Probab. Theory Relat. Fields 132, 39–73 (2005)
Digital Object Identifier (DOI) 10.1007/s00440-004-0388-1

E. Breuillard

Distributions diophantiennes et théorème limite local sur $\mathbb{R}^d$

Received: 16 August 2003 / Revised version: 14 July 2004 /
Published online: 9 October 2004 –

Abstract. We study the speed of convergence of $n^{d/2} \int f d\mu^{*n}$ in the local limit theorem on $\mathbb{R}^d$ under very general conditions upon the function f and the distribution μ. We show that this speed is at least of order $\frac{1}{n}$ and we give a simple characterization (in diophantine terms) of those measures for which this speed (and the full local Edgeworth expansion) holds for smooth enough f. We then derive a uniform local limit theorem for moderate deviations under a mild moment assumption. This in turn yields other limit theorems when f is no longer assumed integrable but only bounded and Lipschitz or Hölder. We finally give an application to equidistribution of random walks.

1. Introduction

Soit $S_n = X_1 + \cdots + X_n$ la somme de n variables aléatoires centrées à valeurs dans $\mathbb{R}^d$ qui sont indépendantes et de même loi μ. Le problème du théorème limite local est de préciser le comportement asymptotique quand n tend vers l'infini de l'espérance

$$\mathbb{E}(f(S_n)) = \int f d\mu^n$$

quand f est une fonction définie sur $\mathbb{R}^d$. Lorsque f est la fonction indicatrice d'un intervalle borné I, cette espérance est la probabilité de retour $\mathbb{P}(S_n \in I)$ dans I au temps n. Le théorème limite local (voir [Bre]) affirme que si la loi μ possède un moment d'ordre 2 fini et si son support n'est pas contenu dans une classe d'un sous-groupe fermé propre de $\mathbb{R}^d$ ("non-lattice case"), alors $n^{d/2}\mathbb{E}(f(S_n))$ converge vers l'intégrale de f sur $\mathbb{R}^d$ par rapport à un multiple de la mesure de Lebesgue dès que f est continue et *à support compact*.

Il est naturel de se demander s'il on peut préciser le comportement asymptotique de $\mathbb{E}(f(S_n))$, et en particulier estimer la vitesse de convergence dans le théorème limite local. La question de la vitesse de convergence a été relativement peu abordée dans la littérature. Sous l'hypothèse que la loi μ est absolûment continue par rapport à la mesure de Lebesgue (ou plus généralement sous la condition de Cramér, i.e. le module de la fonction caractéristique de μ reste éloigné de 1 de façon uniforme hors d'un voisinage de 0), on dispose du développement classique

E. Breuillard: DMA, Ecole Normale Supérieure, 45, rue d'Ulm, Paris Ve, France
e-mail: breuillard@ens.fr

de Edgeworth (voir [Fel]), qui fournit tous les termes suivants de l'asymptotique de $\mathbb{P}(S_n \in I)$ (sous réserve de l'existence de moments). Dans le cadre de la théorie du renouvellement, les questions de vitesse ont été étudiées en partie, notamment dans [Car2] où Carlsson obtient, pour le renouvellement sur $\mathbb{R}$, une asymptotique très précise de $\nu(]-\infty, x])$ quand x tend vers $+\infty$, où ν est la mesure de renouvellement. Il montre aussi que si ν vérifie une condition diophantienne, la convergence en est d'autant plus rapide. Mais le cas multi-dimensionnel et en particulier le théorème local (qui n'est qu'un cas particulier du théorème de renouvellement pour la chaîne mixte (S_n, n) en dimension $d+1$) n'a pas fait l'objet d'une étude détaillée à notre connaissance. Il se trouve que la vitesse de convergence dépend de façon significative de la distribution des X_i, et non plus seulement de l'existence de moments. Nous donnons, dans la première partie de cet article, la condition exacte sur la loi μ pour que la vitesse de convergence soit optimale (i.e. en $\frac{1}{n}$), et en définitive pour que le développement de Edgeworth tout entier ait lieu lorsque f est suffisamment différentiable ainsi que sa transformée de Fourier. Cette condition est d'ordre diophantien : elle demande que la variable aléatoire X_1 ne puisse être très bien approximée par une variable à valeurs dans une réunion $H + \mathbb{Z}v$ d'hyperplans affines de $\mathbb{R}^d$. On dira que μ est *diophantienne*. On donne le développement de Edgeworth en en précisant les premiers termes au théorème 3.1. Nous notons que seules des puissances entières de $\frac{1}{n}$ interviennent dans ce développement.

L'autre volet de cet article est consacré à l'étude de l'asymptotique de la suite $\mathbb{E}(f(S_n))$ pour des fonctions f qui ne sont plus nécessairement intégrables sur $\mathbb{R}^d$. Cette étude est motivée par des problèmes d'équidistribution de marches aléatoires. Etant donnée la trajectoire d'un flot issu d'un point x dans un espace X, on montre que si cette trajectoire est équidistribuée pour une certaine mesure finie sur X, alors toute marche aléatoire centrée évoluant le long de cette trajectoire s'équidistribue de la même façon (voir le corollaire 6.1). Dans ce problème, nous avons besoin de pouvoir comparer la marche S_n à la marche gaussienne correspondante sur un domaine plus large que celui fourni par le théorème limite local. C'est l'objet du théorème local de Stone (théorème 4.3) qui donne une estimation uniforme sur une boule dont le rayon est de l'ordre de $\sqrt{n}$. On en déduit en premier lieu une forme très simple du reste dans l'approximation gaussienne en fonction de f et de sa variation totale (voir théorème 5.2). On étend ensuite le théorème de Stone aux écarts modérés, i.e. à un domaine de taille $\sqrt{cn \log n}$, sous l'hypothèse de l'existence d'un moment d'ordre $c+2$ (théorème 4.2). Ceci permet de montrer que lorsque μ est diophantienne et sous une condition de moment, on a bien l'équivalent attendu de $\mathbb{E}(f(S_n))$ pour toute fonction f bornée sur $\mathbb{R}^d$ et suffisamment différentiable. Plus précisément, si f prend des valeurs positives, quand $n \to +\infty$

$$\frac{\mathbb{E}(f(S_n))}{\int_{\mathbb{R}^d} f(x\sqrt{n})p(x)dx} \to 1 \tag{1}$$

où p est la densité de la gaussienne associée (voir théorème 5.3). Lorsque μ satisfait la condition de Cramér, l'équivalent (1) a lieu pour toutes les fonctions höldériennes bornées. Il y a une compétition entre la régularité de μ et celle de f. Sans hypothèse de régularité sur μ ni sur f, cet équivalent est faux. Ainsi pour tout entier $p \geq 0$ il existe des distributions (que l'on peut même choisir diophantiennes) et des fonctions

C^p, intégrables, et dont toutes les dérivées d'ordre $\leq p$ tendent vers 0 à l'infini, pour lesquelles la limite (1) n'a pas lieu (voir l'exemple 5.4). En revanche, sans hypothèse sur μ autre que celles du théorème local cette fois, on montre que si le dénominateur dans (1) tend vers 0, alors le numérateur aussi. Plus généralement, si les moyennes de Cesaro de f sur de grands rectangles convergent, alors la suite $\mathbb{E}(f(S_n))$ tend vers la même limite, pour toute fonction f uniformément continue et bornée (voir théorème 6.1). Il est assez remarquable que l'on puisse déterminer dans ce cas le comportement de $\mathbb{E}(f(S_n))$ sous ces seules hypothèses. L'application à l'équidistribution des marches aléatoires en découle aussitôt.

2. Notations et préliminaires

Bien que la discussion se déroulera le plus souvent (pour alléger les notations) dans le cas plus simple des variables aléatoires réelles, on indiquera aussi dans la dernière section comment généraliser ces résultats au cas multi-dimensionnel.

Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes et de même loi. Dans cet article, on désignera par μ la mesure de probabilité sur $\mathbb{R}$ correspondant à la loi commune des variables X_n. Tout le long de ce texte, on fera l'hypothèse que la mesure μ possède un moment d'ordre 2 fini, c'est-à-dire que $\int_{\mathbb{R}} x^2 d\mu(x) < \infty$.

On notera $S_n = X_1 + \cdots + X_n$ la somme partielle des variables X_i jusqu'au temps n et μ^{*n}, ou simplement μ^n, la loi de la variable aléatoire S_n, à savoir la convoluée n fois de la mesure μ. La convolution des mesures μ et ν est notée $\mu * \nu$. De même μ^{-1} désigne la mesure adjointe à μ : $\mu^{-1}(A) := \mu(-A)$ pour tout borélien A.

On emploiera également dans la suite les notations de théorie de la mesure, comme $\mu(A)$ et $\int f(x)d\mu(x)$, et les notations probabilistes, comme $\mathbb{P}(A)$ et $\mathbb{E}(f(X_1))$, en passant de l'une à l'autre dès que cela semble mieux adapté.

On désignera par $\{x\} = d(x, \mathbb{Z})$ la distance du réel x à un entier le plus proche, entier que l'on notera $[x]$.

On notera $\widehat{f}$ la transformée de Fourier de la fonction f, à savoir

$$\widehat{f}(x) = \int e^{-itx} f(t)dt$$

Aussi pour une mesure de probabilité μ, loi de la variable aléatoire X, on note $\widehat{\mu}$ sa fonction caractéristique $\widehat{\mu}(x) = \mathbb{E}(e^{ixX})$ et on notera σ_p le moment d'ordre p de μ s'il existe, i.e.

$$\sigma_p = \int x^p d\mu(x)$$

La formule d'inversion de Fourier s'écrit alors $\widehat{\widehat{f}}(x) = 2\pi f(-x)$.

On note $\mathbb{L}^1(\mathbb{R})$ l'espace de Banach des classes de fonctions intégrables, muni de la norme usuelle $\|\cdot\|_1$. De même, $\|\cdot\|_\infty$ désigne la norme pour la convergence uniforme des fonctions. On notera $\langle f, g\rangle$ le produit scalaire usuel sur $\mathbb{L}^2(\mathbb{R})$.

On dira que μ est *apériodique* si son support n'est pas contenu dans une progression arithmétique. Cela revient à dire que

$$\forall x \in \mathbb{R}\backslash\{0\} \ |\widehat{\mu}(x)| < 1.$$

On dit que μ satisfait la *condition de Cramér* si elle vérifie l'une des trois conditions équivalentes suivantes

$$\sup_{|x|>1} |\widehat{\mu}(x)| < 1 \Leftrightarrow \lim \inf_{|x|\to+\infty} |1-\widehat{\mu}(x)| > 0 \Leftrightarrow \lim \inf_{|x|\to+\infty} \int \{xa\} d\mu(a) > 0$$

3. Distributions diophantiennes

Dans cette section, on introduit une certaine condition sur la mesure μ qui, lorsqu'elle est satisfaite (et seulement dans ce cas), permet d'obtenir une vitesse de convergence optimale dans le théorème limite local pour les fonctions à support compact suffisamment régulières. Il s'agit en réalité d'une condition diophantienne sur μ, stipulant qu'en moyenne la variable aléatoire X de loi μ ne peut être très bien approchée par les points d'une progression arithmétique réelle. Comme nous l'expliquerons ci-dessous, cette condition est vérifiée dans "la plupart" des cas.

3.1. Définitions

Un nombre réel a est habituellement appelé diophantien si $\{qa\} \geq C/|q|^l$ pour tout entier non nul q et pour un certain $C > 0$ et un certain entier $l \geq 0$ fixés, où, rappelons-le, $\{x\} = d(x, \mathbb{Z})$ désigne la distance du réel x à l'entier le plus proche. Cette définition dépend du choix du réseau $\mathbb{Z}$ des entiers dans $\mathbb{R}$. De façon similaire, sur la droite affine cette fois, on dira qu'un ensemble de points S est **diophantien** s'il est mal approximé par une progression arithmétique réelle, c'est-à-dire plus précisément s'il existe $C > 0$ et $l \geq 0$ tels que

$$\inf_{y\in\mathbb{R}} \sup_{a\in S} \{xa + y\} \geq C/|x|^l$$

pour tout réel x assez grand (on dit alors que S est $2l$ -diophantien). Ainsi le réel a est diophantien si et seulement si le triplet $\{0, 1, a\}$ est diophantien, et tout ensemble de points dont une partie est diophantienne est lui-même diophantien. Par analogie, on a la définition suivante.

Définition 3.1. *Soit μ une mesure de probabilité borélienne sur $\mathbb{R}$ et l un réel ≥ 0. On dit que μ est l-* ***diophantienne*** *s'il existe $C > 0$ tel que pour tout $x \in \mathbb{R}$ assez grand en valeur absolue,*

$$\inf_{y\in\mathbb{R}} \int \{xa + y\}^2 d\mu(a) \geq \frac{C}{|x|^l}$$

On dit que μ est ***diophantienne*** *si elle est l-diophantienne pour un réel $l \geq 0$.*

Cette notion signifie donc que la moyenne du carré des écarts à toute progression arithmétique est minorée par une puissance fixée du pas de la progression. L'introduction du carré permet d'avoir la caractérisation simple ci-dessous, mais ne joue pas un role primordial. Notons aussi que la condition "0-diophantienne" est équivalente à la condition de Cramér énoncée dans la section précédente.

Avant d'aller plus loin, donnons quelques définitions équivalentes de la notion de mesure diophantienne :

Proposition 3.1. *Soit $l \geq 0$. Les assertions suivantes sont équivalentes :*

(i) la mesure μ est l-diophantienne
(ii) il existe un réel $C > 0$ tel que pour tout x assez grand (en valeur absolue), on a

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{|x|^l}$$

(iii) la mesure symétrisée $\mu * \mu^{-1}$ est l -diophantienne
(iv) il existe $C > 0$ tel que pour tout $x \in \mathbb{R}$ assez grand en valeur absolue,

$$\int \{x(a-b)\}^2 d\mu(a) d\mu(b) \geq \frac{C}{|x|^l}$$

Preuve. L'équivalence entre (i) et (iv) est immédiate. Clairement l'assertion (ii) est équivalente à la même assertion pour la mesure symétrisée $\mu * \mu^{-1}$. En tenant compte de cette remarque, on obtient aisément l'équivalence entre (ii) et (iv) si l'on note qu'il existe deux constantes positives c_1 et c_2 telles que

$$c_1\{x\}^2 \leq 1 - \cos(2\pi x) \leq c_2\{x\}^2$$

L'équivalence de (iii) avec les assertions précédentes est évidente au vu de (ii). □

On voit aussi instantanément que μ est diophantienne si et seulement si μ^p est diophantienne pour un entier $p \geq 1$ quelconque. Observons que si μ est diophantienne, alors son support l'est aussi. On note aussi que si la mesure μ n'est pas étrangère à la mesure de Lebesgue, d'après le lemme de Riemann-Lebesgue, μ satisfait la condition de Cramér donc est 0 -diophantienne.

Remarque 3.1. *Notons que l'existence d'un $l > 0$ tel que* $\liminf_{|x|\to+\infty} \left|x^l(1-\widehat{\mu}(x))\right| > 0$ *est en général une condition strictement plus faible que la condition "μ diophantienne", contrairement à ce qui se passe quand $l = 0$. Pour voir cela, il suffit de considérer une mesure supportée par deux points* $\{1, a\}$ *avec a diophantien.*

Supposons maintenant que μ est étrangère à la mesure de Lebesgue. Soit $\Omega(\mu)$ l'ensemble des réels z tels que

$$\lim_{\varepsilon\to 0} \frac{\mu(I_\varepsilon(z))}{2\varepsilon} = +\infty$$

où $I_\varepsilon(z)$ représente l'intervalle ouvert de longueur 2ε centré au point z. On sait que $\mu(\Omega(\mu)) = 1$ (voir par exemple [Rud]). On peut alors énoncer la

Proposition 3.2. *Supposons que μ soit étrangère à la mesure de Lebesgue. S'il existe un ensemble fini S inclus dans $\Omega(\mu)$ qui soit diophantien, alors μ est diophantienne. Si μ est à support fini, alors μ est l-diophantienne si et seulement si support S de μ est l-diophantien.*

Preuve. Soit X une variable aléatoire de loi μ. Soit $l \geq 0$ et $C > 0$ tels que

$$\inf_{y \in \mathbb{R}} \max_{s \in S} \{sx + y\} \geq C/|x|^l$$

pour tout x assez grand. Comme $S \subset \Omega(\mu)$ est fini, pour tout x assez grand et $s \in S$, $\mathbb{P}(|X - s| < C/2|x|^{l+1}) > 1/|x|^{l+1}$. Pour x grand et $y \in \mathbb{R}$, on peut trouver $s \in S$ tel que $\{sx + y\} \geq C/|x|^l$. Donc si $|X - s| < C/2|x|^{l+1}$, alors $\{xX + y\} \geq C/2|x|^l$. D'où

$$\mathbb{E}(\{xX + y\}^2) \geq \mathbb{P}(|X - s| < C/2|x|^{l+1}) \frac{C^2}{4|x|^{2l}} \geq \frac{C^2}{4|x|^{3l+1}}.$$

Le reste de la proposition est immédiat. □

Rappelons que pour presque tout choix de réels a, b et c, par rapport à la mesure de Lebesgue, le triplet $\{a, b, c\}$ est diophantien. Cependant, il est facile de construire un exemple de mesure μ à support dans $[0, 1]$, qui soit à la fois apériodique et non diophantienne, et même dont le support soit diophantien. Dans le premier exemple, la mesure μ est atomique de support $[0, 1]$, dans le second, elle est diffuse et son support contient $[0, 1]$.

Exemple 3.1. Soit E_n, n entier ≥ 1, l'ensemble des rationnels r dont l'écriture en fraction irréductible est $r = \frac{p}{n}$ pour un certain entier p premier à n et compris entre 1 et n. Soit μ une mesure de probabilité qui assigne à E_n le poids $\frac{1}{2^{n!}}$ si $n \geq 2$ et assigne un même poids aux points de E_n et X_1, X_2 des variables aléatoires indépendantes de loi μ. Clairement

$$\mathbb{E}(\{n!(X_1 - X_2)\}^2) \leq 2\mathbb{P}(n!X_1 \notin \mathbb{Z}) \leq 2\textstyle\sum_{p>n} 1/2^{p!} \leq 1/2^{n!}$$

donc μ n'est pas diophantienne mais son support est l'intervalle $[0, 1]$ tout entier.

Exemple 3.2. Soit $(n_i)_{i \geq 1}$ une suite d'entiers ≥ 1 telle que $n_{i+1} \geq 2^{n_i}$. Soit $E = \{n_i\}_i$ la partie de $\mathbb{N}$ correspondante et K l'ensemble des réels x qui s'écrivent

$$x = \sum_{i \in I} \frac{1}{2^{n_i}}$$

pour un choix quelconque d'une partie I de l'ensemble E, la somme valant 0 si $I = \emptyset$. On voit que K est un compact inclus dans $[0, 1]$ sans point isolé, totalement discontinu et de mesure de Lebesgue nulle (Cantor). Il est possible de construire une fonction continue f croissante de l'intervalle $[0, 1]$ dans lui-même, avec $f(0) = 0$ et $f(1) = 1$, telle que la dérivée f' est nulle en tout point de $[0, 1] \backslash K$ et $+\infty$ en tout point de K. La mesure de Stieljes associée μ est une mesure de probabilité diffuse et supportée par le compact K. Comme K n'est pas diophantien, on voit que μ n'est pas diophantienne. On peut aussi construire une mesure diffuse non-diophantienne dont le support contient tout l'intervalle $[0, 1]$ (par exemple en prenant la somme $c \cdot \sum \mu * \delta_{x_n}/2^{k_n}$ où $(x_n)_n$ est une suite dense dans $[0, 1]$ convenablement choisie et $k_n \uparrow +\infty$, $c > 0$).

3.2. Une classe de fonctions analytiques de type exponentiel

Dans ce paragraphe, on introduit une classe d'exemples de fonctions sur $\mathbb{R}$ qui nous seront utiles dans les paragraphes suivants, en particulier en tant que source de contre-exemples. Il s'agit de fonctions analytiques sur $\mathbb{R}$ dont la transformée de Fourier est à support compact.

Soit $(a_n)_{n\in\mathbb{Z}}$ une suite bornée quelconque de nombres complexes et p un entier ≥ 1. On lui associe la fonction d'une variable complexe $\beta(z)$ définie comme suit

$$\beta(z) = \frac{1}{2^{2p}} \sin^{2p}(\pi z) \sum_{n\in\mathbb{Z}} \frac{a_n}{(z-n)^{2p}}.$$

C'est une fonction holomorphe sur $\mathbb{C}$ qui vérifie $|\beta(z)| \leq Ce^{2p\pi|z|}$ où C est une constante dépendant de $\sup_{n\in\mathbb{Z}} |a_n|$. Donc β est entière et de type exponentiel. Restreinte à la droite réelle, β est analytique et bornée.

Remarquons d'abord que si la suite $(a_n)_n$ est dans $\ell^1(\mathbb{Z})$ alors $\beta \in \mathbb{L}^1(\mathbb{R})$. Il résulte alors du théorème de Paley-Wiener que la transformée de Fourier $\widehat{\beta}(t)$ sur $\mathbb{R}$ est une fonction continue à support compact inclus dans $[-2p\pi, 2p\pi]$ et

$$\int \widehat{\beta}(t)dt = 2\pi\beta(0) = 2\pi\left(\frac{\pi}{2}\right)^{2p} a_0.$$

Supposons maintenant que la suite $(a_n)_n$ soit un $O(1/n^{2p})$ au voisinage de l'infini. Alors, on voit aisément que $\beta(x)$ est elle-même un $O(1/\,|x|^{2p-1})$ au voisinage de l'infini, lorsque $x \in \mathbb{R}$. La transformée de Fourier $\widehat{\beta}$ est donc au moins de classe C^l dès que $2(p-1) > l$.

Revenons au cas où la suite $(a_n)_n$ est seulement bornée et supposons de plus que $a_n \geq 0$ pour tout n. Alors on a

$$\beta(x) \geq a_{[x]}$$

où $[x]$ désigne l'entier le plus proche du réel x. En effet, $\beta(x) \geq a_{[x]} \left(\frac{\sin \pi x}{2\{x\}}\right)^{2p} \geq a_{[x]}$.

3.3. Développements de Edegeworth locaux pour les mesures diophantiennes

Lorsque μ est diophantienne et admet un moment d'ordre $r+2$, alors $\sqrt{n}\mathbb{E}(f(S_n))$ admet un développement asymptotique avec un reste de l'ordre de $o(1/n^{r/2})$ dès que f est assez régulière. C'est l'objet des deux théorèmes qui suivent. Ceux-ci généralisent aux mesures diophantiennes le développement de Edgeworth classique pour les densités qui est valide sous l'hypothèse que $\widehat{\mu}$ est intégrable (voir [Fel]).

On pose, pour une fonction numérique f

$$C_r(f) = \max_{0\leq j\leq r+1} \left\|x^j f\right\|_1, \; C^k(f) = \max_{0\leq j\leq k} \left\|f^{(j)}\right\|_1, \; C_r^k(f) = C^k(f) + C_r(f)$$

Théorème 3.1. *Soit r un entier ≥ 0. Supposons que la mesure de probabilité μ est centrée et possède un moment d'ordre $r+2$ fini. Si de plus μ est l-diophantienne pour $l \geq 0$, alors pour toute fonction f de classe C^k avec $k > l(r+1)/2+1$ et telle que $C_r^k(f) < +\infty$, on a le développement asymptotique (de Edgeworth) suivant*

$$\sqrt{n}\mathbb{E}(f(S_n)) = \sum_{p=0}^{[r/2]} \frac{1}{n^p} \langle f, Q_p \rangle_{L^2(\mathbb{R})} + C_r^k(f) \cdot o\Big(\frac{1}{n^{r/2}}\Big)$$

où le $o()$ ne dépend que de r et de la loi μ et les Q_p sont des polynômes de degré $\leq 2p$ dont les coefficients dépendent des moments de μ d'ordre $\leq p+2$.

Notons que seules des puissances entières interviennent dans le développement (comparer avec Feller [Fel] chap. 16 (2.12) et (4.10), au (2.13) il faut en fait lire H_6 au lieu de H_3 dans l'expression de P_4). De plus on calcule aisément les premiers termes. Ainsi

$$\sqrt{2\pi\sigma_2}Q_0(X) = 1, \sqrt{2\pi\sigma_2}Q_1(X) = -\frac{1}{\sigma_2}X^2 - \frac{\sigma_3}{\sigma_2^2}X + \frac{1}{8}\left(\frac{\sigma_4}{\sigma_2^2} - 3\right) - \frac{5}{24}\frac{\sigma_3^2}{\sigma_2^3}$$

Preuve. On écrit

$$\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) = \sqrt{n}\sqrt{\frac{\sigma_2}{2\pi}} \int \widehat{f}(x)\widehat{\mu}(x)^n dx \tag{2}$$

La preuve reprend les techniques de Fourier classiques exposées dans [Fel]. On commence par traiter dans le lemme suivant la partie de l'intégrale correspondant aux grandes valeurs de x. Ce lemme nous sera aussi utile plus tard.

Lemme 3.1. *Soit μ une mesure de probabilité sur $\mathbb{R}$. On suppose que μ est l-diophantienne. Soit $r > 0$ un réel positif. Alors il existe un réel $D(r,\mu) > 0$ tel que pour tout $D > D(r,\mu)$, on a*

$$\int_{|t| \geq \sqrt{\frac{D\log n}{n}}} \widehat{f}(t)\widehat{\mu}^n(t)dt = C^k(f) \cdot o_\mu\Big(\frac{1}{n^r}\Big)$$

uniformément pour toute fonction f de classe C^k telle que $C^k(f) < +\infty$, où $k > lr+1$.

Preuve. Si $l = 0$ la loi μ satisfait la condition de Cramér, donc le lemme est clairement vrai. Supposons $l > 0$. Puisque μ est apériodique (car diophantienne) et $\widehat{\mu}$ continue, pour tout $\varepsilon > 0$ et $x_0 > 0$, il existe c_0, $0 < c_0 < 1$, tel que $|\widehat{\mu}(x)| \leq c_0$ quel que soit x avec $\varepsilon < |x| < x_0$. D'où

$$\left|\int_{\varepsilon<|x|<x_0} \widehat{f}(x)\widehat{\mu}(x)^n dx\right| \leq c_0^n \left\|\widehat{f}\right\|_\infty x_0 \leq c_0^n C^k(f) x_0.$$

Puisque μ est l-diophantienne, il existe un réel x_0 tel que si $|x| \geq x_0$ alors

$$|\widehat{\mu}(x)| \leq \exp(-C/|x|^l)$$

pour une certaine constante $C > 0$. Comme f est C^k et $C^k(f) < +\infty$, on a $\left|\widehat{f}(x)\right| \leq \left\|f^{(k)}\right\|_1 / |x|^k$. Il vient

$$\left|\int_{|x|>x_0} \widehat{f}(x)\widehat{\mu}(x)^n dx\right| \leq \int_{|x|>x_0} \frac{C^k(f)}{|x|^k} \exp(-Cn/|x|^l) dx \tag{3}$$
$$\leq C^k(f) \frac{C'}{n^{(k-1)/l}} \int_0^{+\infty} \frac{e^{-u}}{u^{(l+1-k)/l}} du$$

où C' est une constante indépendante de f. En choisissant $k > lr + 1$, la quantité ci-dessus est au plus de l'ordre de $C^k(f) \cdot o(1/n^r)$ quand n croît.

Passons maintenant à la partie de l'intégrale où $|x| < \varepsilon$. Puisque μ a un moment d'ordre 2 fini, on peut trouver $\varepsilon > 0$ et $c > 0$ tels que si $|x| < \varepsilon$, on a $|\widehat{\mu}(x)| \leq \exp(-cx^2)$. Fixons $D > r/c$. Alors si $\sqrt{D \log n} < |x| < \varepsilon\sqrt{n}$, on a

$$\left|\widehat{\mu}\left(\frac{x}{\sqrt{n}}\right)^n\right| \leq \exp(-cD\log n) \leq \frac{1}{n^{cD}} = o\left(\frac{1}{n^r}\right).$$

Il vient

$$\left|\int_{\sqrt{D\log n}<|x|<\varepsilon\sqrt{n}} \widehat{f}\left(\frac{x}{\sqrt{n}}\right)\widehat{\mu}\left(\frac{x}{\sqrt{n}}\right)^n \frac{dx}{\sqrt{n}}\right| \leq C^k(f) \cdot o\left(\frac{1}{n^r}\right)$$

□

On peut donc maintenant se concentrer sur la partie de l'intégrale où $|x| < \sqrt{D \log n}$. On peut écrire au voisinage de zéro

$$\widehat{f}(x) = \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!} x^j + \frac{x^{r+1}}{(2r+1)!}\phi(x)$$

où ϕ est une fonction bornée par $\left\|\widehat{f}^{(r+1)}\right\|_\infty \leq C_r^k(f)$. Il vient

$$\int_{|x|<\sqrt{D\log n}} \widehat{f}\left(\frac{x}{\sqrt{n}}\right)\widehat{\mu}\left(\frac{x}{\sqrt{n}}\right)^n dx = \sum_{j=0}^{r} \frac{\widehat{f}^{(j)}(0)}{j!n^{j/2}} \Phi_j + \frac{1}{n^{(r+1)/2}} \Phi_{r+1} \tag{4}$$

où l'on a noté

$$\Phi_j = \int_{|x|<\sqrt{D\log n}} x^j \widehat{\mu}\left(\frac{x}{\sqrt{n}}\right)^n dx$$

et

$$\Phi_{r+1} = \frac{1}{(r+1)!} \int_{|x|<\sqrt{D\log n}} x^{r+1} \phi\left(\frac{x}{\sqrt{n}}\right)\widehat{\mu}\left(\frac{x}{\sqrt{n}}\right)^n dx$$
$$|\Phi_{r+1}| \leq C_r^k(f) \int |x|^{r+1} e^{-cx^2} dx$$

Comme μ est centrée et possède un moment d'ordre $r+2$ fini, il existe une fonction ψ_1 bornée continue et valant 0 en 0, telle qu'au voisinage de 0

$$\log \widehat{\mu}(x) = x^2\Big(-\frac{\sigma_2}{2!} + P_r(x)\Big) + x^{r+2}\psi_1(x)$$

où P_r est un polynôme de degré $\leq r$ tel que $P_r(0) = 0$ dont les coefficients dépendent seulement des moments de μ jusqu'à l'ordre $r+2$. Il vient,

$$\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n = e^{-\frac{\sigma_2 x^2}{2}} \exp\left(x^2 P_r\Big(\frac{x}{\sqrt{n}}\Big) + \frac{x^{r+2}}{n^{r/2}}\psi_1\Big(\frac{x}{\sqrt{n}}\Big)\right) \tag{5}$$

Si $r = 0$, on obtient directement le résultat du théorème en appliquant le théorème de convergence dominée de Lebesgue. Si $r \geq 1$, pour tout x tel que $|x| < \sqrt{D\log n}$, l'expression ci-dessus à l'intérieur de exp est un $O(\log^{3/2}(n)/\sqrt{n})$, d'où

$$\begin{aligned} e^{\frac{\sigma_2 x^2}{2}}\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n = &\sum_{m=0}^{r}\frac{1}{m!}\left[x^2 P_r\Big(\frac{x}{\sqrt{n}}\Big)\right]^m + \frac{x^{r+2}}{n^{r/2}}\psi_1\Big(\frac{x}{\sqrt{n}}\Big) \\ &+O_{\mu,r,D}\Big(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}}\Big) \end{aligned}$$

En regroupant les termes de la somme ci-dessus qui sont en facteur de $1/n^{i/2}$, on obtient des polynômes A_i (de degré au plus $3i$) tels que

$$e^{\frac{\sigma_2 x^2}{2}}\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n = \sum_{i=0}^{r}\frac{1}{n^{i/2}}A_i(x) + \frac{x^{r+2}}{n^{r/2}}\psi_1\Big(\frac{x}{\sqrt{n}}\Big) + O_{\mu,r,D}\left(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}}\right) \tag{6}$$

En tenant compte de (6), l'expression (4) devient

$$\sum_{i=0}^{r}\sum_{j=0}^{r}\frac{1}{n^{(i+j)/2}}\frac{\widehat{f}^{(j)}(0)}{j!}\int_{|x|<\sqrt{D\log n}} x^j A_i(x)e^{-\frac{\sigma_2 x^2}{2}}dx + C_r(f)\cdot o_{\mu,r,D}\Big(\frac{1}{n^{r/2}}\Big)$$

En prenant D assez grand, on peut remplacer l'intégrale ci-dessus par une intégrale sur $\mathbb{R}$ tout entier. Notons maintenant que le polynôme A_i a même parité que i et dépend des moments de μ jusqu'à l'ordre $i+2$. Donc lorsque $i+j$ est impair,

$$\int_{\mathbb{R}} x^j A_i(x)e^{-\frac{\sigma_2 x^2}{2}}dx = 0$$

Il ne reste plus que des puissances entières dans le développement. Grâce au lemme ci-dessus, la preuve du théorème est maintenant complète. □

Le théorème qui suit donne l'analogue pour les mesures diophantiennes du développement asymptotique classique pour les densités donné dans [Fel]. La suite de polynômes (P_p) ci-dessous est celle de [Fel] XVI Theorem 2 (pour le calcul des premiers termes, on renvoie à [Cra2]) :

Théorème 3.2. *Soit r un entier ≥ 0 et l un réel ≥ 0. Supposons que la mesure de probabilité μ est centrée et possède un moment d'ordre $r+2$ fini. Si de plus μ est l-diophantienne, alors pour toute fonction f de classe C^k avec $k > l(r+1)/2+1$ et telle que $C^k(f) < +\infty$, on a le développement asymptotique (de Edgeworth) suivant*

$$\mathbb{E}(f(S_n)) = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{\mathbb{R}} f(x\sqrt{n}) P_p(x) g(x) dx + C^k(f) \cdot o\Big(\frac{1}{n^{(r+1)/2}}\Big)$$

où le $o()$ ne dépend que de r et de la loi μ, où g est la densité de la gaussienne associée à μ dans le théorème de la limite centrale et où les P_p sont des polynômes de degré $\leq 3p$ dont les coefficients ne dépendent que des moments de μ jusqu'à l'ordre $p+2$.

La preuve reprend les mêmes étapes que celles du théorème précédent. En particulier, elle résulte aisément de la combinaison du lemme 3.1 avec le lemme suivant :

Lemme 3.2. *Soit μ une mesure centrée sur $\mathbb{R}$ admettant un moment d'ordre $r+2$. Soit D un réel assez grand ($D > r/\sigma_2$ par exemple suffit). Alors on a, uniformément pour toute fonction intégrable f,*

$$\int_{|t|\leq\sqrt{\frac{D\log n}{n}}} \widehat{f}(t)\widehat{\mu}^n(t)dt = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int f(x\sqrt{n}) P_p(x) g(x) dx + \|f\|_1 \cdot o_{r,D,\mu}\Big(\frac{1}{n^{(r+1)/2}}\Big)$$

où l'on a gardé les notations du théorème 3.2.

Preuve. On peut reprendre l'estimée (6) qui s'écrit, sous les hypothèses du lemme :

$$\widehat{\mu}\Big(\frac{t}{\sqrt{n}}\Big)^n = e^{-\frac{\sigma_2 t^2}{2}} \sum_{p=0}^{r} \frac{1}{n^{p/2}} A_p(t) + \frac{t^{r+2}}{n^{r/2}} e^{-\frac{\sigma_2 t^2}{2}} \psi_1\Big(\frac{t}{\sqrt{n}}\Big) + O\left(\frac{\log^{3(r+1)/2}(n)}{n^{(r+1)/2}}\right) e^{-\frac{\sigma_2 t^2}{2}}$$

Il vient alors, comme $\widehat{g}(t) = e^{-\frac{\sigma_2 t^2}{2}}$,

$$\int_{|t|\leq\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big)\widehat{\mu}^n\Big(\frac{t}{\sqrt{n}}\Big)dt = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{|t|\leq\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big) A_p(t)\widehat{g}(t)dt + o\Big(\frac{1}{n^{r/2}}\Big)\|f\|_1$$

On sait que le degré de A_p ne dépasse pas $3p$, donc d'après le choix de D, pour tout $p \leq r$

$$\left|\int_{|t|\geq\sqrt{D\log n}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big) A_p(t)\widehat{g}(t)dt\right| \leq \|f\|_1\, o_r\Big(\frac{1}{n^{r/2}}\Big)$$

On peut donc substituer l'intégrale sur $|t| \leq \sqrt{D \log n}$ ci-dessus par une intégrale sur tout $\mathbb{R}$. Mais on a $t^p \widehat{g}(t) = (-i)^p \widehat{g^{(p)}}(t)$ pour tout entier $p \geq 0$, d'où

$$\int_{\mathbb{R}} \widehat{f}\Big(\frac{t}{\sqrt{n}}\Big) A_p(t) \widehat{g}(t) dt = \sqrt{n} \int_{\mathbb{R}} f(x\sqrt{n}) P_p(x) g(x) dx$$

où $P_p(x)$ est le polynôme tel que $2\pi A_p(-i\frac{d}{dx})g(x) = P_p(x)g(x)$. On a donc bien le résultat souhaité. □

3.4. Caractérisation des mesures diophantiennes par la vitesse de convergence

Dans ce paragraphe, on démontre que, pour une mesure de probabilité μ centrée, la convergence dans le théorème limite local est en général plus lente que ou égale à $O(1/n)$ pour une fonction f à support compact fixée. On montre aussi que cet ordre de grandeur est atteint pour des fonctions suffisamment régulières si et seulement si la mesure μ est diophantienne. On va démontrer ainsi une réciproque au théorème 3.1. Plus précisément, on a la caractérisation suivante des mesures diophantiennes :

Théorème 3.3. *Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, et ayant un moment d'ordre 3 fini. Soit ν la mesure gaussienne qui lui est associée par le théorème de la limite centrale (i.e. μ est dans le domaine d'attraction de ν). Alors μ est diophantienne si et seulement s'il existe un entier $k_0 \geq 0$ tel que pour toute fonction f à support compact et de classe C^k $(k \geq k_0)$ sur $\mathbb{R}$, on ait*

$$\sup_{t \in \mathbb{R}} \left| \int f(t+\cdot) d\mu^n - \int f(t+\cdot) d\nu^n \right| = O\Big(\frac{1}{n}\Big)$$

où $f(t+\cdot)$ désigne la translatée de f par le réel t et où O dépend de μ et de f.

Preuve. Supposons d'abord μ diophantienne. Elle est donc apériodique. On peut reprendre à l'identique le début de la preuve du théorème (3.1). Notons pour simplifier $f_t = f(t+\cdot)$. On remarque d'abord que $\widehat{f_t}(x) = e^{ixt}\widehat{f}(x)$. On obtient successivement pour $\varepsilon > 0$ et $x_0 > 0$ comme dans la preuve ci-dessus

$$\left| \int_{\varepsilon < |x| < x_0} \widehat{f_t}(x) \widehat{\mu}(x)^n dx \right| \leq c^n \, \|f\|_1$$

pour $c \in]0,1[$

$$\left| \int_{|x| > x_0} \widehat{f_t}(x) \widehat{\mu}(x)^n dx \right| \leq \sup_{0 \leq p \leq k} \left\| f^{(p)} \right\|_1 \frac{cste}{n^{(k-1)/l}}$$

et

$$\left| \int_{\sqrt{D \log n} < |x| < \varepsilon\sqrt{n}} \widehat{f_t}\Big(\frac{x}{\sqrt{n}}\Big) \widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n dx \right| \leq \|f\|_1 \cdot o\Big(\frac{1}{n}\Big) \tag{7}$$

Ces estimations sont valides uniformément en $t \in \mathbb{R}$ et sont aussi satisfaites par ν à la place de μ. On peut donc se concentrer sur le terme

$$\int_{|x| < \sqrt{D \log n}} \widehat{f_t}\Big(\frac{x}{\sqrt{n}}\Big) \left(\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n - e^{-\frac{\sigma_2 x^2}{2}} \right) dx$$

que l'on majore uniformément en t par

$$\|f\|_1 \int_{|x|<\sqrt{D\log n}} \left|\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n - e^{-\frac{\sigma_2 x^2}{2}}\right| dx$$

Mais, puisque μ a un moment d'ordre 3 fini, on peut écrire

$$\widehat{\mu}(x) = \exp\Big(-\frac{\sigma_2}{2!}x^2 - i\frac{\sigma_3}{3!}x^3 + x^3\psi_0(x)\Big)$$

pour une certaine fonction ψ_0 qui tend vers 0 quand x tend vers 0. Alors lorsque $|x| < \sqrt{D\log n}$,

$$\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n - e^{-\frac{\sigma_2 x^2}{2}} = \frac{1}{\sqrt{n}}e^{-\frac{\sigma_2 x^2}{2}}\left(-i\frac{\sigma_3}{3!}x^3 + x^3\psi_1\Big(\frac{x^3}{\sqrt{n}}, \frac{x}{\sqrt{n}}\Big)\right) \tag{8}$$

pour une certaine fonction ψ_1, avec $\psi_1(u) \to 0$ si $u \to 0$. Il vient

$$\int_{|x|<\sqrt{D\log n}} \left|\widehat{\mu}\Big(\frac{x}{\sqrt{n}}\Big)^n - e^{-\frac{\sigma_2 x^2}{2}}\right| dx \le \frac{c_0(\mu)}{\sqrt{n}} \tag{9}$$

où $c_0(\mu)$ est une constante ne dépendant que de μ. Donc il vient

$$\sup_{t\in\mathbb{R}}\left|\int f(t+\cdot)d\mu^n - \int f(t+\cdot)d\nu^n\right| \le \frac{c(\mu)}{n}\left(\|f\|_1 + \sup_{0\le p\le k}\left\|f^{(p)}\right\|_1\right)$$

pour une certaine constante $c(\mu)$ ne dépendant que de μ.

Maintenant, passons à la réciproque. Nous allons d'abord établir la proposition suivante qui concerne le cas particulier des mesures symétriques :

Proposition 3.1. *Soit ν une mesure de probabilité sur $\mathbb{R}$ qui possède un moment d'ordre 3 fini. Soit $\mu = \nu * \nu^{-1}$. On suppose de plus que μ n'est pas diophantienne. Soit S_n la marche aléatoire associée à μ. Alors pour tout entier $p \ge 0$ et tout $\varepsilon > 0$, il existe une fonction f de classe C^p et à support compact telle que*

$$\lim_{n\to+\infty}\sup n^{\varepsilon}\left|\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) - \int f\right| > 0$$

Preuve. On fixe l'entier p et on considère la fonction β du paragraphe 3.2 associée à la suite $(a_n)_{n\in\mathbb{Z}}$ définie par $a_0 = 0$ et $a_n = \frac{1}{|n|^p}$ pour $n \neq 0$. On prend $f = \widehat{\beta}$. C'est une fonction de classe C^p à support compact et d'intégrale nulle. De plus $\widehat{f}(x) = 2\pi\beta(x)$ est réel et $> \frac{1}{[x]^p}$ pour tout x tel que $[x] \neq 0$. Comme dans la preuve du théorème ci-dessus, on écrit

$$\sqrt{n}\sqrt{2\pi\sigma_2}\mathbb{E}(f(S_n)) = \sqrt{n}\sqrt{\frac{\sigma_2}{2\pi}}\int \widehat{f}(x)\widehat{\mu}(x)^n dx$$

Comme ci-dessus (μ a un moment d'ordre 3), on voit que

$$\int_{|x|\le\varepsilon} \widehat{f}(x)\widehat{\mu}(x)^n dx = o\Big(\frac{1}{n}\Big)$$

Maintenant, fixons un entier $l \geq \frac{p}{\varepsilon}$. Remarquons que $\widehat{\mu}(x) \geq 0$, car on a supposé $\mu = \nu * \nu^{-1}$. Puisque μ n'est pas diophantienne, il existe une suite non bornée $(x_m)_m$ telle que

$$\widehat{\mu}(x_m) \geq e^{-\frac{1}{|x_m|^l}}$$

Montrons que chaque x_m est inclus dans un intervalle I_m de longueur au moins $\frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}$ tel que tout $x \in I_m$ vérifie $\widehat{\mu}(x) \geq e^{-\frac{4}{|x_m|^l}}$. En effet, supposons que ce ne soit pas le cas, alors $\widehat{\mu}(x)$ admet un maximum dans l'intervalle $]a_m, b_m[$ où $b_m - a_m \leq \frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}$ et $x_m \in]a_m, b_m[$ et $\widehat{\mu}(a_m) \leq e^{-\frac{4}{|x_m|^l}}$, $\widehat{\mu}(b_m) \leq e^{-\frac{4}{|x_m|^l}}$. Soit y_m l'abscisse de ce maximum. Alors

$$\widehat{\mu}(x) \geq \widehat{\mu}(y_m) - \frac{\sigma_2}{2}|x - y_m|^2$$

car la dérivée seconde $\widehat{\mu}(x)''$ est bornée par $\sigma_2 = \mathbb{E}(X^2)$. D'où $\widehat{\mu}(a_m) \geq e^{-\frac{1}{|x_m|^l}} - \frac{1}{|x_m|^l} \geq e^{-\frac{3}{|x_m|^l}}$, dès que m est assez grand. Ce qui fournit une contradiction.

Finalement, il vient

$$\begin{aligned}\int_{|x|\geq\varepsilon} \widehat{f}(x)\widehat{\mu}(x)^n dx &\geq \int_{x\in I_m} \widehat{f}(x)\widehat{\mu}(x)^n dx \\ &\geq \frac{1}{[x_m]^p} e^{-\frac{4n}{|x_m|^l}} \frac{\sqrt{2/\sigma_2}}{|x_m|^{l/2}}\end{aligned}$$

En prenant $n = [x_m]^l$, on obtient

$$\sqrt{n}\int \widehat{f}(x)\widehat{\mu}(x)^n dx \geq \frac{C}{n^{p/l}} \geq \frac{C}{n^{\varepsilon}}$$

pour une certaine constante $C > 0$. □

Passons maintenant au cas général. La mesure μ n'est pas diophantienne, mais est centrée et possède un moment d'ordre 3 fini. Alors la mesure symétrisée $\widetilde{\mu} = \mu * \mu^{-1}$ non plus n'est pas diophantienne, mais elle possède aussi un moment d'ordre 3 fini. D'après la proposition ci-dessus, quel que soit $\varepsilon > 0$ et quel que soit l'entier $k \geq 0$, on peut trouver une fonction f à support compact et de classe C^k sur $\mathbb{R}$, que l'on choisit aussi d'intégrale nulle, telle que

$$\lim_{n\to+\infty} \sup n^{\frac{1}{2}+\varepsilon}\left|\int f d\widetilde{\mu}^n\right| > 0 \tag{10}$$

Or si μ satisfait la conclusion du théorème (3.3), alors

$$\sup_{t\in\mathbb{R}}\left|\int f(t+\cdot)d\mu^n\right| = O\left(\frac{1}{n}\right)$$

car, puisque f est d'intégrale nulle, on a automatiquement

$$\sup_{t\in\mathbb{R}}\left|\int f(t+\cdot)d\nu^n\right| \le \|xf\|_1\, O\Big(\frac{1}{n}\Big)$$

Mais alors

$$\left|\int f d\widetilde{\mu}^n\right| = \left|\int \mathbb{E}(f(S_n-t))d\mu^n(t)\right| \le \sup_{t\in\mathbb{R}}|\mathbb{E}(f(S_n-t))| = O\Big(\frac{1}{n}\Big)$$

ce qui contredit (10) dès que $\varepsilon < \frac{1}{2}$. □

On peut construire des mesures non diophantiennes qui fournissent des vitesses de convergence aussi lentes que l'on veut. A cet égard, on renvoie à l'exemple décrit en (5.4) plus bas.

4. Théorème local pour les écarts modérés

4.1. Loi des écarts modérés

Le développement asymptotique précédent permet d'obtenir simplement un résultat de grandes déviations valable sans faire l'hypothèse habituelle d'existence d'un moment exponentiel. Ce résultat, qui traite des écarts modérés (i.e. du comportement de S_n dans un domaine de taille $\sqrt{cn\log n}$), a été démontré par Rubin et Sethuraman dans [RuS] (voir aussi [Nag2, Nag3], [Mic] et [Sla]). Cependant comme nous en aurons besoin dans la suite de cet article et que l'argument que nous présentons est très différent de l'argument original et se généralise automatiquement en dimension supérieure, nous choisissons de l'écrire ici.

Théorème 4.1 ([RuS]). *Soit r un entier ≥ 1. Supposons que μ est une mesure de probabilité centrée, de variance $\sigma_2 > 0$, qui possède un moment d'ordre $r+2$ fini. Alors pour tout c, $0 < c < r$, on a uniformément en x pour $1 \le x \le \sqrt{c\log n}$*

$$\lim_{n\to+\infty}\frac{\mathbb{P}\left(|S_n| > x\sqrt{\sigma_2 n}\right)}{\mathbb{P}(|N| > x)} = 1 \tag{11}$$

où N est une variable gaussienne normalisée. En particulier

$$\mathbb{P}\left(|S_n| > \sqrt{c\sigma_2 n\log n}\right) \sim \frac{4}{n^{c/2}\sqrt{2\pi c\log n}}$$

Preuve. On peut supposer $\sigma_2 = 1$. La preuve qui suit illustre bien le principe classique de lissage et "délissage" (voir par exemple [Sto3]). Supposons d'abord que la fonction caractéristique $\widehat{\mu}$ est intégrable. Dans ce cas la variable $S_n/\sqrt{n}$ admet une densité continue f_n et l'on dispose du développement du théorème 3.2 qui n'est autre que le développement de Edgeworth classique pour les densités (cf. [Cra2] ou [Fel] XVI, 2) et s'écrit :

$$f_n(t) - g(t) = \sum_{p=1}^{r}\frac{1}{n^{p/2}}P_p(t)g(t)dt + o_{r,\mu}\Big(\frac{1}{n^{r/2}}\Big)$$

où g est la densité gaussienne normalisée. En intégrant sur $|t| < x$, il vient

$$\mathbb{P}\left(|N| > x\right) - \mathbb{P}\left(|S_n^{\mu}| > x\sqrt{n}\right) = K_n^{\mu}(x) + o\Big(\frac{\sqrt{\log n}}{n^{r/2}}\Big) \tag{12}$$

où $K_n^{\mu}(x) = \sum_{p=1}^{r} \int_{|t|<x} P_p g/n^{p/2}$ ne dépend que des moments de μ jusqu'à l'ordre $r+2$. Or on peut trouver une autre probabilité μ' dont la fonction caractéristique est intégrable, qui possède un moment exponentiel fini et dont tous les moments jusqu'à l'ordre $r+2$ coïncident avec ceux de μ. Pour μ' on dispose alors de la théorie des grandes déviations classiques due à Cramér. En particulier la limite uniforme (11) a bien lieu pour S_n' à la place de S_n et on a (voir [Cra1] ou [Fel] XVI, 6):

$$\mathbb{P}\left(|N| > \sqrt{c\log n}\right) \sim \frac{4}{n^{c/2}\sqrt{2\pi c\log n}} = o\Big(\frac{1}{n^{c/2}}\Big)$$

Mais puisque (12) est aussi valide pour μ' et que $K_n^{\mu} = K_n^{\mu'}$ d'après le choix de μ' , on obtient uniformément en x pour $1 \leq x \leq \sqrt{c\log n}$

$$\lim_{n\to+\infty} \frac{K_n^{\mu}(x)}{\mathbb{P}\left(|N| > x\right)} = 0. \tag{13}$$

En revenant à (12) pour μ cette fois, on obtient bien la limite (11) souhaitée. Ceci termine la preuve dans le cas où $\widehat{\mu}$ est intégrable.

Dans le cas général, il faut reprendre l'argument précédent pour la famille de mesures $\mu_\varepsilon = \mu * g_\varepsilon$ où g_ε est la gaussienne d'écart type ε. On remarque d'abord que l'estimation (12) ci-dessus reste valable uniformément quand appliquée aux mesures μ_{ε_n} si l'on a $1 > \varepsilon_n \geq \sqrt{D_0\log n}/\sqrt{n}$ pour une certaine constante $D_0 = D_0(\mu) > 0$. Pour voir cela, il suffit de reprendre la preuve du théorème 3.2, c'est-à-dire celle des lemmes 3.2 et 3.1. Au lemme 3.2 appliqué à μ_{ε_n}, on remarque que le reste est un $o(1/n^{D_0/2})$ dès que $\varepsilon_n \geq \sqrt{D_0\log n}/\sqrt{n}$. De même, au lemme 3.1, le reste est uniforme quand ε est petit. On obtient donc (12) uniformément pour les μ_{ε_n}. Dans la suite, on pose $\mu_n = \mu_{\varepsilon_n}$ avec $\varepsilon_n = \sqrt{D_0\log n}/\sqrt{n}$.

Ensuite on vérifie que $K_n^{\mu}(x)$ et $K_n^{\mu_n}(x)$ sont comparables. Plus précisément, en considérant le calcul fait au lemme 3.2, on remarque que

$$h_{\mu_n}(t) = \int h_{\mu}(t - \varepsilon_n y)g(y)dy$$

où $h_{\mu}(t) = \sum_{p=1}^{r} P_p(t)g(t)/n^{p/2}$, $K_n^{\mu}(x) = \int_{|t|\leq x} h_{\mu}(t)dt$. En scindant cette intégrale en deux parties $|y| \leq x$ et $|y| > x$, on vérifie grâce à (13), que pour tout $c < r$ on a uniformément en x pour $1 \leq x \leq \sqrt{c\log n}$

$$\lim_{n\to+\infty} \frac{K_n^{\mu_n}(x)}{\mathbb{P}\left(|N| > x\right)} = 0$$

et

$$\lim_{n\to+\infty} \frac{\mathbb{P}\left(|S_n^{\mu_n}| > x\sqrt{\sigma_2(\mu_n)n}\right)}{\mathbb{P}\left(|N| > x\right)} = 1$$

Comme $\sigma_2(\mu_n) = 1 + \varepsilon_n^2$, et d'après le choix de ε_n on peut remplacer $\sigma_2(\mu_n)$ par 1 dans la limite précédente.

Finalement, on écrit pour $x \to +\infty$

$$\mathbb{P}\left(|S_n^{\mu_n}| > x\sqrt{n}\right) = \int_{|t|\leq 2x} \mathbb{P}\left(|S_n^{\mu}| > (x + \varepsilon_n t)\sqrt{n}\right) g(t)dt + o(\mathbb{P}(|N| > x))$$

D'où

$$\liminf \frac{\mathbb{P}\left(|S_n^{\mu}| > (1 - 2\varepsilon_n)x\sqrt{n}\right)}{\mathbb{P}(|N| > x)} \geq 1$$

et

$$\limsup \frac{\mathbb{P}\left(|S_n^{\mu}| > (1 + 2\varepsilon_n)x\sqrt{n}\right)}{\mathbb{P}(|N| > x)} \leq 1$$

Mais d'après le choix de ε_n, $\mathbb{P}(|N| > x(1 \pm 2\varepsilon_n))/\mathbb{P}(|N| > x)$ tend vers 1 uniformément quand $1 \leq x \leq \sqrt{c \log n}$. De plus (11) est valable pour uniformément pour x borné d'après le théorème de la limite centrale. Cela termine la preuve. $\square$

4.2. *Une version uniforme du théorème local*

Le théorème limite local classique (voir [Bre]) donne un équivalent quand n tend vers l'infini de la probabilité $\mathbb{P}(S_n \in I)$ où I est un intervalle borné de $\mathbb{R}$. Analogue local du théorème de Rubin et Sethuraman, le théorème suivant permet d'estimer $\mathbb{P}(S_n \in I + x)$ lorsque x varie de façon quelconque dans une boule de rayon au plus $\sqrt{cn \log n}$, où la constance c dépend du plus grand moment fini de μ. On obtient donc un théorème local pour les écarts modérés en supposant seulement un moment fini d'ordre assez grand. En particulier, dans le résultat suivant, on ne fait pas d'hypothèse sur la régularité de μ. Les cas où μ est supportée sur $\mathbb{Z}$ ou possède une densité bornée ont déjà été traités dans [Amo] et [Sla]. Notons ν la mesure gaussienne sur $\mathbb{R}$ associée à μ dans le théorème de la limite centrale

Théorème 4.2. *Si μ est une mesure centrée et apériodique sur $\mathbb{R}$ de variance σ_2 et admettant un moment d'ordre $r \geq 2$. Soit $s > 0$ et $I = I_s = [-s, s[$ un intervalle borné centré en* 0. *Alors pour tout $x \in \mathbb{R}$*

$$\lim_{n\to+\infty} \frac{\mu^n(I + x)}{\nu^n(I + x)} = 1. \tag{14}$$

De plus, si $r > 2$ (resp. $r = 2$) pour tout $c \in]0, r - 2[$, la limite (14) *est uniforme en x et s quand $|x| + s \leq \sqrt{c\sigma_2 n \log n}$ (resp. $|x| + s = O(\sqrt{n})$) et s est minoré par un réel > 0.*

Si de plus $\widehat{\mu}$ vérifie la condition de Cramér, alors il existe $\delta = \delta(\mu) > 0$ tel que la limite (14) *est uniforme quand $|x| + s \leq \sqrt{c\sigma_2 n \log n}$ (resp. $|x| + s = O(\sqrt{n})$ si $r = 2$) mais $s > e^{-\delta n}$.*

Cet énoncé sera la clé des théorèmes qui vont suivre dans l'étude de l'asymptotique de $\mathbb{E}(f(S_n))$ pour des fonctions qui ne sont pas à support compact ou bien pas intégrables.

Dans le cas $r = 2$ ce résultat est dû à Ch. Stone (voir [Sto]). Plus précisément :

Théorème 4.3. (Stone) *Si μ est une mesure centrée et apériodique sur $\mathbb{R}$ admettant un moment d'ordre 2, alors il existe une suite $(\varepsilon_n(\mu))_{n\geq 0}$ ne dépendant que de μ et tendant vers 0 telle que, pour tout intervalle fermé I de $\mathbb{R}$, on a*

$$\sup_{x\in\mathbb{R}} \left|\mu^n(I+x) - \nu^n(I+x)\right| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}}(1+|I|)$$

où ν est la mesure gaussienne sur $\mathbb{R}$ associée à μ dans le théorème de la limite centrale et $|I|$ la mesure de Lebesgue de I.

On remarque en passant que le théorème de la limite centrale pour une telle mesure μ est un corollaire immédiat. Pour des valeurs de x plus grandes que $\sqrt{n}$, Stone a aussi démontré un théorème local pour les grandes déviations sous l'hypothèse habituelle d'existence d'un moment exponentiel fini (voir [Sto2]). Des théorèmes semblables au théorème 4.3 ont été démontrés dans le cadre de la théorie du renouvellement (voir par exemple [Car1], [Nag], [Sta]) et notamment par Höglund qui obtient dans [Hög] un théorème de renouvellement pour les chaînes mixtes possédant à la fois une partie apériodique et une partie arithmétique, cadre qui généralise celui du théorème local. Keener dans [Kee] obtient un analogue du théorème de Stone pour le renouvellement sous la condition de Cramér et avec des conditions de moments légèrement différentes. Passons à la preuve du théorème 4.2 :

Preuve. La preuve qui suit reprend certaines idées de la preuve de (11) ainsi que de la preuve de Stone dans le cas $r = 2$. Notons $P_n^I(x) = \mu^n(x+I)$ et $Q_n^I(x) = \nu^n(x+I)$. On peut supposer que $\sigma_2 = 1$. On fixe une fonction intégrable paire, positive et continue $h \geq 0$ sur $\mathbb{R}$ dont la transformée de Fourier $\widehat{h}$ est de classe C^∞ et à support compact inclus dans $[-1,1]$ et vérifie de plus $\int h = 1$ (voir le paragraphe 3.2). On introduit aussi les $h_\varepsilon(x) = \frac{1}{\varepsilon}h(\frac{x}{\varepsilon})$ qui forment une famille de Dirac quand ε tend vers 0. La preuve s'effectue en trois étapes : lissage, preuve pour la quantité lissée, et "délissage". Pour plus de clarté, on écrit ces étapes sous la forme de lemmes.

Lemme 4.1. *Il existe une suite $\varepsilon_n(\mu) \to 0$ ne dépendant que de μ telle que pour tout $\varepsilon > 0$ fixé on ait, uniformément en x et s :*

$$\left|P_n^I * h_\varepsilon(x) - Q_n^I * h_\varepsilon(x)\right| \leq \varepsilon_n(\mu) Q_n^I * h_\varepsilon(x) + |I| \cdot o_\varepsilon\left(\frac{1}{n^{(r-1)/2}}\right).$$

De plus si μ satisfait la condition de Cramér, alors on peut choisir ε de la forme e^{-an} pour un certain $a > 0$ petit et on obtiendra toujours un $o(1/n^{(r-1)/2})$ dans le reste ci-dessus.

Etant donnés $\varepsilon > 0$ et $x \in \mathbb{R}$, on pose $f_x(u) = \chi_{I+x} * h_\varepsilon(u)$. Alors $P_n^I * h_\varepsilon = \int f_x d\mu^n$. On peut appliquer à f_x le lemme 3.2 et on obtient pour $D > r/\sigma_2$

$$\int_{|t|\le\sqrt{\frac{D\log n}{n}}} \widehat{f_x}(t)\widehat{\mu}^n(t)dt = 2\pi Q_n^I * h_\varepsilon + \sum_{p=1}^{r-2} \frac{1}{n^{p/2}} \int f_x(u)P_p\Big(\frac{u}{\sqrt{n}}\Big)d\nu^n(u) \\ +|I|\cdot o_{r,D,\mu}\Big(\frac{1}{n^{(r-1)/2}}\Big) \tag{15}$$

Comme les polynômes P_p ne dépendent que de μ et sont de degré au plus $3p$, il existe une constante $C_1 = C_1(\mu)$ telle que

$$\int_{|u|\le\sqrt{Dn\log n}} f_x(u)P_p\Big(\frac{u}{\sqrt{n}}\Big)d\nu^n(u) \le C_1(\sqrt{D\log n})^{3p} Q_n^I * h_\varepsilon(x)$$

d'autre part, comme $P_p(u)g(u) \le e^{-u^2/4}$ dès que u est assez grand (g est la densité de la loi normale),

$$\int_{|u|\ge\sqrt{Dn\log n}} f_x(u)P_p\Big(\frac{u}{\sqrt{n}}\Big)d\nu^n(u) \le \frac{1}{n^{D/4}}\,\|f_x\|_1 \le \frac{1}{n^{D/4}}|I|$$

D'où l'on conclut que pour $D \ge 2r$ et uniformément en $x \in \mathbb{R}$, (15) devient

$$\frac{1}{2\pi}\int_{|t|\le\sqrt{\frac{D\log n}{n}}} \widehat{f_x}(t)\widehat{\mu}^n(t)dt = (1+\varepsilon_n(x))Q_n^I * h_\varepsilon(x) + |I|\cdot o_{r,D,\mu}\Big(\frac{1}{n^{(r-1)/2}}\Big)$$

où la suite $\varepsilon_n(x)$ ne dépend que de μ et tend vers 0 uniformément en x.

Passons maintenant aux grandes valeurs de t. Notons que $\widehat{f_x}(t) = \widehat{\chi_{I+x}}(t)\widehat{h}(\varepsilon t)$:

$$\left|\int_{|t|\ge\sqrt{\frac{D\log n}{n}}} \widehat{f_x}(t)\widehat{\mu}^n(t)dt\right| = \left|\int_{\sqrt{\frac{D\log n}{n}}\le|t|\le 1/\varepsilon} \widehat{f_x}(t)\widehat{\mu}^n(t)dt\right| \\ \le |I|\int_{\sqrt{\frac{D\log n}{n}}\le|t|\le 1/\varepsilon} |\widehat{\mu}(t)|^n\,dt = o_\varepsilon\Big(\frac{1}{n^r}\Big)$$

dès que $D > r/c$ où c est la constante définie dans la preuve du lemme 3.1. On remarque que si μ satisfait la condition de Cramér, alors on peut choisir ε de la forme e^{-an} pour un certain $a > 0$ petit et on obtiendra toujours un $o(1/n^r)$ dans la ligne de calcul ci-dessus. En prenant D assez grand et en combinant les majorations ci-dessus on obtient l'estimation du lemme 4.1. □

Lemme 4.2. *Pour tout $\delta > 0$, il existe $\varepsilon_0 > 0$ tel que si x et s sont des $O(\sqrt{n\log n})$, alors pour tout n assez grand et tout $\varepsilon \in]0,\varepsilon_0[$ on a uniformément en x et $I = I_s = [-s, s[$*

$$\left|Q_n^I * h_\varepsilon(x) - Q_n^I(x)\right| \le \delta\cdot Q_n^I(x)$$

Soit $c > 0$ une constante telle que $|x| + s \leq \sqrt{cn \log n}$. Notons que pour tous réels u et t

$$|g(u + \varepsilon t) - g(u)| = g(u)|e^{\varepsilon ut} e^{-\varepsilon^2 t^2/2} - 1| \leq \frac{\delta}{2} g(u)$$

dès que $|ut| \leq 1$, $|t| \leq 1$ et pour tout ε assez petit (disons $\leq \varepsilon_0$). Il vient

$$\left| Q_n^I(x - \varepsilon y) - Q_n^I(x) \right| \leq \frac{\delta}{2} Q_n^I(x) \tag{16}$$

tant que $\varepsilon < \varepsilon_0$ et $|y| \leq \sqrt{n}/\sqrt{c \log n}$. D'où

$$\left| Q_n^I * h_\varepsilon(x) - Q_n^I(x) \right| \leq \frac{\delta}{2} \cdot Q_n^I(x) + 2\frac{|I|}{\sqrt{n}} \int_{|y| \geq \sqrt{n}/\sqrt{c \log n}} h(y) dy.$$

Mais $\widehat{h}$ est de classe C^∞ donc $h(y) = o(\frac{1}{|y|^k})$ au voisinage de l'infini pour tout $k \geq 0$. Donc

$$\int_{|y| \geq \sqrt{n}/\sqrt{c \log n}} h(y) dy = o\Big(\frac{1}{n^k}\Big). \tag{17}$$

Mais

$$\sqrt{2\pi}\, Q_n^I(x) \geq \frac{|I|}{\sqrt{n}} \frac{1}{n^{c/2}} \tag{18}$$

donc pour tout n assez grand on a le résultat souhaité. □

Nous allons maintenant terminer la preuve du théorème. On suppose désormais que $s = |I|/2 > \varepsilon^{1/3}$, et l'on note $I_\varepsilon = [-s - \sqrt{\varepsilon}, s + \sqrt{\varepsilon}[$ et $I^\varepsilon = [-s + \sqrt{\varepsilon}, s - \sqrt{\varepsilon}[$. Remarquons d'abord qu'il existe $\theta(\varepsilon)$ satisfaisant $\theta(\varepsilon) \to 0$ si $\varepsilon \to 0$ telle que, dans les conditions du lemme 4.2,

$$\left| Q_n^{I_\varepsilon}(x) - Q_n^{I^\varepsilon}(x) \right| \leq \theta(\varepsilon) \cdot Q_n^I(x). \tag{19}$$

Ensuite on a $P_n^{I_\varepsilon} * h_\varepsilon(x) \geq P_n^I(x) \int_{|y| \leq 1/\sqrt{\varepsilon}} h(y) dy$ puis d'après les lemmes 4.1 et 4.2 pour ε assez petit

$$\frac{P_n^I(x)}{1 - \varepsilon} \leq P_n^{I_\varepsilon} * h_\varepsilon(x) \leq (1 + \varepsilon_n(\mu))(1 + \delta) Q_n^{I_\varepsilon}(x) + |I_\varepsilon| \cdot o_\varepsilon\Big(\frac{1}{n^{(r-1)/2}}\Big).$$

Puis en tenant compte de (18) (on a supposé $c \leq r - 2$) et de (19) on obtient que pour tout n assez grand

$$P_n^I(x) \leq (1 + 2\delta) Q_n^I(x). \tag{20}$$

On remarque aussi que si la condition de Cramér est satisfaite, on peut prendre $\varepsilon = e^{-3an}$ pour un certain $a > 0$ donc, pour tout n grand, (20) est vraie uniformément quand $s + |x| \leq \sqrt{cn \log n}$ et $s > e^{-an}$.

De façon similaire, on a

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \leq P_n^I(x) \int_{|y| < 1/\sqrt{\varepsilon}} h(y) dy + \int_{|y| \geq 1/\sqrt{\varepsilon}} P_n^{I^\varepsilon}(x - \varepsilon y) h(y) dy$$

et d'après (20), (19) et (16), en posant $A_n = [1/\sqrt{\varepsilon}, \sqrt{n}/\sqrt{c \log n}]$ dès que ε est assez petit,

$$\begin{aligned}
\int_{|y| \in A_n} P_n^{I^\varepsilon}(x - \varepsilon y) h(y) dy &\leq (1 + 2\delta) \int_{|y| \in A_n} Q_n^{I^\varepsilon}(x - \varepsilon y) h(y) dy \\
&\leq (1 + 2\delta)(1 + \theta(\varepsilon)) \int_{|y| \in A_n} Q_n^I(x - \varepsilon y) h(y) dy \\
&\leq Q_n^I(x)(1 + 2\delta)(1 + \theta(\varepsilon))(1 + \delta) \int_{|y| \geq 1/\sqrt{\varepsilon}} h(y) dy \\
&\leq \delta Q_n^I(x)
\end{aligned}$$

et par (17) pour tout $k \geq 0$ fixé, pour n assez grand

$$\int_{|y| \geq \sqrt{n}/\sqrt{c \log n}} P_n^{I^\varepsilon}(x - \varepsilon y) h(y) dy \leq \frac{1}{n^k}.$$

Finalement par (18) pour ε assez petit et si $s + |x| \leq \sqrt{cn \log n}$, $s > \sqrt{\varepsilon}$, on a pour n grand

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \leq P_n^I(x) + \delta Q_n^I(x). \tag{21}$$

Mais d'après le lemme 4.1 et (18)

$$P_n^{I^\varepsilon} * h_\varepsilon(x) = (1 + o_\varepsilon(1)) Q_n^{I^\varepsilon} * h_\varepsilon(x).$$

puis en appliquant le lemme 4.2 puis (19)

$$P_n^{I^\varepsilon} * h_\varepsilon(x) \geq (1 - |o_\varepsilon(1)|)(1 - \delta) Q_n^{I^\varepsilon}(x) \geq (1 - 2\delta) Q_n^I(x)$$

pour n assez grand. En combinant ceci avec (21) on obtient

$$P_n^I(x) \geq (1 - 3\delta) Q_n^I(x). \tag{22}$$

On a donc bien la majoration (20) et la minoration (22) souhaitées. La preuve du théorème est complète. □

Remarque 4.1. *Lorsque μ est l-diophantienne, il est possible de préciser le comportement asymptotique de la suite $(\varepsilon_n(\mu))_n$ intervenant dans (4.3). Le calcul donne $\varepsilon_n(\mu) = o(\frac{1}{n^{\delta/2}})$ pour tout $0 < \delta < \frac{1}{l}$ si μ possède un moment d'ordre 3. De même, si μ est l -diophantienne, l'équivalent est valable uniformément quand $s > \frac{1}{n^{\delta/2}}$ pour tout $0 < \delta < \frac{1}{l}$.*

Remarque 4.2. *On donne plus bas un exemple (cf. proposition 5.1) qui montre que la convergence vers 0 de $\varepsilon_n(\mu)$ peut être aussi lente que l'on veut si l'on ne fait pas d'hypothèse sur μ. Plus exactement, pour toute suite de réels strictement positifs $\varepsilon_n > 0$ tendant vers 0, on peut trouver une mesure μ (non diophantienne en général) telle que* $\limsup \varepsilon_n(\mu)/\varepsilon_n > 0$.

5. Un théorème limite pour les fonctions bornées

Nous allons maintenant passer au second volet de cet article et nous intéresser au comportement asymptotique de $\mathbb{E}(f(S_n))$ lorsque la fonction f est seulement supposée bornée et assez régulière. Commençons par remarquer que pour toute fonction bornée f ayant une limite l en $\pm\infty$, les moyennes $\mathbb{E}(f(S_n))$ convergent vers l. Cela résulte en effet immédiatement du théorème de la limite centrale. Dans les paragraphes suivants, nous allons montrer que, en supposant μ ou f assez régulière mais sans aucune hypothèse sur le comportement à l'infini de f, ces moyennes sont asymptotiquement indépendantes de la marche aléatoire centrée de variance 1 que l'on considère (théorème 5.3).

Nous commençons dans le premier paragraphe par étendre le théorème local aux fonctions directement Riemann intégrables.

5.1. Fonctions directement Riemann intégrables

Dans ce paragraphe, on s'intéresse à la validité du théorème limite local pour des fonctions qui ne sont plus nécessairement à support compact mais toujours intégrables.

Théorème 5.1. *Soit μ une mesure de probabilité centrée et apériodique sur $\mathbb{R}$ avec moment d'ordre 2 fini. Soit f une fonction Riemann intégrable sur $\mathbb{R}$ telle que*

$$\sum_{n\in\mathbb{Z}} \max_{[x]=n} |f(x)| < +\infty \tag{23}$$

Alors on a le théorème limite local pour f, c'est-à-dire que

$$\lim_{n\to+\infty} \sqrt{2\pi\sigma_2 n} \int f d\mu^n = \int f(x)dx \tag{24}$$

Preuve. On reprend la preuve classique du théorème local (voir [Bre]). Notons $a_n = \max_{[x]=n} |f(x)|$ pour tout entier $n \in \mathbb{Z}$, et considérons la fonction

$$\beta(x) = \frac{1}{2^2} \sin^2(\pi x) \sum_{n\in\mathbb{Z}} \frac{a_n}{(x-n)^2}$$

On peut supposer que la suite (a_n) n'est pas identiquement nulle. D'après le paragraphe (3.2) β est continue, strictement positive sur $\mathbb{R}$, et appartient à $\mathbb{L}^1(\mathbb{R})$. De plus $\beta(x) \geq a_{[x]} \geq |f(x)|$. Enfin $\widehat{\beta}$ est à support compact. Considérons la suite de mesures positives $d\nu_n(x) = \sqrt{2\pi\sigma_2 n}\beta(x)d\mu^n(x)$. Puisque $\widehat{\beta}$ est à support compact, la suite des masses totales $\nu_n(\mathbb{R})$ est bornée. De plus, pour tout réel t_0,

$$\begin{aligned} \int e^{it_0 x} d\nu_n(x) &= \sqrt{2\pi\sigma_2} \int \widehat{\beta}\Big(\frac{t}{\sqrt{n}} - t_0\Big)\widehat{\mu}^n\Big(\frac{t}{\sqrt{n}}\Big)dt \\ &\longrightarrow_{n\to+\infty} \widehat{\beta}(-t_0) = \int e^{it_0 x}\beta(x)dx \end{aligned}$$

On en déduit que la suite de mesures $(\nu_n)_n$ converge (convergence usuelle des mesures) vers la mesure βdx.

La fonction f/β est bien définie, Riemann intégrable et bornée par 1, donc

$$\lim_{n\to+\infty}\sqrt{2\pi\sigma_2 n}\int f d\mu^n = \lim_{n\to+\infty}\int \frac{f}{\beta} d\nu_n(x) = \int f(x)dx$$

ce qu'il fallait démontrer. □

Remarque 5.1. *On qualifie de directement Riemann intégrables les fonctions f Riemann intégrables satisfaisant la condition (23). La notion est introduite par Feller dans [Fel] et étudiée en détail dans [Hin] (la définition de cette notion peut prendre plusieurs formes équivalentes). Le théorème de renouvellement de Höglund [Hög] est énoncé pour toute fonction directement Riemann intégrable. Le théorème 5.1 est donc un cas particulier du résultat de Höglund, bien que la méthode de preuve soit ici différente (comparer prop. 2.2.).*

Remarque 5.2. *En général, il y a des exemples de fonctions f intégrables sur $\mathbb{R}$ et de classe C^p, $p \geq 1$, dont toutes les dérivées jusqu'à l'ordre de p tendent vers* 0 *à l'infini, qui ne vérifient pas* (24). *Plus bas, on donne un exemple d'une telle fonction pour une distribution μ qui est même diophantienne. Cependant, on verra au paragraphe 5.3 que si μ satisfait la condition de Cramér, alors* (24) *a bien lieu dès que f est intégrable et hölderienne sur $\mathbb{R}$.*

5.2. Fonctions intégrables à variation bornée

Le résultat précédent montre en particulier que le théorème local est valide pour les fonctions continues intégrables et décroissantes au loin quand $|x| \to +\infty$. En fait, par une méthode différente, on a aussi le résultat suivant, valide pour les fonctions à variation bornée.

Théorème 5.2. *Soit à nouveau μ une mesure de probabilité centrée et apériodique sur $\mathbb{R}$ ayant moment d'ordre* 2 *fini. Soit $(\varepsilon_n(\mu))_n$ la suite qui tend vers* 0 *intervenant dans l'énoncé du théorème de Stone (théorème 4.3). Alors pour toute fonction f intégrable et à variation bornée sur $\mathbb{R}$,*

$$\left|\int f d\mu^n - \int f d\nu_n\right| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}}\left(\int |f| + Var(f)\right)$$

où $Var(f)$ est la variation totale de f, et $d\nu_n(x) = \frac{\exp(-x^2/2\sigma_2 n)}{\sqrt{2\pi\sigma_2 n}}dx$ est la suite de mesures gaussiennes associée à μ.

Comme on l'a remarqué (cf. remarque 4.1), le comportement de la suite $(\varepsilon_n(\mu))_n$ peut être précisé lorsque l'on suppose que μ est diophantienne. Passons à la preuve du théorème.

Preuve. Clairement, on peut supposer que f est à support compact (considérer $f(x)\chi_{|x|<M}$ et faire tendre M vers l'infini à n fixé et remarquant que $Var(f\chi_{|x|<M}) \to Var(f)$). De même, en approximant f uniformément par des fonctions en

escalier de la forme $\sum a_i \chi_{A_i}$ où $A_i = [x_i, x_{i+1})$ pour une certaine subdivision $x_0 < \cdots < x_N$, on se ramène au cas où f a précisément cette forme. Finalement, puisque $Var(f) = Var(f^+) + Var(f^-)$ et $|f| = f^+ + f^-$, où f^+ et f^- sont respectivement les parties positives et négatives de f, on peut supposer que f ne prend que des valeurs positives, i.e. $a_i \geq 0$ pour tout i.

Soit $0 = f_0 < f_1 < \cdots < f_K$ la suite ordonnée des valeurs prises par la fonction f. Maintenant considérons la famille $(B_j)_j$ de toutes les composantes connexes des ensembles de niveau $\{x | f(x) \geq f_i\}$ pour i variant de 1 à K. Alors

$$f(x) = \sum (f_i - f_{i-1}) \chi_{f(x) \geq f_i} = \sum \beta_j \chi_{B_j}(x)$$

où l'on définit β_j par $f_i - f_{i-1} \geq 0$ si B_j est une composante de l'ensemble $\{x | f(x) \geq f_i\}$. Remarquons que chaque B_j est une union finie d'intervalles A_i adjacents.

D'après le théorème de Stone (4.3), on obtient

$$\left| \int f d\mu^n - \int f d\nu_n \right| \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}} \sum \beta_j (\left| B_j \right| + 1) \leq \frac{\varepsilon_n(\mu)}{\sqrt{n}} \left(\int f + \sum \beta_j \right)$$

Mais la construction des intervalles B_j est de telle sorte que

$$\sum \beta_j = \frac{1}{2} Var(f)$$

En fait, parmi toutes les representations de f en sommes de la forme $\sum \beta_j \chi_{B_j}$ avec des $\beta_j \geq 0$ et des intervalles B_j, celle definie ci-dessus minimise $\sum \beta_j$. $\square$

5.3. Fonctions bornées

Ici nous allons obtenir l'équivalent annoncé dans l'introduction pour des fonctions bornées, sous l'hypothèse que μ est diophantienne (resp. Cramér). Plus précisément, nous avons :

Théorème 5.3. *Soit μ une mesure de probabilité centrée sur $\mathbb{R}$ et l un réel ≥ 0. On suppose de plus que μ possède moment fini d'ordre 4 et on note ν la mesure gaussienne associée. Enfin on suppose que μ est l-diophantienne. Soit k_0 un entier tel que $k_0 > 3l/2 + 1$. Alors toute fonction f définie sur $\mathbb{R}$ et telle que f et toutes ses dérivées jusqu'à l'ordre k_0 sont bornées vérifie la relation suivante*

$$\int f d\mu^n - \int f d\nu^n = o(\int |f| d\nu^n) \tag{25}$$

En particulier, si $f \geq 0$ non identiquement nulle

$$\frac{\int f d\mu^n}{\int f d\nu^n} \to 1 \tag{26}$$

Si l'on suppose de plus que μ vérifie la condition de Cramér, alors (25) et (26) sont vérifiées pour toute fonction f höldérienne bornée sur $\mathbb{R}$.

Preuve. Comme on le vérifie aisément, pour chaque n, on peut trouver deux fonctions f_n et ϕ_n de classe C^{k_0} telles que $f = f_n + \phi_n$ avec f_n à support dans $[-s_n - 1, s_n + 1]$ et $\phi_n(x) = 0$ si $|x| \leq s_n$ où $s_n = 2\sqrt{n \log n}$. On peut de plus choisir f_n telle que $C^{k_0}(f_n) \leq C\sqrt{n \log n}$ pour une certaine constante C. Appliquons alors le théorème 3.2 aux fonctions f_n. Il vient aussitôt

$$\left|\int f_n d\mu^n - \int f_n d\nu^n\right| \leq \varepsilon_n \int |f_n| d\nu^n + C\frac{\sqrt{\log n}}{n}$$

où ε_n est un $O(\frac{\log^3 n}{n})$ indépendant de f. De plus, d'après la loi des écarts modérés (théorème 11) $\mu^n(|x| \geq s_n)$ et $\nu^n(|x| \geq s_n)$ sont des $o(1/n^{3/2})$. Enfin, d'après le théorème local lui-même, il existe $c > 0$ tel que $\int |f| d\nu^n \geq c/\sqrt{n}$ si f n'est pas identiquement nulle. Le théorème s'ensuit aussitôt.

Si l'on suppose que μ satisfait la condition de Cramér, alors on écrit $f = \sum f_i + R_n$ où f_i est la restriction de f à l'intervalle I_i et R_n la restriction au complémentaire de $[-s_n, s_n]$. Ces intervalles I_i sont de longueur $e^{-\delta n}$ et subdivisent $[-s_n, s_n]$ (où $\delta > 0$ est la constante intervenant dans le théorème 4.2). Choisissons dans chaque I_i un point x_i. Alors

$$\int_{I_i} |f - f(x_i)|\, d\mu^n \leq Lip_\alpha(f)|I_i|^\alpha \mu^n(I_i)$$

et de même pour ν^n. D'après le théorème 4.2, il existe une suite $\varepsilon_n \to 0$ telle que pour tout i

$$\left|\mu^n(I_i) - \nu^n(I_i)\right| \leq \varepsilon_n \nu^n(I_i)$$

Enfin

$$\int |R_n| d\mu^n \leq \|f\|_\infty \, \mu^n(x, |x| \geq s_n)$$

Ainsi d'après la loi des écarts modérés 11, pour tout n assez grand,

$$\begin{aligned}\left|\int f d\mu^n - \int f d\nu^n\right| &\leq \varepsilon_n \sum_i f(x_i)\nu^n(I_i) + 2Lip_\alpha(f)e^{-\delta\alpha n} + \|f\|_\infty\, o(1/n^{3/2}) \\ &\leq 2\varepsilon_n \int |f| d\nu^n\end{aligned}$$

□

Remarquons que le théorème reste valable dans le cas où μ satisfait la condition de Cramér pour les fonctions f qui sont höldériennes par morceaux sur $\mathbb{R}$, si ces morceaux sont de longueur bornée inférieurement par rapport à 0 et si les constantes de Hölder sur les morceaux sont uniformément bornées.

Si l'on suppose seulement μ diophantienne et si la fonction f possède un nombre p insuffisant de dérivées successives bornées, alors l'équivalent (26) n'est pas valable (cf. 6.1). Remarquons aussi que supposer $\exists q > 0, 1 - |\widehat{\mu}(t)| > 1/(\log|t|)^q$ pour t grand suffit à la place de la condition de Cramér pour avoir (26) pour les fonctions höldériennes bornées.

5.4. Un exemple

Ici on va construire un exemple d'une mesure centrée diophantienne et à support fini, pour laquelle on peut trouver des fonctions f continues et intégrables sur $\mathbb{R}$ telles que

$$\lim_{n\to+\infty} \sqrt{2\pi\sigma_2 n} \int f d\mu^n = +\infty$$

et qui par suite ne satisfont pas le théorème local. On verra qu'on peut même choisir f de classe C^p et telle que f et ses dérivées jusqu'à l'ordre p soient bornées sur $\mathbb{R}$. La construction de μ nous sera aussi utile en 6.1.

Considérons

$$\mu_\alpha = \frac{1}{2}\delta_0 + \frac{1-\alpha}{2}\delta_\alpha + \frac{\alpha}{2}\delta_{\alpha-1}$$

où $\alpha \in (0, 1)$ et $\alpha \notin \mathbb{Q}$. Alors $\mu = \mu_\alpha$ est apériodique. Ci-dessous, on construit α à partir de $(\varepsilon_n)_{n\geq 0}$ et $(h_n)_{n\geq 0}$, deux suites de nombres > 0 décroissantes vers 0.

On adopte les notations suivantes : $S_N = X_1 + \cdots + X_N = p\alpha - q(1-\alpha) = n\alpha - q$ où $p, q \geq 0$ et $n = p + q \leq N$. C'est-à-dire $n = \#\{i \leq N,\ X_i \neq 0\}$ et $q = \#\{i \leq N,\ X_i = \alpha - 1\}$. On pose :

$$\alpha = \sum_{i\geq 1} 10^{-M_i} \tag{27}$$

et on définit la suite la suite $(M_i)_i$ et la suite auxilliaire $(N_i)_i$ récursivement de la façon suivante :

- $N_0 = 1$, $M_0 = 0$ et $M_1 = 1$
- Pour $i \geq 1$, on choisit N_i assez grand pour que $\varepsilon_{N_i} \leq 10^{-M_i}$ et $N_i > 10^{2M_i}$ et $N_{i+1} > 5N_i$.

Alors on choisit M_{i+1} de sorte que $10^{M_i - M_{i+1}} \leq \frac{h_{N_i}}{N_i}$ et $M_{i+1} - M_i > M_i - M_{i-1}$.

La dernière condition implique que α n'est pas rationnel, et donc que μ_α est apériodique. De plus,

$$\{10^{M_i}\alpha\} \leq 2 \cdot 10^{M_i - M_{i+1}} \leq 2\frac{h_{N_i}}{N_i} \tag{28}$$

où on rappelle que $\{x\} = d(x, \mathbb{Z})$ pour tout $x \in \mathbb{R}$, et $[x]$ est l'entier le plus proche de x. Donc pour tout n entre 0 et N_i on a $\{n10^{M_i}\alpha\} \leq 2h_{N_i}$. Et il vient

$$d(S_{N_i}, 10^{-M_i}\mathbb{Z}) \leq 2 \cdot 10^{-M_i} h_{N_i} \tag{29}$$

Avec ces notations, pour N assez grand (dès que $h_N < 1$), on a aussi

$$|S_N| \leq h_N \Leftrightarrow |n\alpha - q| \leq h_N \Leftrightarrow \{n\alpha\} \leq h_N \text{ et } q = [n\alpha] \tag{30}$$

Remarque 5.3. *Remarquons que μ_α est diophantienne si et seulement si le réel α est diophantien. De plus, on constate d'après la définition de α en (27) que α est diophantien si et seulement si la suite $(M_i)_i$ croît au plus exponentiellement, i.e. $\exists \rho > 1$ tel que $M_i \leq \rho^i$. Enfin, on vérifie que si les suites (ε_n) et (h_n) sont de la forme $\varepsilon_n = 1/n^a$ et $\varepsilon_n = 1/n^b$ avec $a > 0$ et $b > 0$, alors on peut choisir la suite M_i comme ci-dessus et vérifiant $M_i \leq \rho^i$ pour un certain ρ assez grand, donc μ_α est alors diophantienne.*

La construction précédente permet en particulier de montrer le résultat suivant :

Proposition 5.1. *Pour tout choix de suites $(\varepsilon_n)_n$ et $(h_n)_n$ de nombres strictement positifs décroissantes vers 0, il existe un $\alpha \in \mathbb{R}$ et une mesure μ_α définie comme plus haut, telle que*

$$\limsup_{n \to +\infty} \frac{\sqrt{n}}{\varepsilon_n} \mu_\alpha^n([-h_n, h_n]) > 0$$

Ainsi la suite $(\varepsilon_n(\mu))_n$ intervenant dans le théorème local de Stone (4.3) peut décroître vers 0 aussi lentement que l'on veut.

Preuve. On note E_n l'événement $\{X_i = 0 \ N - n \text{ fois}\} \cap \{X_i = \alpha - 1 \ [n\alpha] \text{ fois}\} \cap \{\|n\alpha\| \leq h_N\}$. Alors pour tout N assez grand

$$\begin{aligned}
\mathbb{P}(|S_N| \leq h_N) &\geq \sum_{n=0}^{N} \mathbb{P}(E_n) \\
&\geq \sum_{N/2-\sqrt{N} \leq n \leq N/2+\sqrt{N}} \frac{1}{2^N} C_N^n \alpha^{[n\alpha]} (1-\alpha)^{n-[n\alpha]} C_n^{[n\alpha]} \chi_{\{n\alpha\} \leq h_N}
\end{aligned}$$

Ainsi

$$\mathbb{P}(|S_N| \leq h_N) \geq \frac{c^2}{\sqrt{N}} \sum_{N/2-\sqrt{N} \leq n \leq N/2+\sqrt{N}} \frac{1}{\sqrt{n}} \chi_{\{n\alpha\} \leq h_N}$$

Dans la dernière inégalité, $c > 0$ est une constante, telle que $\frac{1}{2^N} C_N^n > c/\sqrt{N}$ pour tout N et $n \in [N/2 - \sqrt{N}, N/2 + \sqrt{N}]$ et $\alpha^{[n\alpha]}(1-\alpha)^{n-[n\alpha]} C_n^{[n\alpha]} > c/\sqrt{n}$, pour tout n. Elle est déterminée par le théorème limite local pour les distributions non apériodiques (voir [GnK] chap. 9 §49) ou bien directement à partir de la formule de Stirling. Ainsi pour tout N assez grand

$$\mathbb{P}(|S_N| \leq h_N) \geq \frac{c^2}{N} A(N)$$

où

$$A(N) = \# \left\{ n \in [N/2 - \sqrt{N}, N/2 + \sqrt{N}] \ s.t. \ \{n\alpha\} \leq h_N \right\}$$

Avec le réel α défini plus haut à partir des suites $(\varepsilon_n)_n$ et $(h_n)_n$, on a vu en (28) que $\{10^{M_i}\alpha\} \leq 2\frac{h_{N_i}}{N_i}$. Mais, puisque $N_i 10^{M_i}/2 \geq N_i$ et $10^{M_i} < \sqrt{N_i}$, on obtient $A(N_i) \geq \frac{2\sqrt{N_i}}{10^{M_i}}$ et $\frac{A(N_i)}{\sqrt{N_i}} \geq \varepsilon_{N_i}$. Comme cette inégalité en vraie pour tout $i \geq 1$, on a bien la conclusion attendue. □

Passons maintenant à la construction de la fonction f annoncée au début de ce paragraphe. Soit p un entier ≥ 1 donné. Fixons $\varepsilon > 0$ tel que $\varepsilon < 1/(10+6p)$ et notons $a = \frac{1}{2} - 2\varepsilon$ et $b = \frac{a}{p+1}$. Ainsi $b > \varepsilon$. Fixons la suite $(\varepsilon_n)_n$ en posant $\varepsilon_n = n^{-\varepsilon}$ et la suite $(h_n)_n$ par $h_n = 1/n$. D'après la remarque 5.3, la mesure μ_α est diophantienne. Soit $a_i = N_i^{-a}$ et $b_i = N_i^{-b}$. Définissons maintenant la fonction f_i en lui donnant la valeur a_i aux points d'abscisse $k10^{-M_i}$ pour tous les entiers k tels que $k10^{-M_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]$ et posons autour de ces points (i.e. pour $|x| < 10^{-M_i}$) $f_i(k10^{-M_i} + x) = a_i\phi(x/b_i)$ où ϕ est une "fonction plateau" de classe C^∞ à support dans $[-1, 1]$, à valeurs dans $[0, 1]$ et égale à 1 dans un voisinage de 0. On suppose de plus que f_i est nulle partout ailleurs. La définition est cohérente car $b_i = N_i^{-b} \leq N_i^{-\varepsilon} = \varepsilon_{N_i} \leq 10^{-M_i}$. Enfin, on pose $f = \sum_i f_i$. La fonction f est C^∞, à valeurs positives ou nulles, et f ainsi que ses dérivées jusqu'à l'ordre p tendent vers 0 car $a_i/b_i^{p+1} = 1$. Remarquons aussi que f est intégrable car

$$\sum a_i b_i \frac{\sqrt{N_i}}{10^{-M_i}} \leq \sum a_i b_i N_i^{\frac{1}{2}+\varepsilon} \leq \sum N_i^{-(a+b-\varepsilon-\frac{1}{2})} < \infty$$

et d'après le choix de ε et a et b ci-dessus, $a + b > \varepsilon + \frac{1}{2}$ et de plus $N_i \geq 5^i$.

D'autre part, la relation (29) montre que $f(S_{N_i}) = a_i$ dès que $S_{N_i} \in [\frac{1}{2}\sqrt{N_i}, \sqrt{N_i}]$. Donc $\mathbb{E}(f(S_{N_i})) \geq ca_i$ où $c > 0$ ne dépend que de μ_α. Mais $a_i\sqrt{N_i} = N_i^{2\varepsilon} \to +\infty$, ce qui fournit le contre-exemple annoncé en début de paragraphe.

6. Théorème limite pour les fonctions asymptotiquement constantes en moyenne

Pour terminer, nous présentons dans cette section un théorème limite valable sans hypothèse sur μ (autre que celles du théorème local) pour les fonctions dont les moyennes de Cesaro sur de grands intervalles sont asymptotiquement constantes. On en déduit ensuite une application à l'équidistribution des marches aléatoires. A nouveau, la clé du théorème qui suit est le théorème local de Stone :

Théorème 6.1. *Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, apériodique et ayant un moment d'ordre 2 fini. On suppose que f est une fonction unifomément continue et bornée sur $\mathbb{R}$ telle que la limite suivante existe*

$$\lim \frac{1}{T} \int_0^T f(t)dt = l$$

quand $|T| \to +\infty$. Alors

$$\lim_{n\to+\infty} \int f d\mu^n = l$$

La preuve résulte des deux observations suivantes, que l'on présente sous la forme de deux lemmes.

Lemme 6.1. *En reprenant les notations du théorème, pour toute fonction f uniformément continue sur $\mathbb{R}$, on a*

$$\lim_{n\to+\infty}\left|\int f d\mu^n - \int f d\nu^n\right| = 0$$

où ν est la mesure gaussienne $d\nu(x) = \frac{\exp(-x^2/2\sigma_2)}{\sqrt{2\pi\sigma_2}}dx$ associée à μ.

Preuve. Soit $\varepsilon > 0$ et $\omega > 0$ un module de continuité pour f relativement à ε, i.e. $|f(x+u) - f(x)| \leq \varepsilon$ si $|u| \leq \omega$. Fixons $C > \log\frac{1}{\varepsilon}$, et $A_n = \{x, |x| \leq C\sqrt{n}\}$. En appliquant le théorème central limite à la somme de variables aléatoires S_n, on obtient pour tout n assez grand : $\mu^{*n}(A_n^c) \leq \varepsilon$ et de façon similaire $\nu^{*n}(A_n^c) \leq \varepsilon$. On peut subdiviser A_n en $O(\sqrt{n}/\omega)$ intervalles $h_i + I$ de longueur $|I| = \omega$. On obtient

$$\begin{aligned}\left|\int f d\mu^{*n} - \int f d\nu^{*n}\right| &\leq \left|\int_{A_n} f d\mu^{*n} - \int_{A_n} f d\nu^{*n}\right| + 2\varepsilon \\ &\leq \sum_i f(h_i)\left|\mu^{*n}(h_i + I) - \nu^{*n}(h_i + I)\right| + 4\varepsilon \\ &\leq \|f\|_\infty\, O\Big(\frac{\sqrt{n}}{\omega}\Big) \sup_{h\in\mathbb{R}}\left|\mu^{*n}(h+I) - \nu^{*n}(h+I)\right| + 4\varepsilon\end{aligned}$$

On peut maintenant appliquer la version uniforme du théorème limite local de Stone (théorème 4.3), qui affirme que

$$\lim_{n\to\infty}\sqrt{n}\sup_{h\in\mathbb{R}}\left|\mu^{*n}(h+I) - \nu^{*n}(h+I)\right| = 0$$

On obtient ainsi le résultat escompté. □

Le théorème ci-dessus découle aussitôt de la combinaison du lemme précédent et de l'observation suivante :

Lemme 6.2. *Soit p une densité de probabilité continue sur $\mathbb{R}$. On suppose que f est une fonction uniformément continue et bornée sur $\mathbb{R}$ telle que la limite suivante existe*

$$\lim \frac{1}{T}\int_0^T f(t)dt = l$$

quand $|T| \to +\infty$. Alors

$$\lim_{T\to+\infty}\int_{\mathbb{R}} f(Tx)p(x)dx = l \tag{31}$$

Preuve. On choisit une fonction ϕ de classe C^1 sur $\mathbb{R}$ telle que $\|p - \phi\|_1 \leq \varepsilon$. Soit $F(x) = \int_0^x f(t)dt$ une primitive de f. On écrit pour tout $A > 0$

$$\int_{\mathbb{R}} f(Tx)\phi(x)dx = \int_{-A}^{A} f(Tx)\phi(x)dx + \varepsilon_1$$

où $|\varepsilon_1| \leq \|f\|_\infty \int_{|x|\geq A} p(x)dx$. Puis on intègre par parties

$$\int_{-A}^{A} f(Tx)\phi(x)dx = \left[\frac{1}{T}F(Tx)\phi(x)\right]_{-A}^{A} - \int_{-A}^{A} \frac{1}{T}F(Tx)\phi'(x)dx$$

et on fait tendre T vers $+\infty$. Il vient

$$\lim_{T\to+\infty} \int_{-A}^{A} f(Tx)\phi(x)dx = l[x\phi(x)]_{-A}^{A} - l\int_{-A}^{A} x\phi'(x)dx = l\int_{-A}^{A} \phi(x)dx$$

Comme A est arbitrairement grand, ε arbitrairement petit, et f bornée, on conclut aussitôt. □

6.1. Exemple

Les hypothèses du théorème ne peuvent être relâchées. On peut trouver en effet une fonction f bornée, de classe C^∞ et telle que

$$\lim_{n\to+\infty} \sup \int f d\mu_\alpha^n > 0$$

bien que quand $|T| \to +\infty$, $\lim \frac{1}{T}\int_0^T f(t)dt = 0$.

La mesure μ_α est diophantienne (cf. remarque 5.3) et définie à la section précédente à l'aide des suites $\varepsilon_n = n^{-\varepsilon}$, où $\varepsilon = \frac{1}{4}$, et $h_n = e^{-n}$. On peut construire aisément f comme précédemment, par exemple $f = \sum f_i$ en posant $a = 0$ et $b = 1/2$ et $a_i = N_i^{-a}$, $b_i = N_i^{-b}$.

6.2. Equidistribution de marches aléatoires

Le théorème 6.1 permet d'obtenir un résultat d'équidistribution probabiliste dès que l'on dispose du résultat déterministe correspondant. Plus précisément, soit X un espace localement compact et $(\phi_t)_t$ un flot agissant continûment sur X en y préservant une mesure borélienne finie m. Le corollaire qui suit montre que si la trajectoire d'un point $x \in X$ par le flot est équidistribuée par rapport à m, alors toute marche aléatoire centrée le long de cette trajectoire s'équidistribue de la même façon. On fixe une mesure de probabilité μ centrée et apériodique sur $\mathbb{R}$ avec un moment d'ordre 2 fini et S_n la marche aléatoire de loi μ^n associée.

Corollaire 6.1. *Supposons que la trajectoire $(\phi_t)_{t\in\mathbb{R}} \cdot x$ soit équidistribuée par rapport à m, c'est-à-dire*

$$\lim_{|T|\to+\infty} \frac{1}{T}\int_0^T f(\phi_t \cdot x)dt = \int_X f(y)dm(y)$$

pour toute fonction continue à support compact f sur X. Alors on a aussi

$$\lim_{n\to+\infty} \mathbb{E}(f(\phi_{S_n} \cdot x)) = \int_X f(y)dm(y)$$

La preuve découle aussitôt du théorème 6.1 appliquée à la fonction $t \mapsto f(\phi_t \cdot x)$ qui est bornée et unifomément continue sur $\mathbb{R}$, puisque f est à support compact. Dans cet énoncé, le fait que μ est centrée est essentiel.

7. Cas multi-dimensionnel

Tous les résultats précédents s'étendent sans difficulté au cas multi-dimensionnel. On énonce ici les principaux théorèmes valides sur $\mathbb{R}^d$, $d \geq 1$. Le cas échéant, on indique les modifications à apporter aux démonstrations.

On se place sur $\mathbb{R}^d$ où l'on note $\|x\|$ la norme euclidienne et $x \cdot y$ le produit scalaire de deux vecteurs. On considère une mesure de probabilité μ sur $\mathbb{R}^d$ qui est centrée et possède un moment d'ordre 2, c'est-à-dire

$$\int \|x\|^2 \, d\mu(x) < \infty$$

La loi μ se trouve donc dans le bassin d'attraction d'une certaine loi gaussienne, que l'on note ν. De plus on fera l'hypothèse que μ est *apériodique* sur $\mathbb{R}^d$, c'est-à-dire que le support de μ n'est pas contenu dans une classe d'un sous-groupe fermé propre de $\mathbb{R}^d$. Cette condition équivaut à la suivante

$$\forall t \in \mathbb{R}^d \backslash \{0\} \ |\widehat{\mu}(t)| < 1$$

où $\widehat{\mu}(t) = \int e^{it\cdot x} d\mu(x)$.

De manière analogue, on note $\{x\}$ la distance du point $x \in \mathbb{R}^d$ au réseau $\mathbb{Z}^d$ et $[x]$ un point de $\mathbb{Z}^d$ qui minimise la distance à x.

Enfin $S_n = X_1 + \cdots + X_n$ est à nouveau la marche aléatoire somme des variables indépendantes X_i qui sont toutes de même loi μ.

La définition des mesures diophantiennes s'étend aisément au cas multi-dimensionnel. Soit l un réel ≥ 0 :

Définition 7.1. *Une mesure de probabilité sur $\mathbb{R}^d$ est dite l- **diophantienne** s'il existe un réel $C > 0$ tel que pour tout $x \in \mathbb{R}^d$ assez grand en norme,*

$$\inf_{y \in \mathbb{R}} \int \{x \cdot a + y\}^2 d\mu(a) \geq \frac{C}{\|x\|^l}$$

*et μ est dite **diophantienne** si elle est l-diophantienne pour un certain l.*

De même, μ est l-diophantienne si et seulement s'il existe $C > 0$ telle que pour tout x assez loin de zéro

$$|\widehat{\mu}(x)| \leq 1 - \frac{C}{\|x\|^l}$$

Ainsi μ est diophantienne si et seulement si $\mu * \mu^{-1}$ est diophantienne. De même, μ est diophantienne si elle (ou une de ses puissances) n'est pas étrangère à la mesure de Lebesgue. Et lorsque μ est à support fini, la condition devient une condition sur le support S de μ uniquement.

L'analogue multi-dimensionnel de la construction faite au paragraphe (3.2) consiste à considérer les fonctions suivantes. Soit $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, on pose pour $p \geq d$

$$\beta(x) = \left(\sum_{i=1}^{d} \frac{1}{4} \sin^2(\pi x_i) \right)^p \sum_{n \in \mathbb{Z}^d} \frac{a_n}{\|x - n\|^{2p}}$$

Comme auparavant, β est dans $\mathbb{L}^1(\mathbb{R}^d)$ dès que la suite $(a_n) \in \ell^1(\mathbb{Z}^d)$ et dans ce cas $\widehat{\beta}$ est à support compact d'après le théorème de Paley-Wiener. De plus, si a_n est un $O(\frac{1}{\|n\|^{2p}})$ au voisinage de l'infini, alors $\beta(x)$ est un $O(\frac{1}{\|x\|^{2p-d}})$. Ainsi $\widehat{\beta}$ est de classe C^l dès que $2(p-d) > l$.

De plus, si les a_n sont tous ≥ 0 et non tous nuls, alors β est strictement positive sur $\mathbb{R}^d$ et de plus on a $\beta(x) \geq a_{[x]}$.

Pour toute fonction f on définit $C^k(f)$ et $C^k_r(f)$ par

$$C^k(f) = \max_{|\beta|\leq k} \left\| D^\beta f \right\|_1, C^k_r(f) = \max_{|\alpha|\leq r+1\ |\beta|\leq k} \left\{ \int |x_\alpha f|, \left\| D^\beta f \right\|_1 \right\}$$

où $\alpha = (i_1, \ldots, i_l)$ avec chaque i_j entier compris entre 1 et d et $|\alpha| := l$.

L'analogue des théorèmes (3.1) et (3.2) s'énonce de la façon suivante. La preuve est analogue.

Théorème 7.1. *Soit μ une mesure de probabilité centrée sur $\mathbb{R}^d$ ayant un moment d'ordre $r+2$ fini et une matrice de covariance $K > 0$. Soit $l \geq 0$. On suppose μ l-diophantienne. Alors il existe un réel $k_0 = k_0(l) \geq 0$ tel que pour toute fonction f de classe C^k avec $k > k_0$*

(i) si $C^k(f) < +\infty$, alors

$$\mathbb{E}(f(S_n)) = \sum_{p=0}^{r} \frac{1}{n^{p/2}} \int_{\mathbb{R}^d} f(x\sqrt{n}) P_p(x) g(x) dx + C^k(f) \cdot o\Big(\frac{1}{n^{(r+d)/2}}\Big)$$

(ii) si $C^k_r(f) < +\infty$, alors

$$(2\pi n)^{d/2} \sqrt{\det K}\, \mathbb{E}(f(S_n)) = \sum_{p=0}^{[r/2]} \frac{1}{n^p} \left\langle f, Q_p \right\rangle_{L^2(\mathbb{R}^d)} + C^k_r(f) \cdot o(\frac{1}{n^{r/2}})$$

où les P_p (resp. Q_p) sont des polynômes de $\mathbb{R}[x_1, \ldots, x_d]$ de degré total $\leq 3p$ (resp. $2p$) qui ne dépendent que des moments de μ d'ordre $\leq p+2$ (resp. $2p+2$), et les constantes intervenant dans les $o()$ ne dépendent que de r et de μ (g est la densité de la gaussienne ν associée à μ). On a $P_0 = Q_0 = 1$.

De même le théorème (3.3) a la généralisation ci-dessous. Le contre-exemple intervenant dans la preuve se généralise à l'identique en faisant intervenir la fonction β introduite ci-dessus.

Théorème 7.2. *Soit μ une mesure de probabilité sur $\mathbb{R}$ centrée, et ayant un moment d'ordre 3 fini. Alors μ est diophantienne si et seulement s'il existe un entier $k_0 \geq 0$ tel que pour toute fonction f à support compact et de classe C^k ($k \geq k_0$) sur $\mathbb{R}$, on a*

$$\sup_{t\in\mathbb{R}^d} \left| \int f(t+\cdot) d\mu^n - \int f(t+\cdot) d\nu^n \right| = O\Big(\frac{1}{n^{(d+1)/2}}\Big)$$

Le résultat sur les écarts modérés et le théorème local s'étendent eux-aussi avec la même preuve au cas multi-dimensionnel.

Théorème 7.3. *Soit r un entier ≥ 1. Supposons que μ est une mesure de probabilité sur $\mathbb{R}^d$, centrée, de matrice de covariance $K = Id$ et qui possède un moment d'ordre $r+2$ fini. Alors pour tout c, $0 < c < r$, on a uniformément en x pour $1 \leq x \leq \sqrt{c \log n}$*

$$\lim_{n\to+\infty} \frac{\mathbb{P}\left(||S_n|| > x\sqrt{n}\right)}{\mathbb{P}\left(||N|| > x\right)} = 1$$

où N, de loi ν, est une variable gaussienne centrée de matrice de covariance Id.

Le théorème de Stone (4.3) est valide sur $\mathbb{R}^d$ (voir [Sto]) et énonce que pour tout rectangle $R = \prod[-s_i, s_i]$ ($s_1, \ldots, s_d > 0$)

$$\sup_{x\in\mathbb{R}^d} \left|\mu^n(R+x) - \nu^n(R+x)\right| \leq \frac{\varepsilon_n(\mu)}{n^{d/2}}(1+|R|)$$

où $\varepsilon_n(\mu)$ est une suite tendant vers 0 et ne dépendant que de μ. La preuve est parfaitement analogue au cas $d = 1$. Le théorème (4.2) reste aussi vrai sur $\mathbb{R}^d$. Ainsi

Théorème 7.4. *Si μ est une mesure centrée et apériodique sur $\mathbb{R}^d$ et admettant un moment d'ordre $r \geq 2$. On suppose que la matrice de covariance de μ est Id et on note ν la loi gaussienne associée. Soit $R = \prod[-s_i, s_i]$ un rectangle borné centré en 0 et $s = \max s_i$. Alors pour tout $x \in \mathbb{R}^d$*

$$\lim_{n\to+\infty} \frac{\mu^n(R+x)}{\nu^n(R+x)} = 1. \tag{32}$$

De plus, si $r > 2$ (resp. $r = 2$) pour tout $c \in]0, r-2[$, la limite (32) est uniforme en x et s quand $||x|| + s \leq \sqrt{cn \log n}$ (resp. $||x|| + s = O(\sqrt{n})$) et tous les s_i sont minorés par un réel > 0.

Si de plus $\widehat{\mu}$ vérifie la condition de Cramér, alors il existe $\delta = \delta(\mu) > 0$ tel que la limite (32) est uniforme quand $||x|| + s \leq \sqrt{cn \log n}$ (resp. $||x|| + s = O(\sqrt{n})$ si $r = 2$) et tous les s_i sont minorés par $e^{-\delta n}$.

Le théorème (5.1) admet une généralisation évidente à $\mathbb{R}^d$, et la preuve est identique en substituant la fonction β ci-dessus à celle sur $\mathbb{R}$. Ainsi

Théorème 7.5. *Soit f une fonction Riemann intégrable sur $\mathbb{R}^d$ telle que*

$$\sum_{n\in\mathbb{Z}^d} \max_{[x]=n} |f(x)| < +\infty$$

Alors on a le théorème limite local pour f, c'est-à-dire que

$$\lim_{n\to+\infty} (2\pi n)^{d/2}\sqrt{\det K} \int f d\mu^n = \int f(x)dx$$

L'analogue du théorème (5.3) est le

Théorème 7.6. *Soit μ une mesure de probabilité centrée sur $\mathbb{R}^d$ et l un réel ≥ 0. On suppose de plus que μ possède moment fini d'ordre $3+d$ et on note ν la mesure gaussienne associée. Enfin, on suppose que μ est l-diophantienne. Soit k_0 un entier tel que $k_0 > 3l/2+1$. Alors toute fonction f définie sur $\mathbb{R}^d$ et telle que f et toutes ses dérivées jusqu'à l'ordre k_0 sont bornées vérifie la relation suivante*

$$\int f d\mu^n - \int f d\nu^n = o(\int |f| d\nu^n)$$

En particulier, si $f \geq 0$ non identiquement nulle

$$\frac{\int f d\mu^n}{\int f d\nu^n} \to 1$$

Si l'on suppose de plus que μ vérifie la condition de Cramér, alors (25) et (26) sont vérifiées pour toute fonction f höldérienne bornée sur $\mathbb{R}^d$.

Enfin, le théorème de la dernière section s'étend facilement à $\mathbb{R}^d$ en procédant de façon similaire en deux étapes. Le premier lemme utilise la version multi-dimensionnelle du théorème local uniforme de Stone, rappelée au début de cette section. On a

Théorème 7.7. *Soit μ une mesure de probabilité sur $\mathbb{R}^d$ centrée, apériodique et ayant un moment d'ordre 2 fini. On suppose que f est une fonction uniformément continue et bornée sur $\mathbb{R}^d$ telle que la limite suivante existe*

$$\lim \frac{1}{T_1 \cdots T_d} \int_0^{T_1} \cdots \int_0^{T_d} f(t) dt = \ell$$

quand $|T_i| \to +\infty$ pour tout $i = 1, \ldots, d$. Alors

$$\lim_{n \to +\infty} \int f d\mu^n = \ell$$

L'application à l'équidistibution des marches aléatoires s'étend verbatim aux flots multi-dimensionnels.

Remerciments 7.1. Je remercie vivement Martine Babillot pour ses remarques enrichissantes sur une version antérieure de l'article et pour l'attention qu'elle a porté, malgré la maladie, à ce travail.

References

[Amo] Amosova, N.N.: Local limit theorems for probabilities of moderate deviations. Lit. Mat. Sb. **14** (3),401–409 (1974)
[Bre] Breiman, L.: Probability. Addison-Wesley, (1968)
[Car1] Carlsson, H.: Error estimates in d-dimensional renewal theory. Compositio Math. **46** (2), 227–253 (1982)
[Car2] Carlsson, H.: Remainder term estimates of the renewal function. Ann. Probab. **11** (1), 143–157 (1983)

[Cra1] Cramér, H.: Les sommes et les fonctions de variables aléatoires. Paris Hermann, (1938)

[Cra2] Cramér, H.: Elements of probability theory and some of its applications. Cambridge, (1962)

[Fel] Feller, W.: An introduction to probability theory and its applications. Vol. II. Second edition John Wiley & Sons, (1971)

[GnK] Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent random variables. Addison-Wesley, (1954)

[Hin] Hinderer, K.: Remarks on Directly Riemann Integrable Functions. Math. Nachr. **130**, 225–230 (1987)

[Hög] Höglund, T.: A multi-dimensional renewal theorem. Bull. Sc. Math, 2ème série **112**, 111–138 (1988)

[Kee] Keener, R.: Asymptotic expansions in multivariate renewal theory. Stochastic Process. Appl. **34** (1), 137–153 (1990)

[Mic] Michel, R.: Results on probabilities of moderate deviations. Ann. Prob. (2), 349–353 (1974)

[Nag] Nagaev, A.V.: Renewal theorems in $\mathbb{R}^d$, Teor. Veroyatnost. i Primenen. **24** (3), 565–573 (1979)

[Nag2] Nagaev, A.V.: Integral limit theorems with regards to large deviations when Cramér's condition is not satisfied. Prob. Th. Appl. **14**, 51–63 (1969)

[Nag3] Nagaev, A.V.: Large deviations of sums of independant random variables. Ann. Prob. **7** (5), 745–789 (1979)

[RuS] Rubin H. et, Sethuraman, J.: Probabilities of moderate deviations. Sankhya Ser. A **27**, 325–346 (1965)

[Rud] Rudin, W.: Real and Complex Analysis, Addison-Wesley

[Sla] Slastnikov, A.D.: Limit theorems for moderate deviations probabilities. Probab. Theo. and Appl. **23** (2), 325–340 (1978)

[Sta] Stam, A.J.: Renewal theory in r dimensions I. Compositio Math. **21**, 383–399 (1969); II, **23**, 1–13 (1971)

[Sto] Stone, Ch.: A local limit theorem for nonlattice multi-dimensional distribution functions. Ann. Math. Stat **36**, 546–551 (1965)

[Sto2] Stone, Ch.: Ratio local and ratio limit theorems. 5th Berkeley Symp. on Math. Stat. and Prob. Berkeley and Los Angeles, UCP (1966), Vol II, **2**, pp. 217–224

[Sto3] Stone, Ch.: Application of unsmoothing and Fourier Analysis to random walks. Markov Processes and Potential Theory, Madison Wis, 165–192 (1967)