proposition 5.2. Quand θ tend vers l'infini, $\lambda^{(2)}$ tend vers l'infini et $C^{(2)}$ vers 0, de sorte qu'on obtient une valeur de $C^{(3)}$ qui tend vers 0. \square

Le théorème 4.2 est alors simplement la proposition 5.5 appliqué avec $r = d$.

Document IX

Moyennes déviations des mesures empiriques

Moyennes déviations des mesures empiriques

Rémi Peyre

17 août 2007

Introduction

Sur un espace métrique, une façon intéressante de mesurer l'écart entre deux distributions de couplage est la distance de Wasserstein L^1 , ou distance de couplage :

Définition 0.1. Soit (E, d) un espace métrique ; soient μ, ν deux mesures de probabilité sur E (muni de sa tribu borélienne). Un couplage entre μ et ν est une mesure de probabilité ξ sur $E \times E$ dont les deux marginales respectives sont μ et ν . Notons $\mathcal{C}(\mu, \nu)$ l'ensemble des couplages entre μ et ν ; alors la distance de couplage entre μ et ν est définie comme :

$$W^1(\mu, \nu) = \inf_{\xi \in \mathcal{C}(\mu, \nu)} \int_{E \times E} d(x, y) \, d\xi(x, y). \quad (1)$$

Il est clair qu'il s'agit effectivement d'une distance ; je noterai également $\|\nu - \mu\|_{W^1}$ pour $W^1(\mu, \nu)$.

Remarque 0.2. La distance de couplage entre μ et ν est l'énergie que doit dépenser un ouvrier pour transformer μ en ν , sachant que son effort est proportionnel à la quantité de terre qu'il transporte et à la distance sur laquelle il la transporte.

Contrôler la distance de couplage entre μ et ν permet de contrôler la différence entre l'intégrale d'une fonction contre μ et contre ν :

Proposition 0.3. Pour toute fonction f k -lipschitzienne,

$$\left| \int f \, d\nu - \int f \, d\mu \right| \leq k \|\nu - \mu\|_{W^1}. \quad (2)$$

Soit μ une mesure de probabilité sur un espace métrique et $N \geq 1$ un entier ; en tirant N points i.i.d. distribués selon μ , on obtient une mesure empirique $\hat{\mu}^N$, aléatoire, qui approche bien μ dans le sens où on démontre que $\hat{\mu}^N$ converge en probabilité vers μ . Nous aimerions bien avoir des estimées plus fines sur cette convergence, en particulier des estimées non asymptotiques.

La méthode la plus logique consiste alors à passer par les méthodes usuelles de grandes déviations, notamment le théorème de Sanov. Si nous cherchons à contrôler les déviations de $\hat{\mu}^N$ par rapport à une seule fonction f , le résultat optimal est le suivant :

Définition 0.4. *On dit qu'une mesure μ sur un espace métrique E vérifie une inégalité $T_1(\lambda)$ si pour toute mesure de probabilité ν , l'entropie relative de ν par rapport à μ vérifie :*

$$H(\nu|\mu) \geq \lambda \frac{(W^1(\mu, \nu))^2}{2}. \quad (3)$$

Il est possible de montrer, cf. p. ex. [1], que toutes les mesures de probabilité sur E vérifient une même inégalité $T_1(\lambda)$ dès lors que E est borné. Comme nous le disions, l'inégalité T_1 apparaît comme le critère le plus pertinent pour évaluer les déviations contre un fonction :

Théorème 0.5. *Soit μ une mesure de probabilité vérifiant une inégalité $T_1(\lambda)$; alors pour toute fonction f 1-lipschitzienne :*

$$\mathbb{P}\left(\left|\int f d\hat{\mu}^N - \int f d\mu\right| \geq \varepsilon\right) \leq 2e^{-\lambda N \varepsilon^2/2}. \quad (4)$$

En outre, la constante λ qui apparaît dans (4) ne peut être améliorée.

Cela dit, ce théorème ne permet pas de contrôler $W^1(\hat{\mu}^N, \mu)$, car cela reviendrait à borner *simultanément* les écarts $\int f d\hat{\mu}^N - \int f d\mu$ pour *toutes* les fonctions f 1-lipschitziennes !

Dans [2], Villani *et al.* tentent de contourner cet écueil pour obtenir un résultat le plus proche possible du théorème 0.5 avec la distance W^1 . Plus précisément, ils établissent le

Théorème 0.6 (Villani *et al.*). *Soient E un espace métrique borné de dimension $d^{(1)}$, et μ une mesure de probabilité portée par E . On sait que μ vérifie une inégalité $T_1(\lambda)$ pour un λ ne dépendant que de E ; donnons-nous ce λ . Soit $N \geq 1$ un entier ; notons $\hat{\mu}^N$ la mesure empirique de N points i.i.d. tirés selon μ . Alors pour tous $\lambda' < \lambda$ et $d' > d$, on peut trouver une constante C ne dépendant que de λ' , d' et E telle que pour tout $\varepsilon > 0$:*

$$N \geq \frac{C}{\varepsilon^{d'+2}} \quad \Rightarrow \quad \mathbb{P}(W^1(\mu, \hat{\mu}^N) \geq \varepsilon) \leq e^{-\frac{\lambda'}{2} N \varepsilon^2}. \quad (5)$$

¹Voir l'appendice B pour la définition de dimension utilisée dans cet article.

Remarque 0.7. J'ai légèrement modifié l'énoncé du théorème par rapport à [2] afin de le rendre comparable au travail qui va suivre.

Dans la mesure où λ' peut être rendu aussi proche de λ qu'on veut, le théorème 0.6 est pratiquement optimal. Cependant ce résultat n'est pas très intéressant d'un point de vue pratique. En effet, à N fixé, il ne fonctionne que pour $\varepsilon \geq \left(\frac{C}{N}\right)^{1/(d'+2)}$, ce qui, comme nous le verrons, est largement au-delà des déviations typiques de $\hat{\mu}^N$ ⁽²⁾. Notre objectif est donc de trouver une majoration qui fonctionne bien, cette fois-ci, à N fixé quand ε devient grand devant les valeurs typiques de $W^1(\mu, \hat{\mu}^N)$.

1 Le résultat

1.1 Énoncé du théorème

Définition 1.1. Soit X une variable aléatoire réelle positive ; pour $\lambda > 0$ on appellera λ -espérance X , et on notera $E_\lambda(X)$, le nombre :

$$E_\lambda(X) = \frac{1}{\lambda} \ln \mathbb{E}[e^{\lambda X}]. \quad (6)$$

Pour $\lambda = 0$, on notera par ailleurs :

$$E_0(X) = \mathbb{E}[X]. \quad (7)$$

Théorème 1.2. Soit $d > 2$. Soit X un espace métrique borné de dimension d et soit μ une mesure de probabilité sur E . Alors tout pour $N \geq 1$, pour $\hat{\mu}^N$ la mesure empirique à N points associée à μ , on a pour tout $a > 0$:

$$E_a(\|\hat{\mu}^N - \mu\|_{W^1}) \leq \Phi\left(\frac{\lambda D}{N^{1/2}}\right) \frac{D}{N^{1/d}}, \quad (8)$$

où Φ est une certaine fonction croissante purement déterministe (dépendant de d), et D une longueur ne dépendant que de X (et de d).

1.2 Démonstration

Avant de commencer la démonstration proprement dite, nous allons donner et prouver le lemme qui fait tout :

²Par contre, si une variante du théorème 0.6 où $d' + 2$ serait remplacé par $d' \vee 2$ était vraie, alors le résultat serait vraiment utile. Pour ma part, je n'ai pas cherché à optimiser le taux de décroissance exponentielle : c'est moins esthétique du point de vue théorique, mais c'est amplement suffisant pour les applications qu'on peut imaginer, et présente en outre l'avantage d'une démonstration simple.

Lemme 1.3. Soit X un espace métrique de diamètre D , partitionné en une union disjointe $X = X_a \cup X_b$. Soit μ une mesure de probabilité sur X ; soit $N \in \mathbb{N}^*$. Soient k_a et k_b les entiers les plus proches de $N\mu(X_a)$ et $N\mu(X_b)$, avec $k_a + k_b = N$. Pour $i \in \{a, b\}$, notons μ_i la mesure de probabilité définie par $d\mu_i = \mathbb{1}_{X_i} d\mu / \mu(X_i)$ ⁽³⁾, alors on a pour tout $\lambda > 0$:

$$E_\lambda(\|\widehat{\mu}^N - \widehat{\mu}\|_{W^1}) \leq \left(\widetilde{\gamma}(\lambda DN^{-1/2})N^{-1/2} + \frac{1}{2N}\right)D + \frac{k_a}{N}E_{\frac{k_a}{N}\lambda}(\|\widehat{\mu}_a^{k_a} - \mu_a\|_{W^1}) + \frac{k_b}{N}E_{\frac{k_b}{N}\lambda}(\|\widehat{\mu}_b^{k_b} - \mu_b\|_{W^1}), \quad (9)$$

où $\widetilde{\gamma}$ est une certaine fonction croissante définie en appendice A.

Démonstration. Soient x_1, \dots, x_N N points i.i.d. tirés selon μ ; notons Z_a , resp. Z_b , le nombre de ces points situés dans X_a . $Z_a - k_a$ et $Z_b - k_b$ sont deux entiers opposés; supposons pour fixer les idées que ce soit $Z_a - k_a$ qui est positif. On va alors déplacer $Z_a - k_a$ points de X_a vers X_b comme suit : parmi les x_i qui sont dans X_a , on tire au sort $Z_a - k_a$ points qu'on efface, et on ajoute dans X_b le même nombre de points, i.i.d. selon μ_b .

À l'issue de cette manœuvre, on a déplacé $|Z_a - k_a|$ points sur une distance D au plus, soit un coût de transport d'au plus $\frac{D}{N}|Z_a - k_a|$, et on se retrouve avec k_a points i.i.d. selon μ_a dans X_a et k_b points i.i.d. selon μ_b dans X_b , sachant que le nombre (algébrique) de points déplacés, la distribution finale dans X_a et la distribution finale dans X_b sont indépendants. On utilise alors le corollaire A.2 du lemme A.1 en appendice pour obtenir le résultat annoncé. \square

Maintenant, nous allons démontrer le théorème 1.2 :

Démonstration. On suppose données $\{X_\emptyset\}, \{X_a, X_b\}, \{X_{aa}, X_{ab}, X_{ba}, X_{bb}\}, \dots$ des partitions de plus en plus fines de X vérifiant les conditions de la définition B.1, avec les mêmes notations. Soit $N \geq 1$. Pour m un mot de l'alphabet $\{a, b\}$, on définit la mesure de probabilité $\mu_m = \mathbb{1}_{X_m} \mu / \mu(X_m)$. On va également définir des entiers k_m indexés par les mots de l'alphabet $\{a, b\}$ de la sorte :

- $k_\emptyset = N$;
- Pour tout mot m , k_{ma} , resp. k_{mb} , est l'entier le plus proche de $\frac{\mu(X_{ma})}{\mu(X_m)} k_a$, resp. $\frac{\mu(X_{mb})}{\mu(X_m)} k_b$.

Soit maintenant $\lambda > 0$; pour tout entier $n \geq 0$, on a :

Proposition 1.4.

$$E_\lambda(W_\emptyset) \leq \frac{\widetilde{\gamma}\left(\frac{\lambda D}{N^{1/2}}\right)}{1 - 2^{-(1/2-1/d)}} 2^{n(1/2-1/d)} \frac{D}{N^{1/2}} + \frac{2^{n(1-1/d)}}{1 - 2^{-(1-1/d)}} \frac{D}{2N} + \frac{1}{2^{n/d}} D. \quad (10)$$

³Si $\mu(X_i) = 0$ alors μ_i n'est pas définie, mais cela n'a aucune importance car sa contribution à l'équation (9) est alors nulle.

En effet, on a la majoration triviale $E_\lambda(W_\emptyset) \leq D$, puis on raisonne par récurrence sur n à l'aide du lemme 1.3 (en prouvant le résultat pour toutes les valeurs de λ à chaque étape). Dans la récurrence, le seul passage un peu astucieux est de penser à utiliser l'inégalité de concavité :

$$(k_a + k_b)^{1/2} \leq \sqrt{2}(k_a^{1/2} + k_b^{1/2}). \quad (11)$$

Pour conclure, il ne reste plus qu'à optimiser n dans (10). Travaillons dans le cas $\lambda = 0$ (autrement dit, on cherche à optimiser l'espérance de W_\emptyset), dans lequel le préfacteur $\tilde{\gamma}(\lambda D/N^{1/2})$ n'intervient pas. Alors les trois termes de (10) deviennent du même ordre de grandeur simultanément lorsque $2^n \approx N$. Si nous choisissons pour n l'entier le plus proche de $\log_2 N$, on récupère finalement :

$$E_\lambda(W_\emptyset) \leq \left[\alpha \tilde{\gamma} \left(\frac{\lambda D}{N^{1/2}} \right) + \beta \right] \frac{D}{N^{1/d}}, \quad (12)$$

pour des constantes numériques α et β ne dépendant que de d . On en tire immédiatement le résultat annoncé. \square

1.3 Corollaire essentiel

La façon la plus commode d'utiliser le théorème 1.2 est le corollaire suivant :

Corollaire 1.5. *Soit X un espace métrique borné de dimension $d > 2$, μ une mesure de probabilité sur X ; alors on peut trouver une constante C ne dépendant que de X et d telle que pour tout $C' \geq C$:*

$$\forall N \geq 1 \quad \mathbb{P}(\|\hat{\mu}^N - \mu\|_{W^1} \geq \frac{C'}{N^{1/d}}) \leq c_2 \exp(-c_1 N^{1/2-1/d}) \quad (13)$$

pour des constantes $c_1 > 0$ et c_2 ne dépendant pas de μ . En particulier, $\mathbb{P}(\|\hat{\mu}^N - \mu\|_{W^1} \geq C'/N^{1/d})$ décroît plus vite que n'importe quelle fonction puissance.

Remarque 1.6. Ce résultat nous montre que $\|\hat{\mu}^N - \mu\|_{W^1}$ est de l'ordre de $N^{-1/d}$. Cela est optimal puisque par exemple, dans le cas où μ est la mesure uniforme sur $[0, 1]^d$ avec $d \geq 3$ entier, on démontre facilement l'existence d'une constante \tilde{C} telle que pour tout N , on ait :

$$\|\hat{\mu}^N - \mu\|_{W^1} \geq \frac{\tilde{C}}{N^{1/d}} \quad \text{p.s.} \quad (14)$$

En d'autres termes, dès lors qu'on est en dimension $d > 2$, ce qui fait la valeur de $\|\hat{\mu}^N - \mu\|_{W^1}$ n'est pas une hétérogénéité globale de la répartition des x_i , mais simplement le fait qu'on essaie d'approcher une mesure *a priori* continue par une mesure discrète. C'est finalement l'enseignement principal de ce travail.

A Le lemme de contrôle exponentiel

Lemme A.1. *Il existe une fonction purement déterministe $\gamma : \mathbb{R}_+ \rightarrow]0, \infty[$, croissante, telle que pour toute variable aléatoire réelle X centrée vérifiant pour une certaine constante $\sigma \geq 0$:*

$$\forall \lambda \in \mathbb{R} \quad \ln \mathbb{E}[e^{\lambda X}] \leq \frac{\lambda^2 \sigma^2}{2}, \quad (15)$$

on ait pour tout $\lambda \geq 0$:

$$E_\lambda(|X|) \leq \gamma(\sigma\lambda)\sigma. \quad (16)$$

Corollaire A.2. *Soient $N \geq 1$ et $p \in [0, 1]$; alors pour une variable aléatoire X de loi binômiale $\mathcal{B}(N, p)$:*

$$E_\lambda(|X - pN|) \leq \gamma(\lambda\sqrt{N}/2) \frac{\sqrt{N}}{2}. \quad (17)$$

Comme nous nous servirons uniquement de ce corollaire, nous noterons $\tilde{\gamma}$ la fonction définie par $\tilde{\gamma}(x) = \gamma(x/2)$, afin d'alléger les notations.

B Dimension fractale dichotomique

Dans cet article, nous avons utilisé un concept un peu particulier de dimension fractale, que j'appellerai ci la *dimension dichotomique* :

Définition B.1. *Soit X un espace métrique borné. Pour $d > 0$ un réel, on dira que X est fortement de dimension d si, pour une certaine constante $D > 0$, on peut construire des parties $X_\emptyset, X_a, X_b, X_{aa}, \dots$ de X indexées par les mots finis de l'alphabet $\{a, b\}$ telles que :*

1. $X_\emptyset = X$;
2. Pour tout mot m , $X_m = X_{ma} \cup X_{mb}$;
3. Pour tout mot m de longueur ℓ , le diamètre de X_m est majoré par $D/2^{\ell/d}$.

On dira également que X est fortement de dimension 0 quand X est fini.

☞ Dans l'article, nous avons simplement écrit « de dimension ... » pour « fortement de dimension ... ».

Par exemple, toute variété riemannienne compacte de dimension d est fortement de dimension d . Si X est fortement de dimension d , alors sa dimension de Hausdorff est d au plus, mais la réciproque n'est vraie que si X apparaît suffisamment identique à toutes les échelles⁽⁴⁾. C'est néanmoins le cas pour la plupart des fractales usuelles : par exemple, le triangle de Sierpiński a pour dimension dichotomique sa dimension de Hausdorff, soit $\ln 3 / \ln 2$.

⁴Par exemple, l'ensemble des $1/n$ pour $n \in \mathbb{N}^*$, dont les points sont de plus en plus resserrés à mesure qu'on approche de 0, est de dimension dichotomique 1 alors que sa dimension de Hausdorff est 0.

Références

- [1] S. BOBKOV et F. GÖTZE : Exponential integrability and transportation cost related to logarithm Sobolev inequalities. *J. Funct. Anal.*, 163:1–28, 1999.
- [2] F. BOLLEY, A. GUILLIN et C. VILLANI : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Theory and Related Fields*, 137:541–593, 2007.

Document X

Moyennes déviations pour des modèles de champs moyen à chocs

Moyennes déviations pour des modèles de champ moyen à chocs

Rémi Peyre

2 septembre 2007

Résumé

Dans [1], Villani *et al.* utilisent des techniques de concentration de la mesure pour étudier le comportement de certaines classes de modèles de champ moyen en termes de distance de couplage W_1 entre la répartition empirique des particules et leur limite hydrodynamique. Malheureusement, cette technique novatrice, qui donne accès à des renseignements fins sur la mesure empirique, ne semblait pas s'adapter aux cas où l'évolution du système se fait par chocs, l'exemple-type étant l'équation cinétique de Boltzmann. Ici, nous explorons une voie qui permet d'obtenir des estimées W_1 pour les modèles de collision.

Notations

Théorie de la mesure

Soit X un espace mesuré. On utilisera les notations suivantes :

- $\mathcal{M}(X)$ est l'ensemble des mesures signées sur X .
- Pour $M \in \mathbb{R}$, $\mathcal{M}_M(X)$ est l'ensemble des mesures de $\mathcal{M}(X)$ de masse totale M .
- $\mathcal{M}^+(X)$ est l'ensemble des mesures positives finies sur X .
- Pour $M \geq 0$, on pose $\mathcal{M}_M^+(X) = \mathcal{M}^+(X) \cap \mathcal{M}_M(X)$.
- Soient $\mu \in \mathcal{M}(X)$ et $f : X \rightarrow Y$ une application mesurable ; on note $f \circ \mu$ la mesure-image de μ par f .
- Pour $x \in X$, on note δ_x la masse de Dirac en x .
- Pour N points x^1, \dots, x^N de X , on appelle *mesure empirique* des x^i la mesure $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$.

Couplage

Définition 0.1. Soit (X, d) un espace métrique borné muni de sa tribu borélienne. Soit $\mu \in \mathcal{M}_0(X)$, dont on note respectivement μ_+ et μ_- les parties positive et négative. On appelle autocouplage de μ une mesure $\gamma \in \mathcal{M}^+(X \times X)$ dont la première marginale est μ_+ et la seconde μ_- . Le coût de cet autocouplage est la grandeur

$$\int_{X \times X} d(x, y) \, d\gamma(x, y). \quad (1)$$

La norme de Monge de μ désignera alors l'infimum des coûts de tous les couplages de μ possible ; on la notera $\|\mu\|_W$.

Remarque 0.2. En d'autres termes, si μ modélise un sol accumulant des déblais et des remblais, $\|\mu\|_W$ est l'effort que doit fournir un ouvrier qui dépense une certaine énergie par mètre et par volume de terre transporté pour remettre le terrain à plat.

Avis 0.3. $\|\cdot\|$ est effectivement une norme sur $\mathcal{M}_0(X)$.

Définition 0.4. On note $\overline{\mathcal{M}}_0(X)$ la complétion de \mathcal{M}_0 pour $\|\cdot\|_W$. Pour $M \in \mathbb{R}$, on définit la distance de Monge entre deux mesures μ et ν de $\mathcal{M}_M(X)$ comme $\|\mu - \nu\|_W$. Le complété de $\mathcal{M}_M(X)$ pour la distance de Monge est alors noté $\overline{\mathcal{M}}_M(X)$; c'est un $\overline{\mathcal{M}}_0$ -espace affine.

☞ Dorénavant, les $\overline{\mathcal{M}}_M(X)$ et leurs sous-ensembles seront systématiquement munis de la distance de Monge.

Remarque 0.5. Nous avons donc deux notions de convergence dans $\mathcal{M}_1^+(X)$: la convergence faible sur les mesures de probabilité et la convergence sur la distance de Monge.

Avis 0.6. Si X est compact, la convergence pour la distance de Monge sur $\mathcal{M}_1^+(X)$ coïncide avec la convergence faible.

☞ Dans tout l'article, chaque fois que l'on rencontrera des produits d'espaces métriques, ils seront implicitement munis de la distance somme. On notera par ailleurs π_1, π_2, \dots les projections canoniques de l'espace-produit sur les espaces de départ.

Probabilités

- Dans tout le texte, les mesures de probabilité qu'on utilisera seront sans ambiguïté ; on les notera donc génériquement \mathbb{P} , et on notera \mathbb{E} les espérances associées.

- Pour X une variable aléatoire réelle, $\lambda > 0$, on appelle λ -*espérance* de X la quantité :

$$E_\lambda[X] = \frac{1}{\lambda} \ln \mathbb{E}[e^{\lambda X}]. \quad (2)$$

Par continuité, on posera également $E_0[X] = \mathbb{E}[X]$.

- X étant une variable aléatoire, on écrira volontiers « X » pour « la loi de X », le contexte permettant de trancher.
- Une loi \mathbb{P}_1 sur \mathbb{R} est dite *majorée* par une autre loi \mathbb{P}_2 s'il existe des variables aléatoires X_1 et X_2 de lois respectives \mathbb{P}_1 et \mathbb{P}_2 vérifiant $X_1 \leq X_2$. On notera $\mathbb{P}_1 \leq \mathbb{P}_2$.
- $\mathcal{P}(\theta)$ notera un processus de Poisson de paramètre θ .

Constantes

☞ Dans tout l'article, on se fixera un réel $n > 2$; toutes les constantes données seront susceptibles de dépendre implicitement de n .

- On notera génériquement α toute constante purement mathématique.
- On notera génériquement $\Phi(\cdot)$ toute fonction purement mathématique de \mathbb{R}_+ dans $]0, +\infty[$, croissante, et continue en 0 avec $\Phi(0) = 1$.

1 Le modèle

1.1 Données

Soit $n > 2$ un réel fixé pour toute la suite de l'article. On se donne un espace métrique compact X fortement de dimension n au sens de la définition A.1; on note D le diamètre de cet espace.

On se donne deux applications $\tau : X \times X \rightarrow [0, \infty[$ et $h : X \times X \rightarrow \mathcal{M}_1^+(X \times X)$, où τ et h sont symétriques, i.e. :

Hypothèse 1.1. Notant $\mathfrak{t} : X \times X \rightarrow X \times X$ l'application qui envoie (x, y) sur (y, x) , on a $\tau \circ \mathfrak{t} = \tau$ et $h \circ \mathfrak{t} = \mathfrak{t} \circ h$.

On suppose que τ et h vérifient l'

Hypothèse 1.2. 1. τ est bornée; on note $\bar{\tau}$ son supremum.

2. L'application

$$S : X \times X \rightarrow \overline{\mathcal{M}}_0(X \times X) \quad (3)$$

$$(x, y) \mapsto \tau(x, y)(h(x, y) - \delta_{(x,y)})$$

est lipschitzienne; on note κ sa constante de Lipschitz.

☞ Dans toute la suite, μ_0 est une mesure de probabilité fixée sur X . Tout le raisonnement sera valable uniformément en μ_0 .

1.2 Modèle microscopique

On considère $N \geq 1$ particules x^1, \dots, x^N en interaction. On note $\eta = (x^1, \dots, x^N)$ l'état du système, et pour $i < j$, $y^i, y^j \in X$, on pose :

$$\eta_{y^i, y^j}^{i,j} = (x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^{j-1}, y^j, x^{j+1}, \dots, x^N). \quad (4)$$

Alors l'évolution du système se fait de la façon suivante : à $t = 0$, x_0^1, \dots, x_0^N sont distribués indépendamment selon μ_0 , puis l'évolution est markovienne avec pour générateur :

$$\mathcal{L}f(\eta) = \frac{1}{N} \sum_{i < j} \tau(x^i, x^j) \int_{X \times X} (f(\eta_{y^i, y^j}^{i,j}) - f(\eta)) d(h(x^i, x^j))(y^i, y^j). \quad (5)$$

En d'autres termes, de manière poissonnienne avec le taux $\tau(x^i, x^j)N$, le couple de particules (x^i, x^j) se heurte et se redistribue selon la loi $h(x^i, x^j)$, tous les chocs étant indépendants.

☛ On notera μ_t^N la mesure empirique des x^i au temps t .

1.3 Modèle macroscopique

Dans ce modèle, la grandeur d'intérêt est une mesure de probabilité μ sur X qui représente moralement la répartition des particules. L'évolution est une simple équation différentielle dans $\overline{\mathcal{M}}_1(X)$: à $t = 0$, on a $\mu_{t=0} = \mu_0$, puis μ évolue selon l'équation :

$$\partial_t \mu = \overline{S}(\mu), \quad (6)$$

où \overline{S} est définie par :

$$\overline{S}(\mu) = \pi_1 \circ \int_{X \times X} \Delta(x, y) d\frac{\mu \otimes \mu}{2}(x, y) \quad (1). \quad (7)$$

La solution de (6) est unique et définie pour tout temps, en vertu de l'

Avis 1.3. Sous l'hypothèse 1.2-2, \overline{S} est $\kappa/2$ -lipschtzienne.

En outre,

Avis 1.4. Pour tout $t \geq 0$, $\mu_t \in \mathcal{M}_1^+(X)$.

Remarque 1.5. On peut modifier les données du problème pour faire en sorte que τ soit constante égale à $\bar{\tau}$ sans changer les évolutions (5) et (6). En effet, il suffit de considérer à la place de la fonction de sauts h la fonction \tilde{h} définie par :

$$\tilde{h}(x, y) = \frac{\tau(x, y)}{\bar{\tau}} h(x, y) + \left(1 - \frac{\tau(x, y)}{\bar{\tau}}\right) \delta_{(x, y)}. \quad (8)$$

Que τ soit constante nous sera essentiel dans le § 3.1 pour justifier certaine hypothèse d'indépendance ; on le supposera donc.

☛ Nous ferons toujours l'hypothèse que τ est constante, égale à $\bar{\tau}$.

1.4 Limite de champ moyen

Le modèle macroscopique correspond à l'idée intuitive que nous avons de ce que doit être l'évolution d'un grand nombre de particules. En effet, on démontre, cf. [2], que le modèle du § 1.3 est la limite, dans un sens acceptable, de celui du § 1.2 quand $N \rightarrow \infty$:

Théorème 1.6. *Pour tout temps $T \geq 0$, pour toute fonction f lipschitzienne sur X :*

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\sup_{t \in [0, T]} \left| \int_X f d\mu_t - \int_X f d\hat{\mu}_t \right| \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0. \quad (9)$$

Notre objectif est d'établir ici une notion plus forte de limite (à savoir, au sens de la distance de Monge) et de donner une estimation quantitative sur la convergence et les moyennes déviations.

2 Résultats et angle d'attaque

2.1 Énoncé du théorème

Théorème 2.1. *Il existe une constante C telle que pour tout $T \geq 0$:*

$$C' > \frac{e^{\kappa T/2} - 1}{\kappa/2} C^{(2)} \quad \Rightarrow \quad \mathbb{P} \left(\|\hat{\mu}_t^N - \mu_t\|_W \geq \frac{C'}{N^{1/(n+1)}} \right) \Downarrow_{N \rightarrow \infty} 0, \quad (10)$$

où « $f(x) \Downarrow_{x \rightarrow +\infty} 0$ » signifie « $f(x)$ décroît plus vite, quand $N \rightarrow \infty$, que n'importe quel $1/x^q$ ».

2.2 Principe de la démonstration

Comment allons-nous aborder ce problème ? Dans la mesure où on cherche un résultat en distance de Monge, la technique habituelle consistant à prendre une fonction-test et à étudier l'évolution au moyen de martingales ne peut fonctionner. Le principe général doit néanmoins rester le même : lorsqu'un grand nombre de sauts se produisent, leur répartition est à peu près celle prédite par la limite hydrodynamique. En l'occurrence, l'outil que nous allons utiliser est la concentration de la mesure pour les mesures empiriques.

Pour nous accommoder de l'évolution du système avec le temps, nous allons procéder par pas de longueur t_1 . Sur chaque pas, nous allons en fait comparer l'évolution du système stochastique et du système déterministe en supposant les lois de chocs figées, puis nous corrigerons nos calculs pour tenir compte de ce que les lois de chocs évoluent en même temps que le système. Deux effets

antagonistes joueront alors : on a intérêt d'une part à choisir t_1 assez grand pour que la concentration sur les sauts soit pertinente ; d'autre part, si t_1 est trop grand l'approximation effectuée à chaque pas devient trop grossière. Il s'agira donc d'optimiser la valeur de t_1 . Ensuite, quelques manipulations techniques nous permettront de contrôler simultanément la majoration pour tous les temps.

3 Étude sur un pas de temps

Supposons dans cette partie N fixé, et notons pour alléger les notations $W_t = \|\widehat{\mu}_t^N - \mu_t\|$. Dans cette partie, on suppose que $\widehat{\mu}_0^N$ est parfaitement connu. On se fixe un $t_1 > 0$; notre objectif est alors de contrôler W_{t_1} à partir de W_0 .

☞ Dans toute cette partie, t_1 est fixé.

3.1 Découpage en plusieurs termes

Pour contrôler l'écart entre μ_{t_1} et $\widehat{\mu}_{t_1}^N$, on va découper la différence des deux mesures en six termes et appliquer l'inégalité triangulaire. Attendu que la description des termes les plus compliqués nous obligerait à un formalisme plus qu'abscons, nous nous contenterons d'expliquer d'où ces termes proviennent et d'en donner directement un majorant. Le découpage est le suivant :

1. **Divergence de l'ÉDO** : Soit $(\widetilde{\mu}_t)_{0 \leq t \leq t_1}$ vérifiant $\widetilde{\mu}_0 = \widehat{\mu}_0^N$ et évoluant selon (6). On appelle terme de *divergence hydrodynamique* l'expression :

$$\|\widetilde{\mu}_{t_1} - \mu_{t_1}\|_W. \quad (11)$$

2. **Termes du second ordre dans l'ÉDO** : Le terme de *courbure hydrodynamique*, qui traduit l'erreur que l'on fait quand on linéarise l'évolution de (6), est :

$$\|\widehat{\mu}_0^N + t_1 \overline{S}(\widehat{\mu}_0^N) - \widetilde{\mu}_{t_1}\|_W. \quad (12)$$

3. **Artefact d'autochoc** : Une petite astuce vient de ce qu'une particule n'interagit pas avec elle-même ; il s'ensuit un terme d'*autochoc* qui est :

$$\left\| -\frac{t_1}{2N^2} \sum_{i=1}^N S(x^i, x^i) \right\|_W. \quad (13)$$

4. **Concentration** : On considère la mesure \mathfrak{S} sur $\overline{\mathcal{M}}_0(X)$ définie par :

$$\mathfrak{S} = \overline{\tau} t_1 \sum_{i < j} \int_{X \times X} \delta_{\delta_{y^i} + \delta_{y^j} - \delta_{x_0^i} - \delta_{x_0^j}} d(h(x_0^i, x_0^j))(y^i, y^j). \quad (14)$$

Cette mesure a pour masse totale $M = (N - 1)\bar{\tau}t_1/2$; notons $\bar{\mathcal{S}} \in \overline{\mathcal{M}}_0(X)$ sa somme, i.e. :

$$\bar{\mathcal{S}} = \int_{\overline{\mathcal{M}}_0(X)} \sigma \, d\mathfrak{G}(\sigma). \quad (15)$$

Soit $\mathfrak{M} = \lfloor M \rfloor$; on tire maintenant \mathfrak{M} variables i.i.d. $\sigma_1, \dots, \sigma_{\mathfrak{M}}$ à valeurs dans $\overline{\mathcal{M}}_0(X)$ selon la loi $\frac{1}{M}\mathfrak{G}$; dont on note \mathcal{S}^* la somme. Alors le terme de *concentration des sauts* est :

$$\frac{1}{N} \|\mathcal{S}^* - \bar{\mathcal{S}}\|_W. \quad (16)$$

5. Courbure stochastique : Ce terme mesure l'écart entre la situation où les particules x^i et x^j sautent toujours selon la loi $h(x_0^i, x_0^j)$, et la situation réelle où la loi de (x_i, x_j) après une collision à l'instant t est donnée par $h(x_{t-}^i, x_{t-}^j)$, ce qui n'est pas la même chose si une des deux particules i ou j a déjà sauté. Techniquement, on considère \mathfrak{M} points dont le premier est colorié en rouge avec probabilité 0, le second avec probabilité $4/N$, le troisième avec probabilité $8/N$, etc., le tout indépendamment. On note $\Upsilon(4/N, \mathfrak{M})$ la loi du nombre de points rouges; alors le terme de *courbure stochastique* est :

$$\frac{4D}{N} \Upsilon(4/N, \mathfrak{M}). \quad (17)$$

6. Erreur poissonnienne : Ce terme traduit l'écart entre le nombre de sauts qui ont réellement lieu et le nombre qu'on aurait attendu. Plus précisément, c'est la loi de la variable aléatoire :

$$\frac{2D}{N} \left| \mathcal{P}(M) - \mathfrak{M} \right|. \quad (18)$$

Avis 3.1.

$$\|\widehat{\mu}_{t_1} - \mu_{t_1}\|_W \leq (11) + (12) + (13) + (16) + (17) + (18). \quad (19)$$

Remarque 3.2. En outre, on peut faire en sorte que les six termes (11) à (18) soient indépendants, au sens où la somme indépendante de ces six termes majore la loi de $\|\widehat{\mu}_{t_1} - \mu_{t_1}\|_W$.

3.2 Majoration des termes déterministes

Les majorations de (11), (12) et (13) ne présentent pas de difficulté. Déjà, comme on sait que $\bar{\mathcal{S}}$ est $\kappa/2$ -lipschitzienne, on a :

$$(11) \leq e^{t_1 \kappa/2} W_0. \quad (20)$$

Pour majorer (12), on utilise qu'on travaille en fait dans $\mathcal{M}_1^+(X)$, ce qui permet d'utiliser la distance de la variation totale $\|\cdot - \cdot\|_{\text{VT}}$ entre deux mesures.

On remarque alors, d'une part que pour $\mu \in \mathcal{M}_1^+(X)$ on a $\|\bar{S}(\mu)\|_{\text{VT}} \leq 2\bar{\tau}$, d'autre part que \bar{S} est $2\bar{\tau}D$ -lipschitzienne de $(\mathcal{M}_1^+(X), \|\cdot\|_{\text{VT}})$ dans $(\bar{\mathcal{M}}_0(X), \|\cdot\|_W)$, d'où on déduit qu'on a :

$$(12) \leq 2\bar{\tau}^2 D t_1^2. \quad (21)$$

Enfin, le terme d'autochoc se majore trivialement par :

$$(13) \leq \frac{D\bar{\tau}t_1}{N}. \quad (22)$$

3.3 Majoration du terme de concentration

Déjà, on va comparer \mathcal{S}^* à $\mathcal{S} = \frac{\mathfrak{M}}{M}\bar{\mathcal{S}}$ plutôt qu'à $\bar{\mathcal{S}}$ lui-même, ce qui ne nous coûtera qu'un terme $2D/N$ dans (16.)

La mesure \mathfrak{S} est portée par les $\delta_{(y,y')} - \delta_{(x,x')}$, que nous noterons ici $\sigma(x, x', y, y')$. Or nous avons le lemme évident suivant :

Lemme 3.3. *Soient μ et ν deux mesures sur X^4 de même masse totale ; on a :*

$$\left\| \int_{X^4} \sigma(x, x', y, y') d\mu(x, x', y, y') - \int_{X^4} \sigma(x, x', y, y') d\nu(x, x', y, y') \right\|_W \leq \|\pi_1 \circ \nu - \pi_1 \circ \mu\|_W + \|\pi_2 \circ \nu - \pi_2 \circ \mu\|_W + \|\pi_3 \circ \nu - \pi_3 \circ \mu\|_W + \|\pi_4 \circ \nu - \pi_4 \circ \mu\|_W. \quad (23)$$

On se ramène ainsi à devoir majorer, pour μ une mesure de probabilité sur X , la distance de Monge entre cette mesure et la mesure empirique de \mathfrak{M} points de X i.i.d. selon μ .

Pour ce faire, j'ai mis au point dans un autre travail [3] le théorème suivant :

Théorème 3.4. *Si μ est une mesure de probabilité sur X , pour $\mathfrak{M} \geq 1$ et $\hat{\mu}^{\mathfrak{M}}$ la mesure empirique à \mathfrak{M} points associée à μ , on a pour tout $\lambda > 0$:*

$$E_\lambda [\|\hat{\mu}^{\mathfrak{M}} - \mu\|] \leq \alpha\Phi\left(\frac{\lambda D'}{\mathfrak{M}^{1/2}}\right) \frac{D'}{\mathfrak{M}^{1/n}}, \quad (24)$$

où D' est une longueur caractéristique de X .

En outre, on a le lemme suivant pour majorer la somme de plusieurs variables aléatoires non indépendantes :

Lemme 3.5. *Pour des variables aléatoires réelles X_1, \dots, X_I , et des p_i vérifiant, $p_1^{-1} + \dots + p_I^{-1} = 1$, on a pour tout $\lambda > 0$:*

$$E_\lambda \left[\sum_{i=1}^I X_i \right] \leq \sum_{i=1}^I E_{p_i \lambda} [X_i]. \quad (25)$$

Finalement, on récupère la majoration suivante de (16) :

$$(16) \leq \frac{2D}{N} + \alpha \Phi \left(\frac{\lambda D'(\bar{\tau}t_1)^{1/2}}{N^{1/2}} \right) D'(\bar{\tau}t_1)^{1-1/n} N^{-1/n}. \quad (26)$$

3.4 Majoration de la courbure stochastique

Soient $Z \in \mathbb{N}$, $\varepsilon \geq 0$; nous voulons étudier la loi $\Upsilon(\varepsilon, Z)$. On a :

$$\mathbb{E}[e^{\lambda \Upsilon}] = \prod_{i=1}^Z \left(1 + (i-1)(e^\lambda - 1)\varepsilon \right) \leq \prod_{i=1}^Z e^{(i-1)(e^\lambda - 1)\varepsilon} = e^{Z(Z-1)(e^\lambda - 1)\varepsilon/2}, \quad (27)$$

soit

$$E_\lambda[\Upsilon(\varepsilon, Z)] \leq \Phi(\lambda) \frac{Z^2 \varepsilon}{2}. \quad (28)$$

On obtient donc :

$$E_\lambda[(17)] \leq \alpha \Phi \left(\frac{\lambda D}{N} \right) (\bar{\tau}t_1)^2 D. \quad (29)$$

3.5 Majoration de l'erreur poissonnienne

Lemme 3.6. *Pour tout $\theta \geq 0$, pour tout $\lambda \geq 0$:*

$$E_\lambda[|\mathcal{P}(\theta) - \theta|] \leq \Phi(\lambda) \left\{ \frac{\lambda}{2} \theta + \alpha \theta^{1/2} \right\}. \quad (30)$$

Démonstration. Voir annexe B. □

En écrivant que $|\mathcal{P}(M) - \mathfrak{M}| \leq |\mathcal{P}(M) - \mathfrak{M}| + 1$, on obtient finalement que

$$(18) \leq \frac{2D}{N} + \alpha \Phi \left(\frac{\lambda D}{N} \right) \left\{ \frac{\bar{\tau}t_1 \lambda D^2}{N} + \frac{(\bar{\tau}t_1)^{1/2} D}{N^{1/2}} \right\}. \quad (31)$$

4 Optimisation du pas de temps

4.1 Principe

Supposons ici N fixé, moralement grand. Sur un seul pas de temps t_1 , nous avons pu majorer la façon dont la mesure empirique s'écarte de la mesure hydrodynamique par une somme de plusieurs termes dans laquelle t_1 intervient de diverses façons. Plus précisément, dans le cas $\lambda = 0$, nous avons obtenu une expression de la forme :

$$\mathbb{E}[W_{t_1}] \leq e^{t_1 \kappa/2} W_0 + f(t_1). \quad (32)$$

Notre objectif est d'étudier l'évolution du système sur un temps T macroscopique. Supposons pour simplifier que T soit un multiple de t_1 , alors en itérant la relation (32) T/t_1 fois, on obtient :

$$\mathbb{E}[W_T] \leq e^{T\kappa/2}W_0 + f(t_1) \sum_{j=0}^{\frac{T}{t_1}-1} e^{jt_1\kappa/2}. \quad (33)$$

Comme t_1 est destiné à être petit devant les données du problème, on peut majorer sans dégâts importants la somme par une intégrale pour trouver finalement :

$$\mathbb{E}[W_T] \leq e^{T\kappa/2}W_0 + \frac{e^{T\kappa/2} - 1}{\kappa/2} \frac{f(t_1)}{t_1}. \quad (34)$$

Cela nous dit comment optimiser t_1 pour avoir la meilleure majoration possible : il faut chercher à minimiser $f(t_1)/t_1$. On notera que le choix de t_1 ne dépend ni de T ni de W_0 , mais simplement de N .

4.2 Détermination

D'après le travail du §3, $f(t_1)$ s'exprime comme une somme de différents termes dont les exposants en t_1 varient entre 0 et 2, ce qui donnera des exposants entre -1 et 1 pour $f(t_1)/t_1$: il importe donc de ne choisir t_1 ni trop petit ni trop grand. Pour quelle valeur de t_1 l'expression est-elle optimale ? Avec un peu de sens physique, on devine que certains termes sont voués à être négligeables. Ainsi, le terme d'autochoc (13) ne devrait vraisemblablement pas jouer beaucoup, pas plus que l'erreur poissonnienne (18). De même, dans la majoration (26), le terme $2D/N$ est secondaire. Bref, pour des valeurs « raisonnables » des paramètres, on doit avoir :

$$\frac{f(t_1)}{t_1} \lesssim \alpha D' \left(\bar{\tau}^2 t_1 + \frac{\bar{\tau}^{1-1/n}}{(Nt_1)^{1/n}} \right), \quad (35)$$

où nous avons éventuellement remplacé D' par $D' \vee D$, et cette valeur atteint son minimum pour

$$t_1 = \frac{\alpha}{\bar{\tau} N^{1/(n+1)}}. \quad (36)$$

4.3 Majoration sur un pas optimal

On a alors :

$$\begin{aligned} E_\lambda[W_{t_1}] &\leq e^{t_1\kappa/2}W_0 + \left(1 + \Phi\left(\frac{\lambda D'}{N^{(n+2)/2(n+1)}}\right) + \Phi\left(\frac{\lambda D'}{N}\right) \right) \frac{\alpha D'}{N^{2/(n+1)}} \\ &+ \alpha \left\{ \frac{D'}{N^{(n+2)/(n+1)}} + \Phi\left(\frac{\lambda D'}{N}\right) \frac{\lambda D'^2}{N^{(n+2)/(n+1)}} + \Phi\left(\frac{\lambda D'}{N}\right) \frac{D'}{N^{(n+2)/2(n+1)}} \right\}. \quad (37) \end{aligned}$$

Cette formule nous dit notamment que, quand le nombre de particules augmente, l'espérance de W_T va décroître en $1/N^{1/(n+1)}$. Par conséquent, pour étudier les moyennes déviations de W_T , le λ pertinent devra croître plus vite que $N^{1/(n+1)}$.

☛ *Dorénavant, nous prendrons toujours λ de la forme $\beta N^\gamma/D'$, où l'exposant γ est égal à $1/(n+1)+\varepsilon$, ε étant choisi très petit — les calculs montrent qu'il suffit de choisir $\gamma < 1/2$; on peut alors fixer γ arbitrairement dans $] \frac{1}{n+1}, \frac{1}{2} [$.*

Ce choix étant fait pour λ , (37) se résume sous la forme :

$$E_{\frac{\beta N^\gamma}{D'}}[W_{t_1}] \leq e^{t_1 \kappa/2} W_0 + \alpha \Phi_\beta \left(\frac{1}{N} \right) \frac{D'}{N^{2/(n+1)}}, \quad (38)$$

où $(\beta, x) \mapsto \Phi_\beta(x)$ est une fonction purement mathématique, qui, à $\beta \geq 0$ fixé, est croissante sur $[0, +\infty[$ et continue en 0 avec $\Phi_\beta(0) = 1$, et qui, à x fixé, est croissante en β . C'est sous cette forme qu'on utilisera la majoration (37) par la suite.

5 Vers un résultat global

5.1 Enchaînement des pas

Supposons ici N fixé ; nous allons majorer $\mathbb{E}[W_T]$.

Déjà, grâce au théorème 3.4 on a :

$$\mathbb{E}_{\frac{\beta N^\gamma}{D'}}[W_0] \leq \alpha \Phi \left(\frac{\beta}{N^{1/2-\gamma}} \right) \frac{D'}{N^{1/n}}. \quad (39)$$

Supposons dans un premier temps $T = it_1$ pour $i \in \mathbb{N}$. Comme le comportement du système microscopique sur les différents intervalles $[it_1, (i+1)t_1]$ est une chaîne de Markov, la formule (38) donne par récurrence sur i :

$$E_{\frac{\beta N^\gamma}{D'}}[W_T] \leq \alpha \left\{ \Phi \left(\frac{e^{\kappa T/2} \beta}{N^{1/2-\gamma}} \right) \frac{e^{\kappa T/2}}{N^{1/n}} + \frac{e^{\kappa T/2} - 1}{\kappa/2\bar{\tau}} \Phi_{e^{\kappa T/2} \beta} \left(\frac{1}{N} \right) \frac{D'}{N^{1/(n+1)}} \right\}. \quad (40)$$

5.2 Supremum

☛ *Dans ce paragraphe, on se donne un $T \geq 0$ et on choisit arbitrairement la valeur de β . On notera $\Psi(x)$ toute fonction dépendant implicitement de ces paramètres, croissante sur $[0, +\infty[$ et continue en 0 avec $\Psi(0) = 1$.*

À partir de (40), on obtient que pour tout $t \leq T$, on a :

$$E_{\frac{\beta N^\gamma}{D'}}[W_t] \leq \alpha \Psi \left(\frac{1}{N} \right) \frac{e^{\kappa t/2} - 1}{\kappa/2\bar{\tau}} \frac{D'}{N^{1/(n+1)}}. \quad (41)$$

Posons maintenant $C = \alpha D' / \bar{\tau}$ pour la valeur de α apparaissant dans (41), et fixons-nous un $C' > \frac{e^{\kappa T/2} - 1}{\kappa/2\bar{\tau}} C$. Soit t_2 un diviseur de t_1 avec $t_1/t_2 = A$; nous allons majorer la probabilité que le supremum de W_t atteigne ou dépasse C' dans l'intervalle de temps $[(i-1)t_2, it_2]$ pour $i \in \{1, \dots, A\}$.

Pour ce faire, on considère le temps d'arrêt défini comme le premier instant après $(i-1)t_2$ où W_t atteint la valeur $C'/N^{1/(n+1)}$, et on regarde ce qui se passe conditionnellement à ce que ce temps d'arrêt soit $\leq it_2$, événement que nous noterons \mathcal{Z} . Si, à compter du temps d'arrêt, le système suivait une évolution hydrodynamique, on aurait $W_{it_2} \geq e^{-\kappa t_2/2} C'$ étant donné que \bar{S} est $\kappa/2$ -lipschitzienne. En fait, le système dévie légèrement de son régime hydrodynamique, mais pas trop puisqu'il a disposé d'un temps au plus t_2 pour le faire. Par conséquent, on a :

$$E_{\beta \frac{N\gamma}{D'}} [W_{it_2} | \mathcal{Z}] \geq e^{-\kappa t_2/2} C' - \alpha \Psi \left(\frac{1}{N} \right) \frac{e^{\kappa t_2/2} - 1}{\kappa/\bar{\tau}} \frac{D'}{N^{1/(n+1)}}, \quad (42)$$

où on a utilisé implicitement le

Lemme 5.1. *Pour toute variable aléatoire réelle Z , pour tout $\lambda \geq 0$, on a :*

$$\mathbb{E}[e^{-\lambda Z}] \geq \frac{1}{\mathbb{E}[e^{\lambda Z}]}. \quad (43)$$

On en tire alors :

$$\Psi \left(\frac{1}{N} \right) \frac{C}{N^{1/(n+1)}} \geq E_{\beta \frac{N\gamma}{D'}} [W_{t_2}] \geq \mathbb{P}(\mathcal{Z}) \left(e^{-\kappa t_2/2} C' - \alpha \Psi \left(\frac{1}{N} \right) \frac{e^{\kappa t_2/2} - 1}{\kappa/\bar{\tau}} D' \right) \frac{1}{N^{1/(n+1)}}. \quad (44)$$

Pour peu que nous ayons choisi t_2 suffisamment petit pour vérifier $e^{-\kappa t_2/2} C' - \alpha \frac{e^{\kappa t_2/2} - 1}{\kappa/\bar{\tau}} D' > C$, il s'ensuit que $\mathbb{P}(\mathcal{Z})$ décroît au moins aussi vite que $\exp(\text{cte} \cdot N^{\gamma-1/(n+1)})$, et en particulier plus vite que tous les $1/N^q$. En sommant sur tous les $i \in \{1, \dots, A\}$, cela est vrai aussi pour $\mathbb{P}(\sup_{t \in [0, T]} W_t \geq C'/N^{1/(n+1)})$, ce qui est l'énoncé du théorème 1.6.

A Dimension fractale dichotomique

Dans cet article, j'ai utilisé un concept un peu particulier de dimension fractale, que j'appellerai ci la *dimension dichotomique* :

Définition A.1. *Soit X un espace métrique borné. Pour $d > 0$ un réel, on dira que X est fortement de dimension d si, pour une certaine constante $D > 0$, on peut construire des parties $X_\emptyset, X_a, X_b, X_{aa}, \dots$ de X indexées par les mots finis de l'alphabet $\{a, b\}$ telles que :*

1. $X_\emptyset = X$;

2. Pour tout mot m , $X_m = X_{ma} \cup X_{mb}$;

3. Pour tout mot m de longueur ℓ , le diamètre de X_m est majoré par $D/2^{\ell/d}$.

On dira également que X est fortement de dimension 0 quand X est fini.

Avis A.2. Toute variété riemannienne compacte de dimension d est fortement de dimension d .

B Preuve du lemme 3.6

Soient $\theta \geq 0$, $\lambda \geq 0$; on considère une variable aléatoire Z de loi $\mathcal{P}(\theta) - \theta$.
On a :

$$\mathbb{E}[e^{\lambda|Z|}] = \mathbb{E}[e^{\lambda Z} \wedge e^{-\lambda Z}] \leq \mathbb{E}[e^{\lambda Z}] + \mathbb{E}[e^{-\lambda Z}] = e^{(e^\lambda - 1 - \lambda)\theta} + e^{(e^{-\lambda} - 1 + \lambda)\theta}, \quad (45)$$

d'où

$$E_\lambda[|Z|] \leq \frac{\ln 2}{\lambda} + \frac{e^\lambda - 1 - \lambda}{\lambda} \theta. \quad (46)$$

Cette majoration est mauvaise quand λ est petit, mais on sait que $E_\lambda[|Z|]$ est croissant en λ . Par conséquent, en utilisant la valeur de (46) pour $\lambda = (2 \ln 2 / \theta)^{1/2}$, on trouve que pour tout $\lambda \geq 0$:

$$E_\lambda[|Z|] \leq \Phi\left(\frac{1}{\theta}\right) \sqrt{2 \ln 2} \theta^{1/2} + \frac{e^\lambda - 1 - \lambda}{\lambda} \theta. \quad (47)$$

Reste à s'affranchir du $\Phi(1/\theta)$ qui nous empoisonne l'existence. Pour ce faire, on observe qu'on a facilement pour tous $\theta \geq 0$, $\lambda \geq 0$:

$$E_\lambda[|Z|] \leq (1 + e^\lambda) \theta. \quad (48)$$

En séparant les cas $\theta < 1$ et $\theta \geq 1$, on a :

$$\Phi\left(\frac{1}{\theta}\right) \sqrt{2 \ln 2} \theta^{1/2} \wedge (1 + e^\lambda) \theta \leq \Phi(\lambda) \theta^{1/2}, \quad (49)$$

d'où finalement (30).

Références

- [1] F. BOLLEY, A. GUILLIN et C. VILLANI : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Theory and Related Fields*, 137:541–593, 2007.
- [2] H. SPOHN : *Large scale dynamics of interacting particles*. Texts and monographs in Physics. Springer, 1991.
- [3] R. PEYRE : Moyennes déviations des mesures empiriques. Document personnel, 2007.

Document XI

Grand nombre de particules en interaction : la loi de Fourier et l'équation de Boltzmann en champ moyen

Grand nombre de particules en interaction : la loi de Fourier et l'équation de Boltzmann en champ moyen

Rémi Peyre

12 octobre 2007

Résumé

Le présent document présente les questions auxquelles je me suis intéressé depuis le début de ma thèse ; les avancées que j'y ait faites et les problèmes qui me restent à résoudre. Dans une première partie, je parlerai des défis liés à la dérivation de la loi de Fourier à partir d'un modèle microscopique ; dans une seconde, j'aborderai la limite de champ moyen pour des modèles de particules à chocs de type Boltzmann.

Introduction

Depuis la découverte de l'atome (Dalton 1803), nous savons que la matière est constituée, à l'échelle atomique, de particules discrètes évoluant selon des lois élémentaires. Cependant, à l'échelle humaine la matière apparaît au contraire comme *continue*, et son comportement est décrit par d'autres lois exprimées en termes de *champs* (ex : loi de Navier–Stokes, 1822). Comment passe-t-on du cas discret au cas continu ? C'était, en substance, la question soulevée par Hilbert en 1900 dans son sixième problème [1].

Le champ de recherches lié aux questions du comportement-limite d'un grand nombre de particules en interaction est immense. Dans ce document, je voudrais présenter deux thèmes de ce domaine sur lesquels j'ai travaillé depuis le début de ma thèse :

- La **loi de Fourier** (§ 1) donne le comportement du flux de chaleur dans un solide inhomogène ; c'est sans doute l'exemple le plus simple de système déterministe *réversible* dont la limite hydrodynamique soit *irréversible*. J'expliquerai pourquoi la dérivation de la loi de Fourier est certainement difficile, puis je donnerai les arguments heuristiques des physiciens pour prédire la valeur de la conductivité thermique (formule de Green–Kubo). Il semble difficile de trouver de nouveaux résultats d'envergure sur la loi de Fourier, et mes recherches en ce sens ont été infructueuses ; néanmoins les questions soulevées par ce problème

sont liées au cadre plus général de l'étude des décorrélatons spatio-temporelles des systèmes de physique statistique, sur laquelle je présenterai quelques résultats et pistes de recherche.

- Les équations de champ moyen traduisent le comportement d'un grand nombre de particules en l'absence de structure spatiale ; en particulier, **l'équation de Boltzmann en champ moyen** donne le comportement cinétique des particules d'un gaz évoluant par chocs. Dans l'étude de la limite de champ moyen, peu de résultats quantitatifs ont jusque-là été donnés ; j'expliquerai comment j'espère aller plus loin dans cette voie-là, d'abord en termes de distances de couplage, puis au sens de la théorie UCLT.

☛ *Afin de permettre au lecteur de distinguer plus facilement ce qui, dans cette présentation, tient de l'exposition des travaux d'autres auteurs de ce qui relève de mes recherches personnelles, les passages résultant de mes propres travaux ont été imprimés en fontes Times.*

Table des matières

1	Loi de Fourier	226
1.1	Introduction	226
1.1.1	La loi de Fourier en physique	226
1.1.2	Le modèle mathématique	227
1.1.3	Phénomènes de conductivité anormale	229
1.2	La formule de Green–Kubo	230
1.2.1	Se ramener à l'équilibre thermique	230
1.2.2	Fluctuations-dissipation	231
1.3	Décorrélatons spatiales	233
1.3.1	Décorrélatons uniformes en domaine	233
1.4	Décorrélatons temporelles	235
1.4.1	Bruits	235
1.4.2	Hypocoercivité	236
1.4.3	Pistes de recherche	237
2	Équation de Boltzmann en champ moyen	238
2.1	Motivations	238
2.1.1	Équations de champ moyen	238
2.1.2	Limite de champ moyen	239
2.2	Les distances de couplage	240
2.2.1	Contrôle des grandes déviations par couplage	240
2.2.2	Modèles à chocs	243
2.2.3	Limites de l'approche en distance W_1	244
2.3	Théorèmes-limite centraux uniformes	244
2.3.1	Problématique	244
2.3.2	La théorie UCLT	245

2.3.3	Espaces de Sobolev	248
2.3.4	Programme de recherches	249

1 Loi de Fourier

1.1 Introduction

1.1.1 La loi de Fourier en physique

La loi de Fourier, établie expérimentalement par Fourier en 1822, est une des lois phénoménologiques les plus simples qui soient. Elle stipule que, dans un solide initialement porté à une température hétérogène, il s'établit un flux de chaleur linéaire en le gradient de température, qui tend à homogénéiser cette dernière. Concrètement, considérons un solide homogène dont nous notons $\theta(\cdot)$ la température locale⁽¹⁾. Il va régner dans ce solide un flux d'énergie \vec{j} (mesurée en $\text{J.m}^{-2}.\text{s}^{-1}$) déterminé localement par

$$\vec{j} = -\kappa(\theta)\vec{\nabla}\theta, \quad (1.1)$$

où $\kappa(\theta)$ est le *tenseur de conductivité thermique*⁽²⁾, spécifique au solide considéré. D'après le second principe de la thermodynamique, κ est un tenseur (strictement) positif; nous verrons par ailleurs que κ est toujours symétrique.

En fait, le physicien n'observe pas \vec{j} directement, mais simplement son effet sur θ : si e note la densité d'énergie (en J.m^{-3}), on a

$$\partial_t e = -\text{div } \vec{j}, \quad (1.2)$$

et comme d'autre part e est liée à θ par la relation

$$\frac{de}{d\theta} = c(\theta), \quad (1.3)$$

où $c(\theta)$ est une grandeur (strictement) positive, en m^{-3} , appelée *capacité thermique (volumique)*⁽³⁾, l'équation qui régit θ est finalement

$$\partial_t \theta = \frac{1}{c(\theta)} \text{div } (\kappa(\theta)\vec{\nabla}\theta). \quad (1.4)$$

On observera que dans cette équation, seule intervient au final la partie symétrique de κ . Autrement dit, la partie antisymétrique de κ n'a aucun sens expérimental, quoique, nous l'avons déjà dit, elle soit en réalité toujours nulle.

⁽¹⁾Les physiciens mesurent la température en Kelvin, mais nous multiplierons par la constante de Boltzmann $k_B = 1,38 \cdot 10^{-23} \text{J.K}^{-1}$ et exprimerons la température en Joules.

⁽²⁾Quand le solide est isotrope, κ est scalaire.

⁽³⁾Remarquer que (1.3) est au départ une relation de thermodynamique à l'équilibre, bien qu'elle reste vraie dans le cas que nous étudions.

Remarque 1.1. Quand le solide n'est pas homogène, κ ne dépend pas seulement de la température mais aussi de la position x . L'équation (1.4) devient alors

$$\partial_t \theta = \frac{1}{c(x, \theta)} \operatorname{div} (\kappa(x, \theta) \vec{\nabla} \theta). \quad (1.5)$$

Cela n'est pas si évident qu'il y paraît, car le résultat deviendrait faux si on avait « ajouté les x » à partir d'une autre forme de (1.4) : l'écriture (1.4) passe bien au cas non homogène seulement parce qu'elle est écrite avec les grandeurs qui restent continues aux interfaces.

Ce n'est qu'en 1905 que les physiciens ont compris la signification de la chaleur à l'échelle microscopique : notre solide est en fait constitué d'un grand nombre d'atomes en interaction, et la température du matériau indique avec quelle intensité ils vibrent autour de leurs positions d'équilibre [2]. Rappelons en effet que la température θ du matériau est le paramètre, qui croît avec l'énergie volumique du matériau, tel qu'à l'équilibre la probabilité d'observer le système dans une région $d\eta$ de l'espace des phases soit proportionnelle à $\exp(-H(\eta)/\theta)d\eta$, où $H(\eta)$ est l'énergie de l'état η . θ est également l'énergie cinétique moyenne des atomes du système divisée par $d/2$.

On aimerait bien pouvoir prédire le comportement global du système hors équilibre à partir de ses caractéristiques constitutives. Or à l'heure actuelle, on sait faire très peu de choses à ce sujet : on comprend assez bien les systèmes thermodynamiques à l'équilibre, mais les systèmes hors équilibre comme ceux qui nous préoccupent soulèvent nombre de questions sans réponse. J'ai donc tenté d'apporter quelques éclaircissements au « problème de la conduction thermique hydrodynamique ».

1.1.2 Le modèle mathématique

Présentons maintenant le modèle que nous utiliserons par la suite. On travaille en dimension $d \geq 1$; notre solide est un solide cristallin périodique sur le réseau \mathbb{Z}^d , la position du nœud \vec{z} étant notée $X_{\vec{z}}$. Dans chaque maille du réseau se trouvent un nombre fini n d'atomes de masses respectives m_1, \dots, m_n , dont les positions par rapport au nœud correspondant sont notées $\vec{x}_{\vec{z},1}, \dots, \vec{x}_{\vec{z},n}$. On note également $\vec{v}_{\vec{z},i} = \dot{\vec{x}}_{\vec{z},i}$ la quantité de mouvement de l'atome (\vec{z}, i) .

Deux mailles \vec{z} et \vec{z}' sont dites *voisines* lorsque $\|\vec{z}' - \vec{z}\|_\infty = 1$, et seuls les atomes d'une même maille ou de deux mailles voisines interagissent : pour $\|\vec{\eta}\|_\infty \leq 1$, $i, j \in \{1 \dots, n\}$, l'interaction entre l'atome (\vec{z}, i) et l'atome $(\vec{z} + \vec{\eta}, j)$ est décrite par le *potentiel d'interaction* $W_{\vec{\eta},i,j}(\vec{x}_{\vec{z}+\vec{\eta},j} - \vec{x}_{\vec{z},i})$. En outre, chaque atome est « attaché » à sa maille par un *potentiel de rappel* $V_i(\vec{x}_{\vec{z},i})$. Finalement, l'évolution du système est décrite par le hamiltonien (infini) dont l'expression formelle est

$$H = \sum_{\vec{z},i} \left(\frac{m_i}{2} |\vec{v}_{\vec{z},i}|^2 + V_i(\vec{x}_{\vec{z},i}) + \sum_{\vec{\eta},j} W_{\vec{\eta},i,j}(\vec{x}_{\vec{z}+\vec{\eta},j} - \vec{x}_{\vec{z},i}) \right) \quad (4). \quad (1.6)$$

Le modèle exposé ci-dessus possède toutes les caractéristiques intéressantes pour étudier la conduction de la chaleur d'un point de vue mathématique. Quelques commentaires s'imposent à propos de la modélisation physique :

1. On a choisi un solide cristallin car c'est la seule façon mathématique commode de décrire un solide.
2. Ici les potentiels d'interaction sont à portée finie⁽⁵⁾, ce qui est bien utile pour la théorie ; dans la réalité, les potentiels d'interaction sont à portée infinie avec une décroissance plus ou moins rapide. Il n'est pas clair qu'il suffise d'approcher un potentiel décroissant par un potentiel à portée finie pour récupérer toutes les propriétés du système, mais la problématique de la conduction de la chaleur est de toutes façons la même dans les deux cas.
3. Les potentiels de rappel V que nous avons instaurés ne correspondent à aucune réalité physique. Cependant, leur présence est primordiale dans notre modèle pour :
 - ne pas prendre en compte les déformations du solide, et assurer qu'un atome reste toujours sur la même maille ;
 - supprimer l'invariance par translation, ce qui est bien utile pour décrire l'équilibre thermique ;
 - éviter les phénomènes de conduction « anormale » de la chaleur : il a en effet montré qu'en dimension 1 et 2, en l'absence de potentiels de rappel, la conductivité thermique d'un solide tend vers l'infini quand sa taille augmente, mais que cet effet est inhibé par l'introduction d'un rappel [3].

Remarque 1.2. Pour traiter le problème d'un point de vue mathématique, il sera nécessaire d'imposer certaines conditions sur les potentiels V et W , par exemple, qu'ils soient suffisamment réguliers, que V soit confinant à l'infini, et que l'effet de V domine celui de W à grande distance.

Grâce au modèle présenté ci-dessus, on peut donner un sens précis aux différents concepts physiques. Ainsi, on peut associer à chaque atome (\vec{z}, i) son énergie individuelle par

$$E(\vec{z}, i) = \frac{m_i}{2} |\vec{v}_{\vec{z}, i}|^2 + V_i(\vec{x}_{\vec{z}, i}) + \sum_{\vec{\eta}, j} W_{\vec{\eta}, i, j} (\vec{x}_{\vec{z} + \vec{\eta}, j} - \vec{x}_{\vec{z}, i}), \quad (1.7)$$

de sorte que l'énergie totale soit la somme des énergies individuelles de chaque atome. S'il est possible de définir une échelle mésoscopique⁽⁶⁾, on

⁽⁴⁾Avec ces notations, l'interaction entre deux atomes est comptée deux fois, mais il suffit pour y remédier de diviser le potentiel W par deux.

⁽⁵⁾Non que les W soient à support compact, mais au sens où seules les interactions entre mailles voisines sont prises en compte.

⁽⁶⁾Rappel : une dimension mésoscopique est une grandeur très grande devant les caractéristiques microscopiques du système (taille de la maille en l'occurrence), mais très petite devant ses caractéristiques macroscopiques (taille totale du système).

peut alors définir une *densité d'énergie locale* e , d'où une *température locale* θ ⁽⁷⁾. Surtout, on peut donner un sens rigoureux à la densité de flux d'énergie. En effet, celle-ci était jusque-là défini comme une grandeur \vec{j} abstraite vérifiant (1.2). Or, cela est vrai si on pose que le flux d'énergie qui circule entre les atomes (\vec{z}, i) et (\vec{z}', j) (avec $\vec{z} \neq \vec{z}'$) est une mesure vectorielle uniformément concentrée sur le segment $[X_{\vec{z}} X_{\vec{z}'}]$, et dont la masse totale est $\left((\vec{v}_{\vec{z}', j} - \vec{v}_{\vec{z}, i}) \cdot \vec{F}_{(\vec{z}, i) \rightarrow (\vec{z}', j)} \right) \overline{X_{\vec{z}} X_{\vec{z}'}}$, où $\vec{F}_{(\vec{z}, i) \rightarrow (\vec{z}', j)}$ est la *force exercée par* (\vec{z}, i) *sur* (\vec{z}', j) , définie par⁽⁸⁾

$$\vec{F}_{(\vec{z}, i) \rightarrow (\vec{z}', j)} = -\vec{\nabla} W_{\vec{z}' - \vec{z}, i, j}(x_{\vec{z}', j} - x_{\vec{z}, i}) + \vec{\nabla} W_{\vec{z} - \vec{z}', j, i}(x_{\vec{z}, i} - x_{\vec{z}', j}). \quad (1.8)$$

Dorénavant donc, le flux d'énergie sera défini comme ci-dessus, la densité de flux étant sa moyenne volumique locale.

1.1.3 Phénomènes de conductivité anormale

Comment se fait-il que si peu de choses soient comprises sur la conduction de la chaleur? Prenons le plus simple des modèles : le solide harmonique. Dans ce modèle, $n = 1$, $V(\vec{x}) = \alpha |\vec{x}|^2/2$, $W_\eta(\vec{\Delta}x) = \beta_\eta |\vec{\Delta}x|^2/2$. L'évolution du système étant linéaire, elle se décompose en plusieurs modes indépendants : si nous supposons que le volume V du système est très grand, nous observons ainsi un quasi-continuum de modes harmoniques avec une densité $V\rho(\vec{k})$ autour du nombre d'ondes \vec{k} ⁽⁹⁾⁽¹⁰⁾. À l'équilibre thermique à la température θ , chaque mode contient (en moyenne) une énergie $\theta/2$, et surtout *transporte* cette énergie à une vitesse $\vec{v}(\vec{k})$ (génériquement) non nulle : on appelle un tel mode un *phonon*, car c'est l'analogie acoustique du photon. On observe ainsi que le transport de la chaleur s'opère comme un flux de particules *libres*, par conséquent on s'attend à ce que la conductance thermique globale d'un tel solide soit proportionnelle à sa surface, ou, en d'autres termes, que sa conductivité « explose » proportionnellement à sa longueur, ce qui constitue une violation évidente de la loi de Fourier ! C'est effectivement ce qu'ont montré Lieb *et al.* en 1967 [4].

Ainsi l'anharmonicité du solide doit être utilisée de façon fondamentale pour obtenir la loi de Fourier. Mais notons que la simple anharmonicité ne suffit pas : par exemple, si nous considérons en dimension 1 un modèle (sans

⁽⁷⁾La notion de température locale n'est pertinente que si le système est à l'équilibre local, i.e. si le comportement du système à l'échelle mésoscopique est celui d'un système infini à l'équilibre aux paramètres thermodynamiques locaux.

⁽⁸⁾On a deux termes dans la formule, mais ce n'est qu'un artefact de notre formalisme, cf. note (4).

⁽⁹⁾Rappelons au passage que pour un système périodique, l'espace de nombre d'ondes est un tore : c'est \mathbb{R}^d quotienté par le réseau réciproque.

⁽¹⁰⁾Si la densité atomique du solide est ρ_{at} , le nombre de modes étant égal à la dimension du système, cela impose $\int_{\vec{k}} \rho(\vec{k}) dk = 2d\rho_{\text{at}}$.

potentiel de rappel) de « sphères » dures de diamètre nul, le système est équivalent à des particules libres, et conduit au même phénomène de conductivité anormale que le cas harmonique⁽¹¹⁾.

1.2 La formule de Green–Kubo

1.2.1 Se ramener à l'équilibre thermique

La formule de Green–Kubo est une formule heuristique qui permet d'exprimer la conductivité thermique sans avoir à considérer de système hors-équilibre, puisqu'elle se calcule à partir des seules données de l'équilibre. Nous allons présenter ici une dérivation possible⁽¹²⁾ de cette formule.

On considère un système hors-équilibre autour de la température θ_0 , grand mais fini, soumis à un faible gradient de température $\vec{\nabla}\theta$ uniforme. Notant $\beta = \theta^{-1}$ la température inverse, cela correspond à un gradient de β de $\vec{\nabla}\beta = -\vec{\nabla}\theta/\theta^2$. On suppose que le système est au départ dans un équilibre local, de sorte que la probabilité pour un état η du système d'être observée est proportionnelle à

$$\exp\left(-\sum_{\vec{z},i}(\beta_0 + \vec{\nabla}\beta \cdot \vec{x}_{\vec{z}})E(\vec{z},i)\right)d\eta. \quad (1.9)$$

Notons \vec{J} le courant d'énergie total. Naïvement, nous voudrions calculer la valeur moyenne de \vec{J} sous la loi que nous venons de définir, mais les symétries du problème font qu'on trouverait... $\vec{0}$! Quand on y regarde de plus près, on s'aperçoit qu'en fait la conduction de chaleur est un phénomène qui répond avec un certain retard, quoique très faible, au gradient de température qui la provoque. Il nous faut donc calculer le transfert de chaleur sur une durée *mésoscopique* pour trouver une grandeur pertinente.

En introduisant le *moment énergétique* \vec{E} , défini par

$$\vec{E} = \sum_{\vec{z},i} E(\vec{z},i)\vec{OX}_z, \quad (1.10)$$

où O est un point-origine de l'espace, (1.9) se réécrit comme

$$\exp(-\beta_0 H + \vec{\nabla}\beta \cdot \vec{E}), \quad (1.11)$$

où H est l'énergie totale du système. Par la suite, nous considérerons que H est constante : ce n'est pas tout-à-fait vrai, mais on peut s'y ramener en conditionnant.

⁽¹¹⁾Le problème n'est pas intrinsèquement lié à l'absence de potentiel de rappel comme on pourrait le croire d'après la page 228, n° 3, car on peut fabriquer le même genre d'exemple pathologique en toute dimension.

⁽¹²⁾Il en existe au moins trois : outre la présente version présentée dans [3], la version originale de Green & Kubo [5, 6, 7] passe par une technique de fluctuations–dissipation, et il existe également une approche par projection de l'évolution due à Zwanzig & Mori [8, 9]

Comme nous avons supposé que $\vec{\nabla}\theta$ était très petit, (1.11) est aussi $(1 + \vec{\nabla}\beta \cdot \vec{E})e^{-\beta_0 H}$. Si maintenant nous notons \mathbb{P}_0 la distribution de probabilité du système à l'équilibre à la température θ_0 , et Z_0 la fonction de partition associée :

$$Z_0 = \int e^{-\beta_0 H(\eta)} d\eta, \quad (1.12)$$

notant \mathbb{P}_1 la distribution de probabilité proportionnelle à (1.9), la fonction de partition associée est

$$Z_1 = \int e^{-\beta_0 H(\eta)} (1 + \vec{\nabla}\beta \cdot \vec{E}(\eta)) d\eta = Z_0 (1 + \vec{\nabla}\beta \cdot \langle \vec{E} \rangle_0), \quad (1.13)$$

où $\langle \cdot \rangle_0$ note l'espérance sous l'équilibre à θ_0 . En déplaçant l'origine de l'espace, on peut faire en sorte que $\langle \vec{E} \rangle_0 = 0$, d'où finalement

$$d\mathbb{P}_1 = (1 + \vec{\nabla}\beta \cdot \vec{E}) d\mathbb{P}_0. \quad (1.14)$$

Soit T une durée mésoscopique. Le flux total de chaleur entre l'instant 0 et l'instant T est

$$\int_0^T \vec{J}(t) dt = \int_0^T \dot{\vec{E}}(t) dt = \vec{E}(T) - \vec{E}(0), \quad (1.15)$$

donc son espérance sous \mathbb{P}_1 est

$$\left\langle (1 + \vec{\nabla}\beta \cdot \vec{E}(0)) (\vec{E}(T) - \vec{E}(0)) \right\rangle_0. \quad (1.16)$$

Comme nous l'avons dit, $\langle \vec{E}(0) \rangle_0 = \vec{0}$, en outre, $\langle \cdot \rangle_0$ est invariant par évolution, donc $\langle \vec{E}(T) \rangle_0 = \vec{0}$ et (1.16) est égal à

$$\frac{1}{\theta_0^2} \vec{\nabla}\theta \cdot \left\langle \vec{E}(0) \otimes (\vec{E}(0) - \vec{E}(T)) \right\rangle_0, \quad (1.17)$$

d'où

$$\kappa = \frac{1}{V\theta_0^2} \left\langle \vec{E}(0) \otimes \frac{1}{T} (\vec{E}(0) - \vec{E}(T)) \right\rangle_0. \quad (1.18)$$

Ainsi, il nous faut comprendre comment \vec{E} évolue pour calculer κ .

1.2.2 Fluctuations-dissipation

Comment donc évolue \vec{E} ? (1.18) vient de nous apprendre que \vec{E} avait tendance à revenir vers $\vec{0}$, cependant nous savons qu' \vec{E} n'est pas identiquement nul : au mouvement de *dissipation* doit donc s'ajouter un mouvement de *fluctuation* qui le compense exactement à l'équilibre⁽¹³⁾.

Nous allons faire quelques hypothèses très naturelles :

⁽¹³⁾Ce principe est extrêmement général et constitue un des piliers de la théorie de la réponse linéaire. Le lecteur trouvera un énoncé général du théorème de fluctuation–dissipation dans [10].

1. Sur des échelles de temps mésoscopiques, l'évolution de \vec{E} est markovienne ;
2. La relaxation de \vec{E} vers l'équilibre est linéaire ;
3. Les fluctuations de \vec{E} sont un bruit blanc gaussien ;
4. Les phénomènes de fluctuation et de dissipation sont indépendants.

Commençons par étudier les fluctuations de \vec{E} . On rappelle que $\dot{\vec{E}} = \vec{J}$, lequel \vec{J} a plus de sens physique que \vec{E} . Sur un temps mésoscopique, les fluctuations prédominent sur la relaxation et \vec{J} apparaît donc comme un bruit blanc temporel. En outre, la nature *locale* du hamiltonien (1.6) fait que les valeurs de la densité de flux d'énergie dans des zones suffisamment éloignées peuvent être considérées comme indépendantes, de sorte que \vec{J} peut être considéré comme un bruit blanc spatial. Ainsi, notant V le volume du système considéré, une grandeur comme $\langle \vec{J}(0) \otimes \vec{J}(0) \rangle_0 / V$ est-elle en fait indépendante de V — c'est simplement la covariance volumique de \vec{J} — ; on la considèrera donc comme une propriété de l'équilibre du système *infini*, bien que nous continuerons à utiliser la même notation.

Quelle est la covariance temporelle de \vec{J} ? C'est le tenseur C tel que, pour un temps mésoscopique T , on ait

$$\left\langle \left(\int_0^T \vec{J}(t) dt \right) \otimes \left(\int_0^T \vec{J}(t) dt \right) \right\rangle_0 = TC, \quad (1.19)$$

soit

$$C = \frac{1}{T} \left\langle \int_0^T \vec{J}(0) \otimes \left(\int_{-t}^{T-t} \vec{J}(u) du \right) dt \right\rangle_0 \quad (1.20)$$

par invariance de $\langle \cdot \rangle_0$ sous l'évolution. Comme \vec{J} se décorrèle rapidement, pour $t \neq 0, T$ on a pratiquement

$$\left\langle \vec{J}(0) \otimes \left(\int_{-t}^{T-t} \vec{J}(u) du \right) dt \right\rangle_0 = \left\langle \vec{J}(0) \otimes \left(\int_{-\infty}^{\infty} \vec{J}(u) du \right) dt \right\rangle_0, \quad (1.21)$$

et finalement

$$C = \int_{-\infty}^{\infty} \langle \vec{J}(0) \otimes \vec{J}(t) \rangle_0 dt. \quad (1.22)$$

Pour la relaxation, notons R le tenseur régissant la relaxation, de sorte qu'on ait en notation d'Itô, \vec{W} notant un mouvement brownien standard,

$$d\vec{E} = -R\vec{E}dt + C^{1/2}d\vec{W}_t. \quad (1.23)$$

Alors l'expression (1.18) se réécrit comme

$$\kappa = \frac{1}{V\theta_0^2} \langle \vec{E} \otimes (R\vec{E}) \rangle_0. \quad (1.24)$$

Observons que R est nécessairement symétrique : en effet, l'évolution du système sous $\langle \cdot \rangle_0$ est invariante par renversement du temps, ce qui fait que l'évolution (1.23) doit être *réversible*, ce qui équivaut à dire que R est symétrique.

Reste à traduire l'équilibre entre fluctuations et dissipation : d'après (1.23), on a

$$\left\langle \frac{d}{dt} (\vec{E} \otimes \vec{E}) \right\rangle_0 = -2 \left\langle \vec{E} \otimes (R\vec{E}) \right\rangle_0 + C. \quad (1.25)$$

Comme la moyenne d'une dérivée temporelle est nulle, il s'ensuit que

$$\left\langle \vec{E} \otimes (R\vec{E}) \right\rangle_0 = \frac{1}{2} C, \quad (1.26)$$

d'où finalement la formule de Green-Kubo :

« **Théorème** » 1.3.

$$\kappa(\theta_0) = \frac{1}{2\theta_0^2} \int_{-\infty}^{\infty} \left\langle \frac{\vec{J}(0) \otimes \vec{J}(t)}{V} \right\rangle_0 dt. \quad (1.27)$$

Corollaire 1.4. *Le tenseur κ est symétrique.*

La formule de Green-Kubo nous montre que l'étude de la conductivité thermique passe par celle des décorrélations spatio-temporelles du flux de chaleur \vec{J} dans le système à l'équilibre ; nous allons donc tenter d'établir des résultats sur celles-ci.

1.3 Décorrélations spatiales

1.3.1 Décorrélations uniformes en domaine

Quand le système que nous étudions est à l'équilibre, son état est décrit par la statistique de Boltzmann ; l'étude mathématique de tels systèmes est une branche en plein dynamisme de la physique statistique avec notamment les résultats récents sur le modèle d'Ising.

Pour diverses raisons techniques, nous n'allons traiter ici que le cas de modèles discrets. On considère ici le système suivant :

Modèle 1.5. *L'état du système est décrit par un ensemble de spins $\sigma_{\vec{r}}$ indexés par les points du réseau \mathbb{Z}^d , où chaque spin a ses valeurs dans un espace d'états Σ fini. Chaque spin a une énergie propre donnée par $h_0(\sigma)$ d'une part ; d'autre part, pour $\vec{r} \neq \vec{j}$ avec $\|\vec{r} - \vec{j}\|_\infty = 1$, le spin en \vec{r} interagit avec le spin en \vec{j} avec une énergie d'interaction donnée par $h_{\vec{j}-\vec{r}}(\sigma_{\vec{r}}, \sigma_{\vec{j}})$ où h_0 et les $h_{\vec{e}}$ pour $\|\vec{e}\|_\infty = 1$ sont des fonctions de Σ , resp. $\Sigma \times \Sigma$, dans \mathbb{R} .*

Notation 1.6. *Dans le modèle 1.5, pour $E \subset \mathbb{Z}^d$ on note σ_E l'état de l'ensemble des sites de E .*

Le résultat suivant est classique [11, 12] :

Théorème 1.7. *Le système du modèle 1.5 a une unique mesure de Gibbs pour la température θ dès lors que θ est suffisamment élevée ; on note alors celle-ci \mathbb{P}_θ .*

En outre, si θ est suffisamment élevée il y a décroissance exponentielle du système au sens suivant : pour toute partie $Z \subset \mathbb{Z}^d$, pour tous $x, y \notin Z$, on a presque-sûrement que

$$\frac{d\mathcal{L}oi(y|x, Z)}{d\mathcal{L}oi(y|Z)} \geq 1 - e^{-\lambda(\theta)\|y-x\|_\infty} \quad (1.28)$$

pour un $\lambda(\theta) > 0$. De plus, $\lambda(\theta)$ tend vers $+\infty$ quand $\theta \rightarrow +\infty$.

(1.28) est la formulation la plus classique qu'on entend par « il y a décroissance exponentielle des corrélations » ; Dobrushin et Shlosman [13] ont établi l'équivalence entre ce critère et une multitude d'autres.

La formule (1.28) nous donne, *stricto sensu*, la décroissance des corrélations à deux points. Si nous voulons en déduire des renseignements sur la décroissance au sens de la variation totale $|\cdot|_{\text{VT}}$ entre des sous-ensembles quelconques de points, nous obtenons des estimées qui font apparaître la taille de ces sous-ensembles, par exemple :

Théorème 1.8. *Si la formule (1.28) est satisfaite, alors pour tous ensembles finis de points $E, F \subset \mathbb{Z}^d$, notant $d(E, F)$ la distance entre E et F pour la distance $\|\cdot - \cdot\|_\infty$, on a*

$$\left| \mathcal{L}oi(\sigma_{E \cup F}) - \mathcal{L}oi(\sigma_E) \otimes \mathcal{L}oi(\sigma_F) \right|_{\text{VT}} \leq \#E \#F e^{-\lambda(\theta)d(E, F)}. \quad (1.29)$$

Il est facile de voir que le facteur $\#E \#F$ ne peut être retiré, comme le montre le contre-exemple suivant :

Exemple 1.9. Pour un modèle d'Ising sous-critique (i.e., à haute température), avec E et F deux hyperplans parallèles (non confondus) de \mathbb{Z}^d , on a

$$\left| \mathcal{L}oi(\sigma_{E \cup F}) - \mathcal{L}oi(\sigma_E) \otimes \mathcal{L}oi(\sigma_F) \right|_{\text{VT}} = 1, \quad (1.30)$$

soit la valeur maximale imaginable : il n'y a aucune décroissance au sens de la variation totale.

Dans le genre de cas qui nous intéresse, on cherche à étudier des décroissances au sens L^2 , ce qui est plus faible que les décroissances en variation totale ; par contre, on aimerait établir des résultats concernant de vastes zones, à cause de la dynamique du système qui fait dépendre l'état futur d'un point de l'état présent de ses voisins.

J'ai conjecturé le résultat suivant :

Question 1.10. Si θ est suffisamment élevée, existe-t-il une constante $\lambda'(\theta) > 0$ telle que, pour tous sous-ensembles $A, B \subset \mathbb{Z}^d$ disjoints, pour toutes fonctions f, g σ_A -, resp. σ_B -mesurables, dans $L^2(d\mathbb{P}_\theta)$, on ait

$$|\mathbb{E}[fg] - \mathbb{E}[f]\mathbb{E}[g]| \leq e^{-\lambda'(\theta)}? \quad (1.31)$$

Plusieurs arguments tendaient à indiquer que ce résultat était vrai ; néanmoins une consultation approfondie ⁽¹⁴⁾ ne m'a permis de trouver nulle part ce résultat (ou un résultat l'impliquant) dans la littérature.

Je me suis alors attelé au problème, et j'ai pu finalement y répondre par l'affirmative :

Théorème 1.11 (voir document VIII). La réponse à la question 1.10 est positive ; en outre, on a $\lambda'(\theta) \xrightarrow{\theta \rightarrow \infty} \infty$.

1.4 Décorrélations temporelles

1.4.1 Bruits

Le côté temporel des décorrélatons, essentiel pour étudier par exemple la finitude de l'intégrale de Green–Kubo, est une autre paire de manches ! Pour simplifier l'étude du problème, on ajoute généralement un bruit dans la dynamique du hamiltonien (1.6). Voici, par violence croissante, les différents types de bruit qu'on peut ajouter :

1. **Aucun bruit** : on conserve la dynamique hamiltonnienne telle quelle. Évidemment, c'est ce modèle-là qui est le but ultime d'un point de vue physique, mais mathématiquement il semble intraitable, car dans la mesure où le système est déterministe et réversible il n'y aura aucune décorrélation dans un sens général ; seules certaines observables (en l'occurrence, ce serait \vec{J}), peut-être, vont décorréler, et encore sait-on que ce ne sera pas le cas dans tous les modèles (cf. § 1.1.3).
2. **Bruit conservatif** : on ajoute à la dynamique un terme de bruit qui n'ajoute ni n'enlève aucune énergie au système, par exemple on ajoute aux forces s'exerçant sur une particule un bruit blanc (de Stratonovitch) orthogonal à sa vitesse. Même si ce modèle n'a pas de justification physique manifeste, il est très intéressant d'un point de vue mathématique car il donne des décorrélatons temporelles au système tout en conservant l'essence du problème de la conduction thermique hydrodynamique. De ce fait, il reste également hélas difficile à traiter, notamment parce que l'existence d'une grandeur conservative mène à une famille de mesures d'équilibre plutôt qu'à une seule.
3. **Bain de Langevin** : dans ce modèle, on ajoute une force $\vec{F}_{z,i}$ s'appliquant à l'atome (z, i) et valant

$$\vec{F}_{z,i} = \sigma \frac{d\vec{W}_{z,i}}{dt} - \frac{\sigma^2 m_i}{4d\theta} \vec{v}_{z,i}, \quad (1.32)$$

⁽¹⁴⁾Communications personnelles auprès de V. Beffara, W. Werner, T. Bodineau, S. Shlosman et B. Zegarlinski.

les $W_{z,i}$ notant des mouvements browniens standard indépendants, où θ est la température à laquelle on travaille. La seule différence avec le bruit conservatif est qu'ici le bruit blanc est isotrope et peut donc ajouter ou enlever de l'énergie au système, avec les avantages et les inconvénients dont nous venons de parler dans le n° 2.

4. **Dynamique de Glauber :** dans cette dynamique, on oublie la structure cinétique du système (i.e., la relation $\vec{v}_i = \dot{x}_i$) et on regarde le système comme une particule évoluant par une ÉDS *du premier ordre* dans un espace d'états de grande dimension avec le potentiel H . La dynamique sera plus facile à comprendre en dimension finie : s'il y a seulement N atomes, l'état $\eta = (x_1, \vec{v}_1, \dots, x_N, \vec{v}_N)$ du système se voit comme un point de l'espace euclidien ⁽¹⁵⁾ \mathbb{R}^{2Nd} , qu'on soumet à l'évolution

$$d\vec{\eta} = -\frac{\alpha}{2} \vec{\nabla} H dt + (\alpha\theta)^{1/2} d\vec{W}_t. \quad (1.33)$$

On connaît alors des résultats très puissants sur la convergence du système vers l'équilibre, notamment en termes de trou spectral. Cela dit, quand on s'intéresse comme ici à des systèmes physiques régis par une équation *du second ordre*, (1.33) ne signifie plus rien et n'est bonne selon moi qu'à jeter aux orties.

☛ Dans la suite de cette section, nous ne nous intéresserons qu'au bain de Langevin.

1.4.2 Hypocoercivité

Le bain de Langevin doit assurer les décorrélations du système et la convergence de sa loi vers la mesure \mathbb{P}_θ . En termes mathématiques, cela s'interprète par la

Question 1.12. Notons \mathcal{L} l'opérateur d'évolution correspondant au bain de Langevin (§ 1.4.1, n° 3). Existe-t-il des constantes $C \geq 1$ et $\lambda > 0$ telles que, pour toute fonction $f \in L^2(\mathbb{P}_\theta)$ avec $\mathbb{E}_\theta(f) = 0$, on ait

$$\|e^{t\mathcal{L}} f\|_{L^2(\mathbb{P}_\theta)} \leq C e^{-\lambda t} \|f\|_{L^2(\mathbb{P}_\theta)} ? \quad (1.34)$$

(1.34) signifie qu'« aux grandes échelles de temps, c'est comme si \mathcal{L} était contractant dans L^2/\mathbb{R} ». Mais il faut bien noter qu'en fait \mathcal{L} n'est *pas* contractant puisque la diffusion (le bruit blanc) n'agit que sur les vitesses ; ainsi, pour f ne dépendant que des positions, on a $\langle \mathcal{L}f, f \rangle_{L^2(\mathbb{P}_\theta)} = 0$.

Cette situation a été baptisée *hypocoercivité* par Villani dans [14], où il donne des critères en termes de commutateurs des parties symétrique et antisymétrique de l'opérateur \mathcal{L} pour assurer que celui-ci soit hypocoercif, avec

⁽¹⁵⁾Ici se cache une « hérésie » physique, à savoir qu'il n'y a pas de structure euclidienne naturelle sur l'espace des phases dans la mesure où les positions et les vitesses n'ont pas la même homogénéité physique ; on sous-entend donc qu'on a choisi une constante de temps qui permet d'identifier les distances à des vitesses.

des constantes explicites. Ici, nous avons une situation hypocoercive où les critères de [14] ne s'appliquent plus car nous sommes en dimension infinie ; il faut donc trouver de nouvelles techniques pour répondre à la question 1.12.

1.4.3 Pistes de recherche

La première idée pour étudier l'hypocoercivité en dimension infinie est d'utiliser une technique de couplage, comme suit.

Dans le cas d'un système d'une seule particule, on peut *coupler* l'évolution stochastique de deux particules en les soumettant au *même* bruit blanc, et on constate alors que ces deux particules tendent à coalescer dans l'espace des phases. Malheureusement, la distance naturelle $(|\overline{x_1 x_2}|^2 + \alpha |\overline{v_2} - \overline{v_1}|^2)^{1/2}$ entre les particules décroît de manière apparemment assez chaotique ; en particulier elle ne décroît pas à tout instant ce qui empêche d'utiliser un lemme de Gronwall. L'idée consiste alors à changer la façon de mesurer la distance en introduisant un produit scalaire équivalent au produit standard, mais contenant un *terme croisé* $-\overline{x_1 x_2} \cdot (\overline{v_2} - \overline{v_1})$. Pour un choix judicieux d'un tel produit scalaire, il est alors possible de minorer la décroissance de la distance entre les deux particules par un terme purement déterministe et d'en déduire la décroissance exponentielle de la distance entre les particules au sens du nouveau produit scalaire, et donc également au sens de l'ancien puisqu'on ne perd qu'une constante dans l'affaire. Une fois établie la convergence exponentielle au sens du couplage, des arguments de compacité permettent d'en déduire la convergence L^2 .⁽¹⁶⁾

En dimension infinie, on peut alors envisager de définir le même genre de distance pour obtenir la convergence au sens du couplage. De fait, cela est possible et même facile, mais cette fois-ci la difficulté consiste à relier la convergence au sens du couplage à la convergence L^2 :

Question 1.13. *En dimension infinie, peut-on trouver des conditions simples telles que la convergence au sens du couplage entraîne la convergence L^2 ?*

Autre piste : dans le cas de l'étude de la coercivité (comme la dynamique de Glauber, § 1.4.1, n° 4), les techniques utilisées pour montrer la coercivité en volume infini consistent à *tensoriser* des relations sur un nombre fini d'atomes, grâce à l'utilisation de « relations de balayage » (voir [15]). De fait, dans notre modèle il est facile d'obtenir des résultats d'hypocoercivité pour un nombre fini d'atomes. Peut-être arriverons-nous à tensoriser l'hypocoercivité comme la coercivité :

Question 1.14. *Peut-on utiliser des techniques de tensorisation pour démontrer l'hypocoercivité d'un système infini-dimensionnel comme celui du bain de Langevin ?*

⁽¹⁶⁾Cette technique de couplage peut être vue comme une approche duale des techniques fonctionnelles mises au points par Villani dans [14].

2 Équation de Boltzmann en champ moyen

2.1 Motivations

2.1.1 Équations de champ moyen

L'étude des systèmes composés d'un grand nombre de particules identiques commence en 1872 avec l'équation de Boltzmann. Moyennant certaines hypothèses physiquement réalistes, Boltzmann y décrit en effet l'évolution de la distribution des positions et vitesses des particules d'un gaz peu dense : notant $f(x, \vec{v}, t)$ la densité de particules ayant la position x et la vitesse \vec{v} à l'instant t , l'évolution de f est donnée par

$$\partial_t f = -\vec{v} \cdot \vec{\nabla}_x f + Q(f, f), \quad (2.1)$$

où Q est un opérateur quadratique appelé *noyau de collision*, de la forme :

$$\begin{aligned} Q(f, f)(x, \vec{v}) = & \\ & - f(x, \vec{v}) \int_{\mathbb{R}^3} (f(x, \vec{v} + 2\vec{w}) + f(x, \vec{v} - 2\vec{w})) \left(\int_{\mathbb{S}^2} B_{\vec{w}}(\vec{n}) d\vec{n} \right) d\vec{w} \\ & + \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} f(x, \vec{v}_1) f(x, \vec{v}_2) \left(B_{|\vec{w}|\vec{n}} \left(\frac{\vec{w}}{|\vec{w}|} \right) + B_{|\vec{w}|\vec{n}} \left(-\frac{\vec{w}}{|\vec{w}|} \right) \right) d\vec{w} d\vec{n}, \end{aligned} \quad (2.2)$$

où on a noté pour alléger les notations :

$$\vec{v}_1 = \vec{v} - \vec{w} - |\vec{w}|\vec{n}, \quad (2.3)$$

$$\text{resp. } \vec{v}_2 = \vec{v} - \vec{w} + |\vec{w}|\vec{n}. \quad (2.4)$$

Expliquons les deux termes de (2.1) :

- Le terme $-\vec{v} \cdot \vec{\nabla}_x f$ traduit simplement l'inertie des particules.
- Le terme $Q(f, f)$ traduit l'effet des collisions entre paires de particules ; il est décrit par les fonctions $B_{\vec{w}}$ sur \mathbb{S}^2 — qu'on peut expliciter si on connaît le potentiel d'interaction entre deux particules —, où $B_{\vec{w}}(\vec{n})$ représente la densité pour que deux particules de vitesses respectives $\vec{u} \pm \vec{w}$ se heurtent et repartent avec les vitesses $\vec{u} \pm |\vec{w}|\vec{n}$.

On remarque que l'opérateur Q n'agit sur f qu'au niveau des vitesses ; cela suggère de définir l'équation de Boltzmann *en champ moyen* comme l'équation (2.1) où on ne garde que l'opérateur de collision dans le second membre, en s'intéressant uniquement à la dépendance de f en \vec{v} :

$$\begin{aligned} \partial_t f(\vec{v}) = & \\ & - f(\vec{v}) \int_{\mathbb{R}^3} (f(\vec{v} + 2\vec{w}) + f(\vec{v} - 2\vec{w})) \left(\int_{\mathbb{S}^2} B_{\vec{w}}(\vec{n}) d\vec{n} \right) d\vec{w} \\ & + \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} f(\vec{v}_1) f(\vec{v}_2) \left(B_{|\vec{w}|\vec{n}} \left(\frac{\vec{w}}{|\vec{w}|} \right) + B_{|\vec{w}|\vec{n}} \left(-\frac{\vec{w}}{|\vec{w}|} \right) \right) d\vec{w} d\vec{n}. \end{aligned} \quad (2.5)$$

Physiquement, l'équation (2.5) revient simplement à supposer que la distribution f des positions-vitesses est homogène en x . On peut aussi considérer que tout se passe comme si chaque particule ressentait de la même façon l'effet de toutes les autres particules, si éloignées soient-elles, d'où l'expression *champ moyen*.

2.1.2 Limite de champ moyen

On souhaiterait maintenant déduire cette équation *macroscopique* d'un modèle *microscopique* ; autrement dit, on va considérer un modèle discret à N particules, dont on voudra dire que le comportement à la limite va être celui de l'équation de champ moyen associée quand $N \rightarrow \infty$.

Détaillons un bébé-modèle appelé *modèle de Kac*, introduit par Kac dans [16] : le système microscopique est constitué de N particules $i = 1, \dots, N$ décrites uniquement par leurs vitesses $v_i \in \mathbb{R}$. À chaque couple de particules $\{i, j\}$ est associé un processus de Poisson sur \mathbb{R}_+ d'intensité $1/N$, et lorsqu'il y a un point dans le processus de Poisson $\{i, j\}$, le couple de vitesses (v_i, v_j) change et se replace au hasard sur le cercle de rayon $(v_i^2 + v_j^2)^{1/2}$ dans \mathbb{R}^2 ⁽¹⁾.

☛ On remarquera que le taux de sauts entre deux particules est renormalisé quand le nombre de particules augmente, de sorte que le taux total de collisions subies par une particule tende vers une constante quand $N \rightarrow \infty$.

Quelle limite attend-on pour ce modèle lorsque $N \rightarrow \infty$? D'après la loi des grands nombres, la proportion asymptotique de particules qui subissent un certain évènement devrait être la probabilité de survenue de l'évènement en question. Ainsi, s'il y a une proportion $d\mu(v_1)$ de particules dans l'état v_1 et une proportion $d\mu(v_2)$ de particules dans l'état v_2 , on s'attend à ce qu'à la limite il y ait $d\mu(v_1) d\mu(v_2)$ collisions entre particules de vitesses respectives v_1 et v_2 par nombre de particules et par unité de temps, et même que les collisions de ce type dont les vitesses émergentes soient $(v_1^2 + v_2^2)^{1/2} \cos \theta$ et $(v_1^2 + v_2^2)^{1/2} \sin \theta$ soient au nombre de $d\mu(v_1) d\mu(v_2) d\theta/2\pi$ ⁽²⁾.

Par conséquent, on peut prédire que la répartition des vitesses des particules, considérée comme une mesure sur \mathbb{R} , va tendre quand $N \rightarrow \infty$ vers une distribution *déterministe* régie par l'évolution

$$\partial_t \mu = Q(\mu, \mu), \quad (2.6)$$

où Q est un opérateur de collision quadratique qu'il est possible de définir rigoureusement en considérant l'intégrale de μ contre une fonction-test φ mesurable et bornée

$$\partial_t \langle \mu, \varphi \rangle = -\langle \mu, \varphi \rangle + \frac{1}{2\pi} \int_{\mathbb{R} \times \mathbb{R} \times \mathbb{S}^1} \varphi((v_1^2 + v_2^2)^{1/2} \cos \theta) d\mu(v_1) d\mu(v_2) d\theta. \quad (2.7)$$

⁽¹⁾Ainsi dans ce modèle il y a une quantité conservative, qui est $\sum_{i=1}^N v_i^2$.

⁽²⁾Ici on compte différemment les couples de vitesses émergentes $(v \cos \theta, v \sin \theta)$ et $(v \sin \theta, v \cos \theta)$, bien qu'ils correspondent à la même situation dans la mesure où les particules sont indifférenciables.

De fait, il est possible de démontrer certaines formes plus ou moins fortes de la limite de champ moyen, et ce par des techniques relativement standard. Par exemple, on a :

Théorème 2.1. *On suppose qu'au temps $t = 0$, la distribution empirique des vitesses des particules, notée $\hat{\mu}_0^N$, converge en probabilité vers une distribution de probabilité déterministe μ_0 , au sens où pour toute fonction continue et bornée φ sur \mathbb{R} , on a*

$$\forall \delta > 0 \quad \mathbb{P}\left(|\langle \hat{\mu}_0^N, \varphi \rangle - \langle \mu_0, \varphi \rangle| \geq \delta\right) \xrightarrow[N \rightarrow \infty]{} 0. \quad (2.8)$$

Alors pour tout temps $t > 0$, la distribution empirique des vitesses des particules au temps t converge en probabilité vers la mesure μ_t qui est l'évolution de la mesure μ_0 sous l'équation de champ moyen (2.6).

La démonstration classique de ce résultat (voir par exemple [17, pp. 312–315]) consiste à étudier l'évolution de $\hat{\mu}_t^N$ au moyen de martingales, qui permettent de contrôler la façon dont $\langle \hat{\mu}_t^N, \varphi \rangle$ fluctue autour de son espérance, laquelle est donnée par le générateur du processus de Markov microscopique. Cela permet d'identifier la limite de champ moyen, si elle existe, comme étant nécessairement celle donnée par l'équation (2.6). Ensuite on utilise des arguments topologiques de compacité (critères de tension) pour conclure que $\hat{\mu}^N$ converge effectivement vers la limite de champ moyen.

Néanmoins, ces arguments topologiques ne renseignent pas sur la façon fine dont $\hat{\mu}^N$ converge vers μ . Il serait donc intéressant d'aborder la limite de champ moyen dans un cadre où l'espace des mesures de probabilités serait doté d'une métrique naturelle, en vue :

- d'une part, de donner des résultats en termes de *vitesse* de convergence ;
- d'autre part, d'obtenir des estimées *non asymptotiques*.

2.2 Les distances de couplage

2.2.1 Contrôle des grandes déviations par couplage⁽³⁾

Pour X est un espace mesurable et μ, ν deux mesures de probabilité sur X , on appelle *couplage* entre μ et ν toute mesure π sur $X \times X$ dont la première marginale est μ et la seconde ν ⁽⁴⁾. Notant $\Pi(\mu, \nu)$ l'ensemble des couplages entre μ et ν , si X est doté d'une métrique d on définit la *distance de couplage de Wasserstein* $W_1(\mu, \nu)$ par

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times X} d(x, y) d\pi(x, y), \quad (2.9)$$

⁽³⁾Attention, le mot « couplage » dans ce titre de paragraphe ne renvoie pas au même concept de couplage que dans le titre de section correspondant.

⁽⁴⁾ π peut être vue comme un plan de transport pour déplacer une masse répartie selon μ sur X vers la répartition ν , où $\mathcal{L}oi(\mu(x, \cdot)|x)$ décrirait la façon dont est redistribuée la masse située au point x .

qui est effectivement une distance (sur les mesures de probabilité ayant un premier moment pour d).

La distance de couplage est une façon très naturelle d'évaluer l'écart entre deux mesures ; au demeurant, si X est compact alors la convergence selon W_1 équivaut à la convergence faible. Dans [18], Bolley *et al.* cherchent à préciser la limite de champ moyen en termes de distance de Wasserstein. Ils commencent par donner le résultat de grandes déviations suivant sur les mesures empiriques :

Théorème 2.2. *Soit μ une mesure de probabilité sur \mathbb{R}^d ayant un moment carré-exponentiel ⁽⁵⁾ ; on sait qu'alors μ vérifie une inégalité de Talagrand $T_1(\lambda)$, c.à.d. qu'on a pour un $\lambda > 0$*

$$\forall \nu \text{ probabilité sur } \mathbb{R}^d \quad H(\nu|\mu) \geq \frac{\lambda}{2} (W_1(\nu, \mu))^2, \quad (2.10)$$

où $H(\nu|\mu)$ note l'entropie relative de ν par rapport à μ .

Ici nous prendrons pour λ la valeur optimale dans (2.10). Pour $N \geq 1$, soient X_1, \dots, X_N des variables dans \mathbb{R}^d i.i.d. selon μ et notons $\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$ la mesure empirique associée. Alors pour tout $\lambda' > \lambda$, pour tout $\varepsilon > 0$, on a dès que N est suffisamment grand

$$\mathbb{P}[W_1(\mu, \hat{\mu}^N \geq \varepsilon)] \leq e^{-\lambda' N \varepsilon^2 / 2}. \quad (2.11)$$

Ce résultat est optimal au sens où (2.11) devient fausse dès lors que $\lambda' > \lambda$. En outre, on peut préciser la valeur de N à partir de laquelle (2.11) est vraie : pour tout $d' > d$, il suffit de prendre

$$N \geq N_0(\varepsilon^{-(d'+2)} \vee 1), \quad (2.12)$$

où N_0 est une constante explicite en λ' , d' et un moment carré-exponentiel de μ .

Le théorème 2.2 est un théorème statique, mais on va arriver à en déduire un résultat dynamique grâce au bon comportement de la distance W_1 sous l'évolution stochastique, du moins dans les cas où il n'y a pas de chocs.

Plus précisément, on considère le modèle suivant :

Modèle 2.3. *On se donne un potentiel à une particule V et un potentiel d'interaction à deux particules W symétrique, tous deux définis sur \mathbb{R}^d , supposés suffisamment réguliers. Le modèle microscopique est constitué par N particules identiques dont l'évolution est la somme de trois phénomènes :*

- L'évolution de la particule dans le potentiel V (régie par une ÉDO d'ordre 1) ;
- L'interaction de la particule avec ses voisines, de type champ moyen, régie par W ;
- Un bruit blanc indépendant sur chacune des particules.

⁽⁵⁾I.e., pour $\alpha > 0$ suffisamment petit on a $\mu[\exp(\alpha|x|^2)] < \infty$.

En d'autres termes, la position X_t^i de la particule i à l'instant t évolue suivant l'équation

$$\overrightarrow{dX}_t^i = \sqrt{2}d\overrightarrow{B}_t^i - \overrightarrow{\nabla}V(X_t^i)dt - \frac{1}{N} \sum_{j=1}^N \overrightarrow{\nabla}W(X_t^i - X_t^j)dt, \quad (2.13)$$

où les B^i sont des mouvements browniens indépendants.

Par ailleurs, la position initiale des particules est tirée i.i.d. suivant une certaine loi μ_0 .

L'idée est la suivante : puisque les particules ont été tirées i.i.d., par le théorème 2.2 leur distribution empirique au temps initial est proche de μ_0 au sens de la distance de Wasserstein. Par conséquent, le terme d'interaction d'une particule avec toutes les autres va être proche de ce qu'il serait si les autres particules suivaient exactement la distribution μ_0 . Ainsi, la trajectoire individuelle d'une particule devrait ressembler au processus non-linéaire défini ci-après, introduit par Sznitman [19] :

Définition 2.4. *Le processus non-linéaire associé au modèle 2.3 est la loi de la trajectoire d'une particule X dont la loi initiale est tirée selon μ_0 et dont l'évolution suit l'équation (2.13), sauf que le terme d'interaction $\frac{1}{N} \sum_{j=1}^N \overrightarrow{\nabla}W(X_t^i - X_t^j)$ y est remplacé par ce qu'il serait si les autres particules suivaient exactement la loi μ_t prédite par la limite de champ moyen, i.e. par*

$$\int_{\mathbb{R}^d} \overrightarrow{\nabla}W(X_t - Y)d\mu_t(Y). \quad (2.14)$$

L'intérêt du processus non-linéaire est que la loi au temps t d'une particule suivant ce processus sera exactement μ_t ⁽⁶⁾. Grâce au théorème 2.2, on sait qu'alors la distribution empirique de N particules indépendantes suivant chacune un processus non-linéaire est proche de μ_t . Or au temps initial, nos particules réelles sont distribuées exactement comme ces particules non-linéaires. Nous allons donc *coupler* l'évolution des particules réelles à celle des particules non-linéaires, simplement en leur associant le même terme de bruit. Par un lemme de Gronwall, on démontre alors que la distance W_1 de la distribution des particules réelles à celle des particules virtuelles, et donc à la mesure-limite, va rester petite au cours de l'évolution grâce à la régularité des potentiels.

☛ *Nous renvoyons à [18, théorème 1.9] le lecteur désirant un énoncé et une preuve précis.*

⁽⁶⁾À vrai dire, la définition 2.4 peut en fait servir de définition de l'équation de la limite de champ moyen : il suffit pour ce faire de remplacer « μ_t » par « la loi de la particule au temps t ». De la sorte, on a un processus parfaitement défini, qui ne concerne qu'une particule, mais dont l'équation fait intervenir la loi du processus lui-même à la façon d'une ÉDO non linéaire, d'où l'appellation de « processus non-linéaire ».

2.2.2 Modèles à chocs

Pleins d'enthousiasme, nous désirons alors appliquer les même méthodes à des modèles à chocs comme les équations de Boltzmann et de Kac. Que l'évolution des particules soit saltatoire plutôt que continue change quelques détails techniques, mais ce n'est pas un problème substantiel.

Par contre, quand on tente d'appliquer la méthode de couplage du § 2.2.1, on aperçoit un obstacle majeur : il n'est *pas* possible de coupler l'évolution réelle à celle de N particules non-linéaires indépendantes, car dans le modèle réel les sauts des particules ne sont pas indépendants : en effet, les particules sautent deux par deux !!

Pour contourner cet écueil, j'ai proposé un méthode consistant à abandonner complètement l'idée de couplage, et à plutôt regarder l'ensemble des sauts subis par le système sur un pas de temps. En effet, quand le nombre de sauts devient très grand, par la loi des grands nombres — ou plutôt, encore et toujours par le théorème 2.2 pour des résultats non-asymptotiques — on peut prédire que les sauts vont se répartir à peu près comme prédit par la limite de champ moyen ⁽⁷⁾.

En fait, pour des raisons qui seront expliquées au § 2.2.3, je n'ai pas utilisé le théorème 2.2 mais une variante personnelle (voir document IX). Le résultat que j'ai obtenu — qui n'est qu'un exemple scolaire destiné à montrer les possibilités de ma technique, ainsi largement susceptible de généralisations — est le suivant :

Notation 2.5. Dans l'énoncé du théorème 2.6, pour X un espace métrique compact, $\mathcal{M}_1^+(X)$ note l'ensemble des mesures de probabilité sur X , et l'espace $\mathcal{M}_0(X)$ des mesures signées sur X de masse totale nulle est doté de la norme définie par la distance de couplage W_1 . D'autre part, si X et Y sont deux espaces métriques, $X \times Y$ est muni de la distance somme.

Théorème 2.6 (voir document X). Avec les notations 2.5 supra, on considère une variété riemannienne compacte X de dimension $d > 2$ et deux applications $\tau : X \times X \rightarrow [0, \infty[$ et $h : X \times X \rightarrow \mathcal{M}_1^+(X \times X)$ symétriques ⁽⁸⁾.

On suppose que :

- τ est bornée ; on note $\bar{\tau}$ son supremum.
- L'application

$$\begin{aligned} S : X \times X &\longrightarrow \mathcal{M}_0(X \times X) \\ (x, y) &\mapsto \tau(x, y)(h(x, y) - \delta_{(x,y)}) \end{aligned} \quad (2.15)$$

est lipschitzienne ; on note κ sa constante de Lipschitz.

On considère $N \geq 1$ particules x^1, \dots, x^N en interaction. On note $\eta = (x^1, \dots, x^N)$ l'état du système, et pour $i < j$, $y^i, y^j \in X$, on pose

$$\eta_{y^i, y^j}^{i,j} = (x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^{j-1}, y^j, x^{j+1}, \dots, x^N). \quad (2.16)$$

⁽⁷⁾Ici se place une petite astuce, car le théorème 2.2 concerne des mesures empiriques « statiques » alors que dans notre situation la loi des sauts évolue en même temps que le système. On remédie à ce problème en choisissant un pas de temps suffisamment court pour que le système n'évolue pratiquement pas pendant ce délai. Il y a alors un compromis à faire entre, d'une part cette question de brièveté du pas de temps, d'autre part le besoin de prendre un pas de temps assez long pour que les sauts y soient nombreux et que le résultat du théorème 2.2 soit fin. Au final, il en résulte une petite dégradation entre la vitesse de convergence prédite par le théorème 2.2 et celle obtenue avec notre méthode.

⁽⁸⁾I.e., notant $t : X \times X \rightarrow X \times X$ l'application qui envoie (x, y) sur (y, x) , on a $\tau \circ t = \tau$ et $h \circ t = t \circ h$.

L'évolution du système se fait alors de la façon suivante : à $t = 0$, x_0^1, \dots, x_0^N sont distribués indépendamment selon une mesure de probabilité arbitraire μ_0 , puis l'évolution est markovienne avec pour générateur

$$\mathcal{L}f(\eta) = \frac{1}{N} \sum_{i < j} \tau(x^i, x^j) \int_{X \times X} (f(\eta_{y^i, y^j}^{i, j}) - f(\eta)) d(h(x^i, x^j))(y^i, y^j). \quad (2.17)$$

Alors il existe une constante C indépendante de μ_0 telle que pour tout $T \geq 0$

$$C' > \frac{e^{\kappa T/2} - 1}{\kappa/2} C \quad \Rightarrow \quad \mathbb{P}\left(\|\widehat{\mu}_t^N - \mu_t\|_W \geq \frac{C'}{N^{1/(n+1)}}\right) \underset{N \rightarrow \infty}{\downarrow} 0, \quad (2.18)$$

où « $f(x) \underset{x \rightarrow +\infty}{\downarrow} 0$ » signifie « $f(x)$ décroît plus vite, quand $x \rightarrow \infty$, que n'importe quel $1/x^q$ ».

2.2.3 Limites de l'approche en distance W_1

Dans le théorème 2.6 ci-dessus, la décroissance obtenue est relativement lente puisqu'elle ne se fait qu'à la vitesse $1/N^{1/(d+1)}$, alors qu'on est habitué à des vitesses en $1/N^{1/2}$ (cf. théorème-limite central). De même, dans le théorème 2.2, la concentration obtenue est gaussienne mais n'est valable que pour des déviations au moins de l'ordre de $1/N^{1/(d+2)}$, ce qui est pire encore ! En réalité, ce phénomène de décroissance lente est inhérent au choix de la norme W_1 , comme le montre la proposition élémentaire suivante :

Proposition 2.7. Notons λ la mesure de Lebesgue sur \mathbb{R}^n et soit $U \subset \mathbb{R}^n$ avec $\lambda(U) = 1$. Alors pour tous $X_1, \dots, X_N \in U$, notant $\widehat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$, on a

$$W_1(\lambda|_U, \widehat{\mu}^N) \geq \frac{c_n}{N^{1/n}}, \quad (2.19)$$

où $c_n = \frac{n}{n+1} \left(\frac{(n/2)!}{\pi^{n/2}}\right)^{1/n}$ est une constante absolue.

Démonstration. Considérons un couplage π entre $\widehat{\mu}^N$ et λ , alors pour tout i $\pi_i = \mathcal{L}oi(\pi|X_i)$ est une mesure de probabilité sur \mathbb{R}^n majorée par $N\lambda$. Soit $\bar{\pi}_i(r) = \pi_i(B(X_i, r))$; on a $\int_0^\infty d\bar{\pi}_i(r) = 1$ et $d\bar{\pi}_i(r) \leq N\sigma_n r^{n-1} dr$, où σ_n est la surface de la sphère unité dans \mathbb{R}^n . Sous ces conditions, $\int_{\mathbb{R}^n} |\overrightarrow{X_i y}| d\pi_i(y) = \int_0^\infty r \bar{\pi}_i(r) dr$ est minimal pour $\bar{\pi}_i = \sigma_n r^{n-1} \mathbb{1}_{0 \leq r \leq (\frac{n}{N\sigma_n})^{1/n}}$ et vaut alors $\frac{n}{n+1} \left(\frac{n}{N\sigma_n}\right)^{1/n}$. Comme au final $W_1(\lambda|_U, \widehat{\mu}^N) = \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{R}^n} |\overrightarrow{X_i y}| d\pi_i(y)$, on obtient bien le résultat (2.19) annoncé. \square

2.3 Théorèmes-limite centraux uniformes

2.3.1 Problématique

Rappelons un résultat fondamental dans la théorie du couplage, pour lequel on pourra consulter [20, chap. 5] :

Théorème 2.8 (Kantorovitch–Rubinstein). *Soit X un espace métrique séparable et μ, ν deux mesures de probabilités sur X ayant un premier moment, alors*

$$W_1(\mu, \nu) = \sup \left\{ \left| \int_X f d\nu - \int_X f d\mu \right| ; f \text{ 1-lipschitzienne} \right\}. \quad (2.20)$$

En substance, donc, utiliser la distance de couplage entre deux mesures, c'est regarder la différence maximale de leurs espérances quand on les teste contre les fonctions 1-lipschitziennes.

Il est à noter que si μ est une mesure de probabilité vérifiant une inéagilité $T_1(\lambda)$ ⁽⁹⁾ et f une fonction 1-lipschitzienne, on a pour tout N , notant $\hat{\mu}^N$ la mesure empirique à N points

$$\mathbb{P} \left(\left| \int_X f d\hat{\mu}^N - \int_X f d\mu \right| \geq \varepsilon \right) \leq 2e^{-\lambda\varepsilon^2/2}, \quad (2.21)$$

ce qui est un résultat très fin donnant une convergence en $1/N^{1/2}$: le problème est ce qui se passe quand on veut tester contre *toutes* les fonctions 1-lipschitziennes à la fois.

En fait, si on regarde quelle est la « pire » fonction 1-lipschitzienne pour tester l'écart entre $\hat{\mu}^N$ et μ , on se rend compte que c'est une fonction qui ressemble à $\min_i d(X_i, \cdot)$, qui présente des fluctuations à très petite échelle de façon à être « petite » quand on est près d'un X_i et « grande » ailleurs.

Mais ce genre de fonction-test n'est pas naturel du tout ! En quelque sorte, le problème est que l'espace de toutes les fonctions 1-lipschitziennes est un espace de test « trop gros » : on voudrait plutôt tester contre un espace de fonctions suffisamment régulières, plus petit mais assez riche pour autant.

2.3.2 La théorie UCLT

La théorie des théorèmes-limite centraux uniforme est étudiée de manière très complète par Dudley dans [21]. La question qu'il y résout est la suivante : on considère une mesure de probabilité μ sur un espace mesurable, et une classe de fonctions $\mathcal{F} \subset \mathcal{L}^2(d\mu)$, où $\mathcal{L}^2(\mu)$ est l'ensemble des fonctions réelles mesurables dont le carré est intégrable sous μ ⁽¹⁰⁾. Pour $\hat{\mu}^N$ la mesure empirique à N points, on peut alors voir $\hat{\mu}^N - \mu$ comme une application de \mathcal{F} dans \mathbb{R} , et on se demande si la loi de $N^{1/2}(\hat{\mu}^N - \mu)$ converge, au sens où on pourrait construire sur un même espace de probabilité un processus aléatoire $G_\mu : \mathcal{F} \rightarrow \mathbb{R}$ et des réalisations des $\hat{\mu}^N$ ⁽¹¹⁾ telles qu'on ait

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\sup_{f \in \mathcal{F}} \left| \langle N^{1/2}(\hat{\mu}^N - \mu), f \rangle - G_\mu(f) \right| \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0. \quad (2.22)$$

⁽⁹⁾La définition d'une inéagilité $T_1(\lambda)$ a été donnée dans l'énoncé du théorème 2.2

⁽¹⁰⁾Cet espace n'est pas l'espace habituel $L^2(\mu)$; en fait $L^2(\mu) = \mathcal{L}^2(\mu) / \sim$, où \sim est la relation d'égalité μ -p.p..

⁽¹¹⁾Attention : au départ, la mesure $\hat{\mu}^N$ a été définie comme $\frac{1}{N} \sum_{i=1}^N \delta_{X_i}$ pour des X_i i.i.d. selon μ , mais ici on n'exige pas que ces X_i existent ; on voit juste $\hat{\mu}^N$ comme une mesure aléatoire ayant la loi en question.

La définition (2.22) est une sorte de convergence en loi : en effet,

Théorème 2.9 (Skorokhod). *Soit (E, d) un espace métrique séparable ; soient $(\nu_n)_{n \geq 1}$ et ν_∞ des mesures de probabilité sur X . Il y a équivalence entre les deux propriétés suivantes :*

1. ν_n converge faiblement vers ν_∞ ;
2. On peut construire sur un même espace de probabilité des variables aléatoires $(X_n)_{n \geq 1}$ et X_∞ telles qu'on ait

$$\forall \varepsilon > 0 \quad \mathbb{P}(d(X_n, X_\infty) \geq \varepsilon) \xrightarrow{N \rightarrow \infty} 0. \quad (2.23)$$

Si nous n'avons pas écrit « convergence faible » tout-à-l'heure, c'est parce que les lois considérées sont à valeurs dans l'espace $\mathcal{F}^{\mathbb{R}}$ muni de la (pseudo-)norme du supremum $|\cdot|_{\text{sup}}$, que nous noterons $\mathcal{L}_b(\mathcal{F})$, lequel n'est pas séparable en général.

Pire, la mesure aléatoire $\hat{\mu}^N$ n'est même pas mesurable pour la tribu borélienne de $\mathcal{L}_b(\mathcal{F})$: prenons par exemple le cas où μ est la mesure uniforme sur $[0, 1[$, et \mathcal{F} est l'ensemble des fonctions de Heaviside $\{\mathbb{1}_{[x, 1[} ; x \in [0, 1]\}$. Considérons $\mathcal{A} \subset [0, 1[$ non mesurable, et posons

$$U = \bigcup_{x \in \mathcal{A}} \left\{ y \in \mathcal{L}_b(\mathcal{F}) ; \sup_{f \in \mathcal{F}} |y(f) - f(x)| < \frac{1}{2} \right\}, \quad (2.24)$$

alors U est un ouvert de $(\mathcal{L}_b(\mathcal{F}), |\cdot|_{\text{sup}})$, donc un ensemble mesurable, et pourtant son image réciproque par l'application $x \mapsto \delta_x$ est \mathcal{A} , qui n'est pas mesurable.

Ces questions de mesurabilité amènent à des contorsions techniques assez désagréables, dont j'expliquerai au § 2.3.3 comment je compte les contourner.

Quoi qu'il en soit, on peut commencer par se demander à quoi doit ressembler le processus G_μ qui vérifierait (2.22), et s'il a une chance d'exister. D'après le théorème-limite central unidimensionnel ordinaire, on voit que, si G_μ existe, on aura nécessairement pour toutes fonction $f, g \in \mathcal{F}$

$$G_\mu(g) - G_\mu(f) \sim \mathcal{N}(\text{Var}_\mu[g - f]) : \quad (2.25)$$

cela caractérise une loi sur $\mathcal{F}^{\mathbb{R}}$ qu'on appelle *processus isonormal* sur \mathcal{F} ⁽¹²⁾.

Comme on étudie de la convergence pour la norme $|\cdot|_{\text{sup}}$, on voudrait que p.s., G_μ soit un élément de $\mathcal{L}_b(\mathcal{F})$. En outre, il est naturel d'exiger que G_μ , comme les $N^{1/2}(\hat{\mu}^N - \mu)$ dont il est la limite, soit *prélinéaire*, i.e. vérifie $G_\mu(f+g) = G_\mu(f)$, resp. $G_\mu(\lambda f) = \lambda G_\mu(f)$. Enfin, comme l'écart entre $G_\mu(f)$ et $G_\mu(g)$ est de l'ordre de $\text{Var}[f - g]^{1/2}$ d'après (2.25), on va aussi imposer que G_μ soit presque-sûrement continue de \mathcal{F} doté de la semi-norme $\rho(f, g) = \text{Var}[f - g]^{1/2}$ dans \mathbb{R} .

⁽¹²⁾Quand j'écris « cela caractérise une loi », cela n'est valable que pour la tribu-produit sur $\mathcal{F}^{\mathbb{R}}$, qui n'est pas celle qui nous intéresse ici. Dans les critères que nous allons donner ci-dessous, il faudra donc comprendre G_μ comme *une* réalisation possible dans $\mathcal{L}_b(\mu)$ du processus isonormal.

Définition 2.10. *On dit que \mathcal{F} est prégaussienne pour μ s'il est possible de réaliser un processus isonormal sur \mathcal{F} vérifiant les conditions ci-dessus ; on dit que c'est une classe de Donsker si, en outre, le théorème-limite central est valide au sens de la formule (2.22).*

Dans [21], Dudley parvient à établir des conditions nécessaires et suffisantes pour que \mathcal{F} soit prégaussienne, resp. de Donsker. Nous ne donnerons pas ici les conditions pour que \mathcal{F} soit une classe de Donsker, qui seraient un peu longues à donner, mais nous ferons néanmoins remarquer que le « véritable » obstacle à l'établissement d'un TLC uniforme est avant tout la condition de pré-gaussienté, et que d'ailleurs les conditions pour être de Donsker sont du même type que celles pour être prégaussien.

Théorème 2.11 (voir [21, chap. 2]). *Notons $N(\mathcal{F}, r)$ le nombre minimal d'ensembles de diamètre $\leq 2r$ pour la semi-norme ρ nécessaires pour recouvrir \mathcal{F} . Alors :*

1. *Si \mathcal{F} est bornée pour ρ et si $\int_0 (\log N(\mathcal{F}, r))^{1/2} dr < \infty$, alors \mathcal{F} est pré-gaussienne ;*
2. *« Réciproquement », si $\limsup_{r \rightarrow 0} r^2 \log N(\mathcal{F}, r) = \infty$, alors \mathcal{F} n'est pas pré-gaussienne.*

Armés du théorème 2.11, nous sommes en mesure de comprendre pourquoi les fonctions 1-lipschitziennes étaient un espace de test « trop gros » :

Avis 2.12. *Pour $d \geq 2$, μ la mesure uniforme sur \mathbb{T}^d , l'ensemble des fonctions 1-lipschitziennes sur \mathbb{T}^d n'est pas prégaussien.*

Démonstration. Choisissons une fonction 1-lipschitzienne sur \mathbb{R}^d , à support dans $[0, 1]^d$, d'intégrale nulle, non identiquement nulle, que nous noterons f_0 . Pour un entier $m \geq 2$, on peut découper \mathbb{T}^d en m^d petits hypercubes que nous numérotions arbitrairement $1, \dots, m^d$; pour tout m^d -uplet $(\varepsilon_1, \dots, \varepsilon_{m^d})$ avec $\forall i \varepsilon_i \in \{\pm 1\}$, on définit la fonction 1-lipschitzienne $f_{\varepsilon_1 \dots \varepsilon_{m^d}}$ sur \mathbb{T}^d de sorte que sur l'hypercube i , $f_{\vec{\varepsilon}}$ soit une copie réduite (réduite de sorte à garder la même constante de Lipschitz) de $\varepsilon_i f$.

Entre deux multi-indices $\vec{\varepsilon}$ et $\vec{\varepsilon}'$, on introduit la distance de Hamming d_H qui est le nombre de ε_i différant entre ces deux multi-indices ; alors on a

$$\|f_{\vec{\varepsilon}'} - f_{\vec{\varepsilon}}\|_{L^2(\mu)} = (d_H(\vec{\varepsilon}, \vec{\varepsilon}'))^{1/2} m^{-(d/2+1)} \|f_0\|_{L^2(\mu)} \quad (2.26)$$

— pour alléger les notations, on posera $\alpha = \|f_0\|_{L^2(\mu)}$. Par conséquent, un ensemble de diamètre $2r$ pour ρ contient des $f_{\vec{\varepsilon}}$ dont les indices ont un diamètre au sens de la distance de Hamming qui est majoré par $(2r/\alpha)^2 m^{d+2}$, donc ces indices sont au nombre d'au plus $(m^d + 1)^{(2r/\alpha)^2 m^{d+2}}$.

Mais comme il y a 2^{m^d} fonctions $f_{\vec{\varepsilon}}$, il s'ensuit finalement que le nombre d'ensembles de diamètre $2r$ pour la distance ρ nécessaires pour recouvrir toutes les fonctions 1-lipschitziennes est d'au moins

$$N(\mathcal{L}ip_1, r) \geq 2^{m^d} (m^d + 1)^{(2r/\alpha)^2 m^{d+2}}, \quad (2.27)$$

et prenant $m = r^{-\beta}$ avec $\beta \in]2/d, 1[$, on en déduit que $r^2 \ln N(\mathcal{L}ip_1, r) \rightarrow \infty$ quand $r \rightarrow 0$, ce qui permet de conclure par le théorème 2.11, n° 2. \square

2.3.3 Espaces de Sobolev

Dans le cadre des systèmes particules en interaction, les particules évoluent en général dans une variété riemannienne, ce qui permet d'introduire des classes de fonctions naturelles en termes de régularité.

Considérons une variété riemannienne M compacte de dimension d , dont on note Δ l'opérateur de Laplace–Beltrani (avec la convention de signe « $\Delta = -\sum \frac{\partial^2}{\partial x_i^2}$ ») et μ la mesure de volume. On note $\tilde{L}^2(\mu) = L^2(\mu)/\mathbb{R}$ l'ensemble des fonctions définies à constante additive près de variance finie, muni du produit scalaire associé à la covariance. Δ est un opérateur non-borné sur $L^2(\mu)$ et s'annule sur les constantes ; il passe donc au quotient sur $\tilde{L}^2(\mu)$.

Sur $\tilde{L}^2(\mu)$, il est bien connu que Δ est symétrique et strictement positif

$$\forall f, g \in \tilde{L}^2(\mu) \quad \|\Delta f\|_{L^2}, \|\Delta g\|_{L^2} < \infty \Rightarrow \langle \Delta f, g \rangle_{L^2} = \langle f, \Delta g \rangle_{L^2}; \quad (2.28)$$

$$\exists c > 0 \forall f \in \tilde{L}^2(M) \quad \langle \Delta f, f \rangle \geq c \|f\|_{L^2}^2. \quad (2.29)$$

Pour tout $s \geq 0$, on peut alors définir une pseudonorme $\|\cdot\|_{H^s}$ sur \tilde{L}^2 par

$$\|f\|_s = \|\Delta^{s/2} f\|_{L^2}. \quad (2.30)$$

Les fonctions f pour lesquelles $\|f\|_s < \infty$, munies de la norme $\|\cdot\|_{H^s}$, forment alors un espace de Hilbert qu'on appelle *espace de Sobolev homogène d'ordre s* et qu'on note \dot{H}^s .

\dot{H}^s est alors un sous-espace dense de \tilde{L}^2 , ce qui va nous permettre d'assimiler son dual à un sur-espace de \tilde{L}^2 où ce dernier sera dense : pour ce faire, on remarque qu'un élément f de \tilde{L}^2 s'assimile à un élément du dual de \dot{H}^s via l'application $g \mapsto \langle g, f \rangle_{L^2}$ et que \tilde{L}^2 est alors dense dans $(\dot{H}^s)'$. Ainsi représenté, le dual de \dot{H}^s est noté \dot{H}^{-s} ; c'est un espace de Hilbert séparable.

Ce qui est intéressant, c'est que pour $s > d/2$, une mesure signée sur M de masse totale nulle peut s'identifier à un élément de \dot{H}^{-s} , puisqu'en effet les éléments de \dot{H}^s s'identifient alors à des fonctions, définies à constante additive près, bornées sur M , i.e. à des éléments de $\mathcal{L}_b(M)/\mathbb{R}$. En outre, il n'est pas très difficile de voir que l'espace \dot{H}^s est suffisamment « riche » pour que deux mesures différentes sur M correspondent à deux éléments distincts de \dot{H}^{-s} .

Remarque 2.13. Que H^{-s} soit séparable est tout l'intérêt d'avoir choisi de s'intéresser à des espaces de Sobolev H^s plutôt qu'à $\mathcal{C}^s(M)$: le dual de ce dernier espace, en effet, n'est pas séparable, ce qui donne le même genre de problèmes de mesurabilité que quand on considèrerait les mesures comme des éléments d'un espace \mathcal{L}_b .

Théorème 2.14. *Pour $s > d/2$, alors il existe une (unique en loi) variable aléatoire w sur \dot{H}^{-s} , que nous appellerons le bruit blanc sur M , qui vérifie*

$$\forall f \in H^s \quad \langle w, f \rangle \sim \mathcal{N}(\|f\|_{L^2}^2). \quad (2.31)$$

Démonstration. On se limite ici au cas $d \geq 2$: soient $d \geq 2$ et $s > d/2$. Pour démontrer le théorème 2.14, il nous suffit de montrer que la boule unité de \dot{H}^s , que nous noterons $B(\dot{H}^s)$, est prégaussienne ; à cette fin nous allons vérifier le critère n° 1 du théorème 2.11.

Soit $m \geq 2$ un entier naturel ; comme dans la preuve de l'avis 2.12 on peut mailler \mathbb{T}^d par m^d petits hypercubes Q_1, \dots, Q_{m^d} , dont les sommets sont les m^d points x_1, \dots, x_{m^d} d'un réseau de \mathbb{T}^d .

Si f et g sont deux fonctions de $B(\dot{H}^s)$ avec $|g(x_i) - f(x_i)| \leq \varepsilon$ pour tout $i \in \{1, \dots, m^d\}$, alors d'après les inégalités de Sobolev $\|g - f\|_{L^2} \leq \varepsilon + \alpha/m^s$, où α est une certaine constante finie⁽¹³⁾.

Prenons maintenant $\varepsilon = 1/m^s$. Pour N_1, \dots, N_{m^d} des entiers, on considère l'ensemble des fonctions f de $B(\dot{H}^s)$ telles que pour tout $i \in \{1, \dots, m^d\}$ on ait $|f(x_i) - N_i \varepsilon| \leq \varepsilon/2$; d'après ce que nous venons de dire cet ensemble a un diamètre dans L^2 qui est au plus $\alpha' \varepsilon$, avec $\alpha' = \alpha + 1$.

Remarquons par ailleurs que si on ajoute une constante à chacun des N_i , on tombe sur un ensemble de fonctions qui est le même à constante additive près. Il nous suffit donc de compter le nombre de m^d -uplets \vec{N} , à constante additive près, qui définissent un ensemble de fonctions de $B(\dot{H}^s)$ non vide pour obtenir une majoration de $N(B(\dot{H}^s), \alpha' \varepsilon/2)$.

Pour $f \in B(\dot{H}^s)$, puisque $s \geq 1$, $\vec{\nabla} f$ est bornée dans L^2 ; on en tire que, notant $x_i \sim x_j$ pour dire que x_i et x_j sont voisins dans le réseau que forment les x_i ,

$$\sum_{x_i \sim x_j} (|N_i - N_j| - 1)_+^2 \leq \beta m^{d+2s-2} \quad (2.32)$$

pour une certaine constante β . Comme il y a dm^d arêtes dans le réseau, on en déduit que le nombre de \vec{N} possibles, à constante près, est majoré par $(\beta' m)^{\gamma m^d}$ pour certaines constantes β' et γ .

En récapitulant tout ce que nous venons de dire, on a donc obtenu que

$$N(B(\dot{H}^s), \alpha'/2m^s) \leq (\beta' m)^{\gamma m^d}; \quad (2.33)$$

on en déduit bien le critère n° 1 du théorème 2.11. \square

2.3.4 Programme de recherches

À l'heure actuelle, mon objectif est d'essayer de démontrer des limites de champ moyen pour des modèles à chocs, au sens des espaces de Sobolev, dans un cadre non asymptotique.

Eu égard aux travaux présentés dans le § 2.2, l'étude de la convergence pour des modèles de champ moyen passe d'abord par des résultats « statiques » de convergence des mesures empiriques. On aimerait par exemple obtenir un résultat du genre :

Question 2.15. Soit M une variété riemannienne compacte de dimension d et $s > d/2$. Soit μ une mesure de probabilité arbitraire sur M et $\hat{\mu}^N$ la mesure empirique à N points associée. Existe-t-il des constantes C et λ , si possible indépendantes de μ , telles qu'on ait pour tout $N \geq 1$

$$\forall \varepsilon \geq 0 \quad \mathbb{P}(\|\hat{\mu}^N - \mu\|_{H^{-s}} \geq \varepsilon) \leq C \exp(-\lambda N \varepsilon^2) ? \quad (2.34)$$

⁽¹³⁾À vrai dire, je n'ai pas encore eu le temps de m'instruire assez sur les inégalités de Sobolev pour être totalement certain de la validité de cette affirmation ; néanmoins une autre démonstration utilisant des inégalités de Sobolev plus fortes donne quand même le résultat quitte à dégrader la valeur de s .

Ce genre de résultat ne devrait pas être très difficile à prouver dans la mesure où il semble possible de raffiner les résultats de [21] pour en tirer des estimées non asymptotiques.

Plus difficile est le défi d'appliquer ces résultats aux systèmes de champ moyen : *a priori*, les techniques du § 2.2.2 s'adaptent sans trop de difficulté, mais malheureusement elles ne donneraient qu'une convergence en $N^{-1/3}$ alors qu'on espère du $N^{-1/2}$. Il reste donc encore du travail pour résoudre la

Question 2.16. *Soit un modèle de champ moyen à chocs du genre de celui décrit au § 2.2.2 ; soit $s > d/2$, d étant la dimension de l'espace où évoluent les particules. Soit $T > 0$ un temps arbitraire ; est-il vrai que pour tout $\varepsilon > 0$, pour tout $A > 0$, on a*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \|\widehat{\mu}^N - \mu\|_{H^{-s}} \geq \frac{A}{N^{1/2-\varepsilon}}\right) \downarrow 0, \quad (2.35)$$

avec des estimées non asymptotiques ?

Pour passer au modèle de Kac, une difficulté supplémentaire s'ajoute par rapport aux modèles du § 2.2.2, car l'espace dans lequel évoluent les particules n'est plus compact. Pour l'équation de Boltzmann, c'est pire encore, car le noyau de collision n'est pas borné au sens où le taux de chocs entre deux particules tend vers l'infini quand leurs vitesses relatives tendent vers l'infini. Même, si les potentiels d'interaction entre atomes ne sont pas à support compact, l'équation de Boltzmann donne un taux de collision global *infini* (quoique la majorité des collisions correspondent alors à des déplacements très faibles dans l'espace des vitesses). J'espère néanmoins, à terme, résoudre la

Question 2.17. *Peut-on obtenir un résultat de convergence en $N^{-1/2}$ au sens des espaces de Sobolev H^{-s} pour le modèle de Kac, resp. l'équation de Boltzmann ?*

Références

- [1] D. HILBERT : Sur les problèmes futurs des mathématiques. *In Comptes-rendus du 2^e congrès international des mathématiciens (1900)*. Gauthier-Villars, 1902. Traduction : L. Laugel.
- [2] A. EINSTEIN : Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 549:182–193, 1905.
- [3] S. LEPRI, R. LIVI et A. POLITI : Thermal conduction in classical low-dimensional lattices. *Phys. Reports*, 377(1), 2003.
- [4] E. LIEB, J. LEBOWITZ et Z. RIEDER : Properties of a harmonic crystal in a stationary non-equilibrium state. *Journal of Mathematical Physics*, 8:1073–1078, 1967.
- [5] M. GREEN : Markoff random processes and the statistical mechanics of time-dependent phenomena. *Journal of Chemical Physics*, 20:1281, 1952.

- [6] M. GREEN : Markoff random processes and the statistical mechanics of time-dependent phenomena, II : Irreversible processes in fluids. *Journal of Chemical Physics*, 22:1398, 1954.
- [7] R. KUBO : Statistical-mechanical theory of irreversible processes, I : General theory and simple applications to magnetic and conduction problems. *Journal of the Physical Society of Japan*, 12(6):570–586, 1957.
- [8] R. ZWANZIG : Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983–992, 1961.
- [9] H. MORI : Transport, collective motion, and brownian motion. *Progress of Theoretical Physics*, 33:423–455, 1965.
- [10] H. B. CALLEN et T. A. WELTON : Irreversibility and generalized noise. *Phys. Rev. (2)*, 83:34–40, 1951.
- [11] G. GRIMMETT : *Percolation*. Springer, 1999.
- [12] W. WERNER : Percolation et FK-percolation. Notes de cours de DEA à Paris XI, 2005.
- [13] R. DOBRUSHIN et S. SHLOSMAN : Completely analytical interactions : Constructive description. *Journal of Statistical Physics*, 46(5/6):983–1014, 1987.
- [14] C. VILLANI : Hypocoercivity.
<http://www.umpa.ens-lyon.fr/~cvillani/Cedrif/pre.Hypoco.pdf>, 2006.
- [15] A. GUIONNET et B. ZEGARLINSKI : (sur les inégalités log-Sobolev). Notes d'un cours à l'IHP, 1998.
- [16] M. KAC : Foundations of kinetic theory. *In Proceedings of the third Berkeley symposium on mathematical statistics and probability*, volume 3, pages 171–197, 1953.
- [17] H. SPOHN : *Large scale dynamics of interacting particles*. Texts and monographs in Physics. Springer, 1991.
- [18] F. BOLLEY, A. GUILLIN et C. VILLANI : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Theory and Related Fields*, 137:541–593, 2007.
- [19] A.-S. SZNITMAN : Topics in propagation of chaos. *In École d'été de probabilités de Saint-Flour XIX – 1989*, pages 165–261. Springer, 1991.
- [20] C. VILLANI : Optimal transport, old and new.
<http://www.umpa.ens-lyon.fr/~cvillani/Cedrif/B07B.StFlour.pdf>, 2007.
- [21] R. M. DUDLEY : *Uniform central Limit Theorems*. Numéro 63 de Cambridge Studies in advanced mathematics. Cambridge university Press, 1999.

