

# Théorème central limite presque sûr

Xiuneng Zhu  
sous la direction de Jean Bertoin

juin 2010

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Convergence faible</b>	<b>4</b>
2.1	Définitions et première propriétés . . . . .	4
2.2	Famille tendue et le théorème de Prokhorov . . . . .	8
2.3	L'étude spécifique sur l'espace $C[0, 1]$ . . . . .	9
<b>3</b>	<b>Mouvement brownien, théorèmes divers</b>	<b>11</b>
3.1	Introduction au mouvement brownien . . . . .	11
3.2	Le théorème de Donsker . . . . .	11
3.3	Le théorème de représentation de Skorokhod . . . . .	12
<b>4</b>	<b>L'approche de Brosamler</b>	<b>12</b>
4.1	Le cas d'un mouvement brownien . . . . .	13
4.2	Le cas des v.a. i.i.d . . . . .	17
<b>5</b>	<b>L'approche de Lacey &amp; Philipp</b>	<b>18</b>

# 1 Introduction

Nous connaissons tous le fameux théorème central limite classique dans la théorie de la probabilité, qui affirme que pour une suite i.i.d de v.a. de moyenne 0 et de variance 1, en notant  $S_n = X_1 + \dots + X_n$ , on a

$$\frac{S_n}{\sqrt{n}} \Longrightarrow \mathcal{N}(0, 1),$$

où  $\mathcal{N}(0, 1)$  désigne la loi normale standard. On se pose naturellement la question de savoir s'il y a une convergence presque sûr, comme dans la loi forte des grands nombres. La réponse est non en général, on peut même montrer le théorème suivant, qui implique que par contre, *p.s.*, la suite concernée ne converge pas.

**Théorème 1.1.** Soit  $(X_n)$  une suite i.i.d de v.a. de moyennes nulles et de variances finies, et on note  $(S_n)$  les sommes partielles. Alors *p.s.*,

$$\limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1.$$

Ce théorème, dû a Khinchin et Kolmogorov, porte le nom de la loi du logarithme itéré. Le but de cet exposé est de présenter un théorème central limite presque sûr, qui donne une convergence faible a presque tout point en donnant un poids harmonique à chaque terme. Plus précisément, soit  $S_k$  la  $k$  ième somme partielle d'une suite i.i.d de variables aléatoires  $X_i$  d'espérance 0, de variance 1 et qui admet un moment  $(2 + \delta)$  fini,  $\delta > 0$  définies sur un espace de probabilité  $(\Omega, \mathcal{F}, P)$ , alors il existe un ensemble de probabilité 1 tel que pour tout  $\omega$  dedans, on a

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta_{\frac{S_k(\omega)}{\sqrt{k}}} \Longrightarrow \mathcal{N}(0, 1),$$

où la convergence en jeu est la convergence faible des mesures sur  $\mathbb{R}$ . En fait on peut démontrer une version fonctionnelle plus puissante de ce théorème, qui joue un rôle similaire au théorème de Donsker par rapport au TCL classique. Enfin pour le démontrer on propose deux approches: celle de Brosamler, qui va être expliquée assez brièvement, traite d'abord le cas du mouvement brownien et puis étend au cas général en utilisant le théorème de

représentation de Skorokhod; et celle de Lacey & Phillip, plus courte, basée sur un principe d'invariance, qui va être expliquée en détails. Le plan du texte est le suivant.

Cette première partie joue un rôle introductif du texte.

Dans la deuxième partie on se rappelle quelques résultats importants sur la convergence faible des mesures. La notion d'une famille tendue de mesures est introduite, et puis on fait une brève étude sur l'espace  $C[0, 1]$  et les mesures sur cet espace.

La troisième partie est consacrée à des rappels sur le mouvement brownien, puis des théorèmes utilisés dans le reste du texte: le théorème de Donsker, le théorème de représentation de Skorokhod, etc.

La quatrième partie correspond à la démonstration due à Brosamler. Il y a une partie importante de cette preuve portant sur des estimations et inégalités concernant le mouvement brownien, qui s'avèrent assez lourdes, on se contentera donc d'avoir les idées principales et on omettra certaines démonstrations techniques.

La cinquième partie présente l'approche de Lacey & Philipp, relativement simple, basée sur un principe d'invariance presque sûr.

Enfin on termine avec deux appendices, la première donne une démonstration de la loi des grands nombres pour les v.a. quasi-orthogonales, et la deuxième introduit le théorème ergodique.

Ce mémoire de maîtrise a été réalisé sous la direction de Jean Bertoin, que je remercie pour ses nombreuses corrections, sa patience et sa disponibilité. Je remercie également Marie Thérêt et Amandine Véber pour leur aide précieuse.

## **2 Convergence faible**

### **2.1 Définitions et première propriétés**

On travaille sur un espace métrique séparable complet  $(S, d)$ , qui est muni naturellement la tribu borélienne, i.e. la plus petite tribu contenant les ouverts. On note  $\mathcal{M}$  l'ensemble des mesures finies sur  $(S, d)$ , et  $C_b(S)$  l'ensemble des fonctions continues bornée à valeurs dans  $\mathbb{R}$ . Alors on définit la convergence faible des éléments de  $\mathcal{M}$  comme suit

**Définition 2.1.** Soit  $(P_n)$  et  $P$  éléments dans  $\mathcal{M}$ , on dit que  $P_n$  converge faiblement vers  $P$ , ou  $P_n \Rightarrow P$ , si pour toute fonction  $f$  dans  $C_b(S)$ , on a

$$\int_S f dP_n \longrightarrow \int_S f dP.$$

On peut identifier les éléments dans  $\mathcal{M}$  aux éléments dans le dual topologique de  $C_b(S)$ , et la convergence faible correspond exactement à la convergence faible-\*. Remarquons que  $\mathbb{R}$  muni de la métrique usuelle et  $C[0, 1]$  muni de la métrique de la convergence uniforme sont des espaces métriques séparables complets, et c'est justement ces deux espaces qu'on va explorer. Donnons tout de suite une caractérisation de la convergence faible.

**Théorème (de Portmanteau) 2.2.** Les 4 assertions suivantes sont équivalentes

(i)  $P_n \Longrightarrow P$ .

(ii) Pour tout ouvert  $G$  dans  $S$ , on a

$$\liminf P_n(G) \geq P(G).$$

(iii) Pour tout fermé  $F$  dans  $S$ , on a

$$\limsup P_n(F) \leq P(F).$$

(iv) Pour tout ensemble mesurable  $A$  de  $S$  qui vérifie  $P(\partial A) = 0$ , on a

$$\lim P_n(A) = P(A).$$

Si on dispose de deux espaces métriques,  $(S, d)$  et  $(S', d')$  chacun muni de sa tribu borélienne, et une fonction  $h$  continue de  $S$  dans  $S'$ , en notant  $Ph^{-1}$  la mesure image de  $P$  par  $h$ , on a le théorème suivant, simple mais fort utile

**Théorème (the mapping theorem) 2.3.** Lorsqu'il y a une mesure  $P$  sur  $S$  et une fonction mesurable de  $S$  dans  $S'$ , on note  $Ph^{-1}$  la mesure image sur  $S'$ , alors si  $P_n \Longrightarrow P$  et  $h$  continue, on a  $P_n h^{-1} \Longrightarrow Ph^{-1}$ .

*Démonstration* On vérifie  $\forall f \in C_b(S')$ ,

$$\int_{S'} f d(P_n h^{-1}) = \int_S f \circ h dP_n \rightarrow \int_S f \circ h dP = \int_{S'} f d(Ph^{-1}),$$

d'ou le résultat.

Maintenant on cherche à trouver une métrique qui métrise la convergence faible des mesures de probabilités. Certes, il en existe plusieurs; on introduit d'abord la métrique de Prokhorov. On note  $A^\varepsilon$  l'ensemble des points de  $S$  qui ont une distance  $< \varepsilon$  à  $A$ .

**Définition 2.4.** Pour deux mesures de probabilité  $P$  et  $Q$  sur  $S$ , on définit la distance

$$\rho(P, Q) = \inf \{ \varepsilon > 0 \mid P(A) \leq Q(A^\varepsilon) + \varepsilon, \text{ pour tout borélien } A \}.$$

et le théorème suivant affirme l'intérêt de cette définition

**Théorème 2.5.**  $\rho$  définit une distance sur l'ensemble des mesures de probabilité.

Une autre distance provient de l'espace  $BL(S)$ , définie comme suit:

**Définition 2.6.** On note  $BL(S)$  l'ensemble des fonctions sur  $S$  à valeurs dans  $R$  qui sont bornées et qui vérifient

$$\|f\|_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} < \infty.$$

On munit  $BL(S)$  la norme

$$\|f\|_{BL} = \|f\|_\infty + \|f\|_L.$$

Encore une fois, l'ensemble des mesures de probabilité s'injecte naturellement dans le dual topologique de  $BL(S)$ . On peut donc munir une norme sur l'ensemble des mesures, qui est justement la norme usuelle dans  $BL(S)'$ , soit

$$\|P\| = \sup_{\|f\|_{BL} \leq 1} \left| \int_S f dP \right|,$$

et la distance associée, notée  $\beta$ ,

$$\beta(P, Q) = \sup_{\|f\|_{BL} \leq 1} \left| \int_S f dP - \int_S f dQ \right|.$$

On va voir que par une façon surprenante, pour les mesures de probabilité, la convergence faible dans  $BL(S)'$ , la convergence faible dans  $C_b(S)'$  et la convergence forte dans  $BL(S)'$  coïncident.

**Théorème 2.7.** Soit  $(S, d)$  un espace métrique qui est toujours supposé séparable complet, et  $P_n, P$  des mesures de probabilité sur  $S$ , les assertions suivantes sont équivalentes

- (i)  $P_n \implies P$ .
- (ii)  $\int_S f dP_n \longrightarrow \int_S f dP, \forall f \in BL(S)$ .
- (iii)  $\beta(P_n, P) \longrightarrow 0$ .
- (iv)  $\rho(P_n, P) \longrightarrow 0$ .

*Démonstration* Evidemment (i) implique (ii). Comme  $S$  est séparable et complet, par le théorème 2.9.,  $P$  est tendue, donc si on se donne  $\varepsilon > 0$ , il existe  $K$  compact de  $S$  tel que  $P(K) \geq 1 - \varepsilon$ . L'ensemble  $B := \{f \mid \|f\|_{BL} \leq 1\}$  restreint à  $K$  forme un compact dans l'espace  $(C(K), \|\cdot\|_\infty)$  par le théorème d'Arzela-Ascoli. Donc en tenant compte le fait que la norme  $\|\cdot\|_{BL}$  contrôle aussi le taux de croissance de la fonction, on a pour un  $k$  fixé, il existe  $f_1, f_2, \dots, f_k$  telles que pour toute fonction  $f$  dans  $B$ , il existe  $1 \leq j \leq k$  tel que  $\|f - f_j\|_{\infty, K} \leq \varepsilon$ , et

$$\sup \{ |f(x) - f_j(x)| \mid x \in K^\varepsilon \} \leq 3\varepsilon.$$

Soit  $g(x) := 0 \vee 1 - d(x, K) / \varepsilon$ , alors  $g \in BL(S)$  et  $1_K \leq g \leq 1_{K^\varepsilon}$ , donc pour  $n$  assez grand,  $P_n(K^\varepsilon) \geq \int_S g dP_n \geq 1 - 2\varepsilon$ . Enfin, pour chaque  $f$  et  $f_j$  déterminée ci-dessus,

$$\begin{aligned} \left| \int_S f d(P - P_n) \right| &\leq \int_S |f - f_j| d(P + P_n) + \left| \int_S f_j d(P_n - P) \right| \\ &\leq 2(P_n + P)(S/K^\varepsilon) + (3\varepsilon) \cdot 2 + \left| \int_S f_j d(P_n - P) \right| \\ &\leq 12\varepsilon + \left| \int_S f_j d(P_n - P) \right| \leq 13\varepsilon \end{aligned}$$

pour chaque  $j = 1, \dots, k$  et pour  $n$  suffisamment grand par (ii). D'où (iii).

Ensuite, pour montrer que (c) implique (b) : Etant donnée un borélien  $A$  et  $\varepsilon > 0$ , on pose  $f(x) := 0 \vee (1 - d(x, A) / \varepsilon)$ . Alors  $\|f\|_{BL} \leq 1 + \frac{1}{\varepsilon}$ . Pour

toutes mesures de probabilité  $P$  et  $Q$  sur  $S$ ,

$$\begin{aligned} Q(A) &\leq \int_S f dQ \\ &\leq \int_S f dP + \left(1 + \frac{1}{\varepsilon}\right) \beta(P, Q) \\ &\leq P(A^\varepsilon) + \left(1 + \frac{1}{\varepsilon}\right) \beta(P, Q). \end{aligned}$$

Ainsi  $\rho(P, Q) \leq \max\left(\varepsilon, \left(1 + \frac{1}{\varepsilon}\right) \beta(P, Q)\right)$ . Donc si  $\beta(P, Q) \leq \varepsilon^2$ ,  $\rho(P, Q) \leq \varepsilon + \varepsilon^2$ . Comme  $\rho(P, Q) \leq 1$ , et  $\rho(P, Q) \leq 2\beta(P, Q)^{\frac{1}{2}}$  par conséquent, dans tous les cas, d'où (d).

Pour montrer que (d) implique (a), en utilisant le théorème de Portman-teau, soit  $A$  tel que  $P(\partial A) = 0$  et  $\varepsilon > 0$ . Alors pour  $0 < \delta < \varepsilon$  et  $\delta$  assez petit,

$$P(A^\varepsilon \setminus A) < \varepsilon$$

et

$$P(A^{c\varepsilon} \setminus A^c) < \varepsilon$$

Puis pour  $n$  assez grand,

$$P_n(A) \leq P(A^\delta) + \delta \leq P(A) + \varepsilon$$

et

$$P_n(A^c) \leq P(A^{c\delta}) + \delta \leq P(A^c) + 2\varepsilon$$

Donc on a  $|(P - P_n)(A)| \leq 2\varepsilon$ , et enfin en tendant  $\varepsilon$  vers 0, on a  $P_n(A) \longrightarrow P(A)$ , d'où (i)

## 2.2 Famille tendue et le théorème de Prokhorov

**Définition 2.8.** On dit qu'une mesure finie  $\mu$  est tendue si  $\forall \varepsilon > 0$ , il existe  $K$  compact de  $S$  tel que

$$\mu(K^c) \leq \varepsilon.$$

On dit qu'une famille  $(\mu_i)_{i \in I}$  de mesures finies est tendue si  $\forall \varepsilon > 0$ , il existe  $K$  compact de  $S$  tel que  $\forall i \in I$ ,

$$\mu_i(K^c) \leq \varepsilon.$$

Cette définition traduit le fait que les masses des mesures ne fuient pas à l'infini.



**Théorème 2.9.** Si  $(S, d)$  est un espace métrique séparable complet, alors toute mesure finie  $\mu$  est tendue.

*Démonstration* Soit  $\varepsilon > 0$ , on exhibe un  $K$  possédant la propriété voulue. Comme  $S$  est séparable, pour tout  $n$ , il existe une famille dénombrable d'éléments dense, en prenant les boules de rayon  $\frac{1}{n}$  et de centres les éléments dans cette famille, on voit qu'elles recouvrent  $S$ . On peut donc extraire une famille finie de boules  $B_{i_1}, B_{i_2}, \dots, B_{i_n}$  telle que

$$\mu\left(\bigcup_{k \leq n} B_{i_k}\right) \geq \mu(S) - \frac{\varepsilon}{2^n}.$$

On note

$$A = \bigcap_n \bigcup_{k \leq n} B_{i_k},$$

alors l'adhérence de  $A$  est complète et par construction précompacte, donc compacte. On vérifie que

$$\mu(\overline{A}) \geq \mu(S) - \varepsilon,$$

et ceci termine la preuve.

On dit qu'une famille de mesures de probabilité est relativement compacte si elle est séquentiellement relativement compacte pour la topologie faible-\* dans  $C_b(S)$ , i.e. de toute suite on peut extraire une sous-suite qui converge faiblement.

**Théorème (de Prokhorov) 2.10.** Une famille de mesures de probabilité sur  $(S, d)$  séparable complet est tendue si et seulement si elle est relativement compacte.

### 2.3 L'étude spécifique sur l'espace $C[0, 1]$

En suivant les discussions ci-dessus,  $C[0, 1]$  muni de la norme infinie est sans doute un espace métrique séparable complet, et dans ce cadre on peut discuter les convergences faibles des mesures. Mais d'un point de vue probabiliste, c'est tout à fait naturel de considérer la tribu initiale sur  $C[0, 1]$  par rapport aux projecteurs des coordonnées: d'une façon précise, la plus petite tribu rendant mesurables les applications

$$\pi_x : f \mapsto f(x).$$

Le point de vue analytique et le point de vue probabiliste sont étroitement liés, en fait, la tribu borélienne et la tribu initiale coïncident, ce qui nous facilite beaucoup notre travail. Dorénavant, pour vérifier qu'une fonction à valeur dans  $C[0, 1]$  est mesurable, il suffit de vérifier que la composée de cette fonction avec les projecteurs de coordonnée est mesurable, ce qui est plus aisé dans la plupart du temps.

**Théorème 2.11.** La tribu borélienne et la tribu initiale définie ci-dessus coïncident.

Les familles tendues de mesures sur  $C[0, 1]$  possède des critères particuliers. Le critère ci-dessus vient du théorème 12.3 du [1] à une légère modification près.

**Théorème 2.12.** Soit  $(\mu_i)_{i \in I}$  une famille de mesures de probabilité sur  $C[0, 1]$  qui vérifie

(i) la famille de mesures  $(\mu_i \pi_0^{-1})_{i \in I}$  est tendue.

(ii) Il existe une constante  $\gamma > 0$ ,  $\alpha > 1$  et une fonction  $F$  croissante continue sur  $[0, 1]$  telle que  $\forall t_1, t_2 \in [0, 1], i \in I$ ,

$$\int_{C[0,1]} |x(t_2) - x(t_1)|^\gamma \mu_i(dx) \leq |F(t_2) - F(t_1)|^\alpha,$$

alors  $(\mu_i)_{i \in I}$  est tendue.

En plus, on peut montrer un critère important de convergence faible sur  $C[0, 1]$ :

**Théorème 2.13.** Soit  $P_n, P$  des mesures de probabilité sur  $C[0, 1]$ . Si  $(P_n)$  est une famille tendue et si les distributions fini-dimensionnelles de  $P_n$  convergent faiblement vers celles de  $P$ , alors

$$P_n \Longrightarrow P.$$

### 3 Mouvement brownien, théorèmes divers

#### 3.1 Introduction au mouvement brownien

Le mouvement brownien, ou processus de Wiener, est une description mathématique du mouvement aléatoire d'une « grosse » particule immergée dans un fluide et qui n'est soumise à aucune autre interaction que des chocs avec les « petites » molécules du fluide environnant. La définition mathématique précise est la suivante.

**Définition 3.1.** Soit  $(\Omega, \mathcal{F}, P)$  un espace de probabilité. On dit que le processus  $(W_t)_{t \geq 0}$  est un mouvement brownien standard unidimensionnel issu de 0 si

(i)  $\forall 0 < t_1 < t_2 \dots < t_k$ , on a que

$$W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_{k-1}} - W_{t_{k-2}}$$

sont des variables aléatoires gaussiennes indépendantes et de lois

$$W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1}).$$

(ii) Les trajectoires browniennes  $t \mapsto W_t$  sont presque sûrement continues sur  $\mathbb{R}^+$ .

#### 3.2 Le théorème de Donsker

Soit  $(X_n)_{n \geq 1}$  une suite de v.a. définie sur  $(\Omega, \mathcal{F}, P)$ , on définit le processus de ligne brisée associé suivant: Soit  $S_n = \sum_{1 \leq i \leq n} X_i$ , on pose  $s_n(t, \omega)$  un élément dans  $C[0, 1]$  tel que

$$\forall 0 \leq i \leq n, s_n\left(\frac{i}{n}, \omega\right) = \frac{S_i(\omega)}{\sqrt{n}}, \quad (1)$$

et  $s_n$  affine sur  $\left[\frac{i}{n}, \frac{i+1}{n}\right]$ ,  $\forall 0 \leq i \leq n-1$ . Alors on vérifie aisément que  $s_n : \Omega \rightarrow C[0, 1]$  est mesurable. Ceci définit donc un nouveau processus sur  $\Omega$ , et on a

**Théorème (de Donsker) 3.2.** Si  $(X_n)_{n \geq 1}$  est i.i.d de moyenne 0 et de variance 1, alors

$$s_n \Longrightarrow \mu_W.$$

On adopte parfois le nom d'un principe d'invariance à ce théorème car la limite est la même pour toute suite vérifiant les hypothèses. On peut retrouver le TCL à partir du théorème de Donsker, car

$$\pi : f \mapsto f(1)$$

est une fonction continue de  $C[0, 1]$  à  $\mathbb{R}$ , et  $\forall n \geq 1$ ,

$$s_n(1, \omega) = \frac{S_n(\omega)}{\sqrt{n}}.$$

De plus la mesure image de  $\mu_W$  par  $\pi$  est bien  $\mathcal{N}(0, 1)$ , par Théorème 2.3., on trouve *p.s.*,

$$\frac{S_n}{\sqrt{n}} \Longrightarrow \mathcal{N}(0, 1),$$

ce qui est le théorème central limite.

### 3.3 Le théorème de représentation de Skorokhod

**Théorème (de représentation de Skorokhod) 3.3.** Soit  $(X_n)_n$  est une suite i.i.d de v.a. de moyenne 0 et de variance finie, et  $(S_n)_n$  la suite des sommes partielles. Alors il existe une suite croissante de temps d'arrêt dans la filtration naturelle de  $(X_n)$   $\tau_1, \tau_2, \dots$  tels que  $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$  sont i.i.d et

$$(W_{\tau_n})_n = (S_n)_n$$

en loi. De plus on peut imposer que  $\forall n \geq 1$ ,

$$E[\tau_n - \tau_{n-1}] = E[X_1^2].$$

Ce théorème nous permet de montrer d'abord les résultats sur les mouvements browniens, puis l'étendre aux cas des v.a. i.i.d.

## 4 L'approche de Brosamler

On se rappelle le théorème de départ:

**Théorème (central limite presque sûr)** Soit  $(X_n)$  une suite i.i.d de variables aléatoires sur  $(\Omega, F, P)$  de moyenne 0 et de variance 1, qui admet un moment  $(2 + \delta)$  ème fini pour un  $\delta$  strictement positif, et on note  $S_k$  la somme partielle, alors il existe un ensemble  $N$   $P$ -négligeable tel que pour tout  $\omega$  dans  $N^c$ , on a

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} \delta_{\frac{S_k}{\sqrt{k}}} \implies \mathcal{N}(0, 1).$$

Comme une analogie du théorème de Donsker, on va prouver une version fonctionnelle de ce théorème:

**Théorème (central limite presque sûr, version fonctionnelle)** On note encore  $s_n(t, \omega)$  l'élément dans  $C[0, 1]$  défini dans ?, alors sous les hypothèses du théorème précédent, on a  $P$  presque sûrement,

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} \delta_{s_k(\omega)} \implies \mu_W, \quad (2)$$

où  $\mu_W$  désigne toujours la mesure de Wiener sur  $C[0, 1]$ .

On va montrer un résultat analogue pour un mouvement brownien, puis l'étendre au cas général grâce au théorème de représentation de Skorokhod.

#### 4.1 Le cas d'un mouvement brownien

Dans la suite de cette sous-section on suppose que  $W$  un mouvement brownien canonique sur  $(\Omega, \mathcal{F}, P)$ , et  $\forall s > 0$ ,

$$W_t^{(s)} = \frac{1}{\sqrt{s}} W_{st},$$

qui s'agit d'un changement d'échelle, et on sait que  $W^{(s)}$  reste un mouvement brownien canonique. On pose

$$\mu_t(\omega) = \frac{1}{\log t} \int_1^t \frac{\delta_{W^{(s)}(\omega)}}{s} ds$$

qui est la moyenne harmonique des mesures. Le premier résultat qu'on va montrer dans cette sous-section est

**Théorème 4.1.**  $\lim_{t \rightarrow \infty} \mu_t = \mu_W$ .

La convergence en jeu est bien sûr la convergence faible sur  $C[0, 1]$ . Pour montrer ce résultat on s'aide à utiliser la notion d'une famille tendue de mesures introduite dans le deuxième chapitre, plus précisément on a

**Lemme 4.2.** La famille  $(\mu_t, t \geq e)$  est tendue.

On vérifie d'abord que  $\forall t, \mu_t \pi_0^{-1} = \delta_0$ , la famille  $(\mu_t \pi_0^{-1})_{t \geq e}$  est tendue sur  $\mathbb{R}$ . En vertu du Théorème 2.12., en prenant  $F$  une fonction affine croissante,  $\alpha = \frac{3}{2}$  et  $\gamma = 4$ , il suffit de montrer que *p.s.*, il existe une constante  $C$  (dépend de  $\omega$ ) tel que  $\forall u_1, u_2 \in [0, 1]$ ,

$$\sup_{t \geq e} \int_{C[0,1]} |x(u_1) - x(u_2)|^4 \mu_t(\omega)(dx) \leq C(\omega) |u_1 - u_2|^{\frac{3}{2}}.$$

Par définition de  $\mu_t$ , ceci se réécrit comme

$$\sup_{t \geq e} \frac{1}{\log t} \int_1^t \frac{ds}{s} \left| \frac{W_{u_1 s}}{\sqrt{s}} - \frac{W_{u_2 s}}{\sqrt{s}} \right|^4 \leq C(\omega) |u_1 - u_2|^{\frac{3}{2}}.$$

On essaye de majorer vaguement le terme

$$|W_{u_1 s} - W_{u_2 s}|,$$

qui s'agit d'avoir un contrôle global sur la vitesse de croissance d'un mouvement brownien, qui vient du

**Lemme 4.3.**  $\forall \varepsilon < \frac{1}{2}$  et  $p \geq 1$  on a

$$\sup_{0 \leq u_1 < u_2 \leq 1} \frac{|W_{u_2} - W_{u_1}|}{(u_2 - u_1)^\varepsilon} \in L^p(\Omega).$$

On admet ce lemme, qui peut être prouvé par estimation élémentaire ou par la théorie générale des processus gaussiens. A partir de cela, on montre

**Lemme 4.4.**  $\forall p \geq 1$ ,

$$\sup_{t \geq e} \sup_{0 \leq u_1 < u_2 \leq 1} \frac{1}{\log t} \frac{1}{(u_2 - u_1)^{\frac{3}{2}}} \int_1^t \frac{ds}{s} \frac{|W_{u_1 s} - W_{u_2 s}|^4}{s^2} \in L^p(\Omega).$$

*Démonstration* On pose

$$Z_s = \sup_{0 \leq u_1 < u_2 \leq 1} \frac{|W_{u_1 s} - W_{u_2 s}|^4}{s^2 (u_1 - u_2)^{\frac{3}{2}}}.$$

Alors la v.a. dans le lemme est majorée par

$$Z^* = \sup_{t \geq e} \frac{1}{\log t} \int_1^t \frac{Z_s}{s} ds = \sup_{t \geq 1} \frac{1}{t} \int_1^{e^t} \frac{Z_s}{s} ds.$$

Or en posant

$$X_s = \frac{1}{\sqrt{s}} W_s, s > 0,$$

on vérifie aisément que le processus  $X$  est invariant par multiplication par rapport à l'échelle du temps, i.e.  $\forall h > 0, (X_s)_s = (X_{hs})_s$  en loi. On remarque que

$$Z_s = \sup_{0 \leq u_1 < u_2 \leq 1} \frac{|\sqrt{u_1} X_{u_1 s} - \sqrt{u_2} X_{u_2 s}|^4}{(u_1 - u_2)^{\frac{3}{2}}}$$

l'est aussi, et ceci nous fournit des processus stationnaires et surtout une occasion d'appliquer le théorème ergodique. En effet, par le lemme,  $Z_1 \in L^p, \forall p \geq 1$ , donc par le théorème ergodique dominé (c.f. [5], p.189, théorème 2.5.), on a

$$\sup_{n \geq 1} \frac{1}{n} \int_1^{e^n} \frac{Z_s}{s} ds = \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} \int_k^{k+1} Z_{e^u} du \in L^p(\Omega).$$

Maintenant pour un  $t \geq 1$  arbitraire, on écrit

$$\frac{1}{t} \int_1^{e^t} \frac{Z_s}{s} ds \leq \frac{[t] + 1}{t} \frac{1}{[t] + 1} \int_1^{e^{[t]+1}} \frac{Z_s}{s} ds \leq 2 \sup_{n \geq 1} \frac{1}{n} \int_1^{e^n} \frac{Z_s}{s} ds,$$

d'où  $Z^* \in L^p, \forall p \geq 1$ .

Une fonction dans  $L^p$  est en particulier finie p.s.. **On a donc prouvé que la famille  $(\mu_t)_{t \geq e}$  est tendue.**

Maintenant on est prêt pour la démonstration du Théorème 4.1..

*Démonstration du Théorème 4.1.* En vertu du Théorème 2.13., il ne nous reste plus qu'à montrer que p.s., les distributions fini-dimensionnelles de  $\mu_t$

convergent faiblement vers celles de  $\mu_W$ . On utilise le théorème de Lévy et étudie les fonctions caractéristiques. Soient  $k \in N$ ,  $0 < u_1 < u_2 < \dots < u_k \leq 1$  et  $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ , posons

$$\phi(x) = \exp(i(\alpha_1 x(u_1) + \dots + \alpha_k x(u_k))).$$

On veut montrer que *p.s.*,

$$\lim_{t \rightarrow \infty} \int_{C[0,1]} \phi(x) \mu_t(\omega)(dx) = \int_{C[0,1]} \phi(x) \mu_W(dx). \quad (3)$$

Par définition même,

$$\int_{C[0,1]} \phi(x) \mu_t(\omega)(dx) = \frac{1}{\log t} \int_1^t \frac{\Psi_s}{s} ds,$$

où  $\Psi_s = \exp(i(\alpha_1 \sqrt{u_1} X_{u_1 s} + \dots + \alpha_k \sqrt{u_k} X_{u_k s}))$ . Par le même argument qu'avant, on a  $\forall h > 0$ ,  $(\Psi_s)_{s \geq 0} = (\Psi_{hs})_{s \geq 0}$  en loi, donc en utilisant le théorème ergodique et la loi du 0-1 pour le mouvement brownien,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_1^{e^n} \frac{\Psi_s}{s} ds = E[\Psi_1].$$

De plus pour  $t > 0$  arbitraire,

$$\left| \frac{1}{t} \int_1^{e^t} \frac{\Psi_s}{s} ds - \frac{[t]}{t} \frac{1}{[t]} \int_1^{e^{[t]}} \frac{\Psi_s}{s} ds \right| = \frac{1}{t} \left| \int_{e^{[t]}}^{e^t} \frac{\Psi_s}{s} ds \right| \leq \frac{1}{t},$$

donc

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_1^{e^t} \frac{\Psi_s}{s} ds = E[\Psi_1] = \int_{C[0,1]} \phi(x) \mu_W(dx).$$

Enfin *p.s.*, (3) est vrai pour tout  $k \in N$ ,  $0 < u_1 < \dots < u_k \leq 1$  dans  $\mathbb{Q}$  et  $\alpha_1, \dots, \alpha_k$  dans  $\mathbb{Q}$ . On a donc le résultat.

Par un effort, on peut montrer une version discrète du théorème précédent, dont on omet ici la preuve:

**Théorème 4.5.** On pose

$$v_n(\omega) = \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta_{W^{(k)}(\omega)},$$

alors

$$\lim_{n \rightarrow \infty} v_n = \mu_W, \text{ p.s..}$$



## 4.2 Le cas des v.a. i.i.d

Soit  $(X_n)$  une suite i.i.d de v.a. vérifiant les hypothèses dans le théorème. D'après le théorème de représentation de Skorokhod, on peut trouver un autre espace de probabilité, un mouvement brownien  $W$  dessus et une suite croissante de temps d'arrêt  $\tau_1, \tau_2, \dots$  tels que  $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$  sont i.i.d et

$$(W_{\tau_n})_n = (S_n)_n$$

en loi. En utilisant l'hypothèse que  $X_n$  admet un moment  $(2 + \delta)$  fini, on peut imposer de plus que

$$E \left[ \tau_1^{1+\delta} \right] < \infty.$$

Sous ces hypothèses on admet le résultat suivant

### Lemme 4.6.

$$\lim_{k \rightarrow \infty} \frac{|W_{\tau_k} - W_k|}{\sqrt{k}} = 0, p.s.$$

Avant de commencer la preuve du théorème principal, on cite enfin un lemme, dont la démonstration est claire:

**Lemme 4.7.** Soient  $(X_n)$  et  $(Z_n)$  deux suites de v.a. à valeurs dans  $C[0, 1]$ . Alors si

$$\lim_{n \rightarrow \infty} \|X_n - Z_n\|_\infty = 0, p.s.$$

alors *p.s.*, pour toute fonction  $f : C[0, 1] \rightarrow \mathbb{R}$  bornée et uniformément continue,

$$\lim_{n \rightarrow \infty} \left| \frac{1}{\log n} \sum_{k=1}^n \frac{f(X_k)}{k} - \frac{1}{\log n} \sum_{k=1}^n \frac{f(Z_k)}{k} \right| = 0.$$

On identifiera le processus  $(S_n)$  et le processus  $(W_{\tau_n})$ . Soit  $(s_n)$  le processus associé à  $(S_n)$  dans le théorème de Donsker. D'après le lemme, pour démontrer le résultat final, il suffit d'établir

### Théorème 4.8.

$$\lim_{n \rightarrow \infty} \|W^{(n)} - s_n\|_\infty = 0, p.s.. \quad (4)$$

*Démonstration* Pour faire un pont entre  $W^{(n)}$  et  $s_n$ , on définit aussi un processus de ligne brisée associé à  $W$ : on pose

$$Z^n : \Omega \longrightarrow C[0, 1],$$

tel que

$$Z^n \left( \frac{i}{n} \right) = \frac{W_i}{\sqrt{n}}, i = 0, 1, \dots, n$$

et affine sur les intervalles  $[\frac{i}{n}, \frac{i+1}{n}]$ . Alors (4) vient du fait que

$$\lim_{n \rightarrow \infty} \| W^{(n)} - Z^n \|_{\infty} = 0, p.s. \quad (5)$$

et

$$\lim_{n \rightarrow \infty} \| Z^n - s_n \|_{\infty} = 0, p.s. \quad (6)$$

Pour montrer (6) il suffit d'écrire

$$\| Z^n - s_n \|_{\infty} = \sup_{1 \leq k \leq n} \frac{|W_{\tau_k} - W_k|}{\sqrt{n}}$$

et le terme de droite tend vers 0 par Lemme 4.6.. Pour montrer (5) on écrit

$$\begin{aligned} P \left( \| W^{(n)} - Z^n \|_{\infty} \geq \varepsilon \right) &\leq \sum_{k=0}^{n-1} P \left( \sup_{k \leq t \leq k+1} |W_t - W_k| \geq \frac{\varepsilon \sqrt{n}}{2} \right) \\ &\leq n P \left( \sup_{0 \leq t \leq 1} |W_t| \geq \frac{\varepsilon \sqrt{n}}{2} \right) \\ &\leq C n e^{-\frac{1}{2} \varepsilon^2 n}. \end{aligned}$$

Par le lemme de Borel-Cantelli on a le résultat. Ceci termine la preuve.

## 5 L'approche de Lacey & Philipp

L'approche de Lacey & Philipp diffère de celle de Brosamler par l'utilisation d'un principe d'invariance presque sûr, un cas particulier du Théorème 1 dans [6].

**Principe d'invariance presque sûr** On se donne une mesure de probabilité  $\lambda$  sur  $\mathbb{R}$  vérifiant

$$\int x d\lambda = 0, \int x^2 d\lambda = 1$$

et il existe  $\delta > 0$  tel que

$$\int |x|^{2+\delta} d\lambda < \infty$$

alors il existe un espace  $(\Omega, \mathcal{F}, P)$  et deux suites de v.a. dessus,  $(X_n)$  i.i.d chacune de loi  $\lambda$ , et  $(Y_n)$  i.i.d chacune de loi  $\mathcal{N}(0, 1)$ , vérifiant *p.s.*,

$$\sum_{k=1}^n (X_k - Y_k) = o(\sqrt{n}).$$

Donc il suffit évidemment de montrer

**Théorème 5.1.** Soit  $(X_n)_n$  une suite de v.a. telle qu'il existe une suite i.i.d de gaussiennes  $(Y_n)_n$  et *p.s.*,

$$\sum_{k=1}^n (X_k - Y_k) = o(\sqrt{n}), \quad (7)$$

alors (2) est vérifiée.

La preuve de ce résultat se décompose en 4 étapes, qu'on va détailler dans la suite.

**Etape 1** D'après Théorème 2.7., il suffit d'établir que *p.s.*,  $\forall f \in BL(C[0, 1])$ ,

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} f(s_k(\cdot, \omega)) \longrightarrow \int_{C[0,1]} f d\mu_w.$$

On affirme qu'on peut mettre le *p.s.* derrière le  $\forall f$ . A priori ce n'est pas légitime, car il y a évidemment un nombre non dénombrable d'éléments dans  $BL$ . On s'inspire de la démonstration du Théorème 2.7., et on voit qu'on peut trouver une sous famille dénombrable de  $BL$  telle que si *p.s.* la convergence

a lieu pour les éléments dans cette famille, alors automatiquement la convergence a lieu pour tout élément dans  $BL$ . Donc le problème est transformé à montrer que  $\forall f \in BL, p.s.$ ,

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} f(s_k(\cdot, \omega)) \longrightarrow \int_{C[0,1]} f d\mu_W.$$

**Etape 2** On fixe un  $f \in BL$ . Par (7), en notant  $(t_k)_k$  le processus de lignes brisées associé à  $(Y_k)_k$ , on a  $p.s.$

$$\|s_k(\omega) - t_k(\omega)\|_\infty \leq \frac{\max_{1 \leq i \leq k} |\sum_{j=1}^i (X_j - Y_j)|}{\sqrt{k}} = o(1).$$

Donc  $|f(s_k(\omega) - f(t_k(\omega)))| \leq \|f\|_\infty \|s_k(\omega) - t_k(\omega)\| = o(1)$ , et par un argument classique de type Cesaro, on voit que

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} f(s_k)$$

et

$$\frac{1}{\log n} \sum_{k \leq n} \frac{1}{k} f(t_k)$$

ont la même limite, donc il suffit de montrer le cas d'une suite i.i.d de gaussiennes.

**Etape 3** D'après l'étape 2, notre but est de montrer que  $\forall f \in BL, p.s.$ ,

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} f(t_k(\omega)) \longrightarrow \int_{C[0,1]} f d\mu_W.$$

On se rappelle que le processus  $(t_k)$  provient d'une suite i.i.d de v.a. gaussiennes, donc d'après le théorème de Donsker,

$$t_k \Longrightarrow \mu_W.$$

Par conséquent,

$$E[f(t_k)] \longrightarrow E[f(W)] = \int_{C[0,1]} f d\mu_W.$$

On regarde maintenant la variable aléatoire

$$\xi_k(\omega) = f(t_k(\omega)) - E[f(t_k)]$$

il suffit donc d'établir que *p.s.*,

$$\frac{1}{\log n} \sum_{1 \leq k \leq n} \frac{\xi_k}{k} \longrightarrow 0.$$

**Etape 4** On a besoin d'un lemme.

**Lemme 5.2.** Il existe une constante  $C$  qui ne dépend que de  $\|f\|_{BL}$  telle que  $\forall j < k$ ,

$$|E[\xi_j \xi_k]| \leq C \sqrt{\frac{j}{k}}.$$

*Démonstration* On note  $T_k = \sum_{j \leq k} Y_j$ , et en fixant  $j < k$ , on définit pour  $0 \leq t \leq 1$ ,

$$r(t) = 0, 0 \leq t \leq \frac{j}{k}$$

et

$$r(t) = t_k(t) - \frac{T_j}{\sqrt{k}}, \frac{j}{k} \leq t \leq 1.$$

Alors ponctuellement, *p.s.*,

$$|r(t, \omega) - t_k(t, \omega)| \leq \max_{i \leq j} \frac{|T_i(\omega)|}{\sqrt{k}} + \frac{|T_j|}{\sqrt{k}}.$$

Donc, comme  $T_i$  suit une loi normale, il existe une constante indépendante de  $j$  et  $k$ ,

$$\sqrt{k} E[\|t_k - r\|_\infty] \leq C \sqrt{j},$$

et comme  $f$  lipschizienne on a une constante  $C'$ ,

$$E[\|f(t_k) - f(r)\|] \leq C' \sqrt{\frac{j}{k}}$$

On pose

$$Z_l = \sum_{4^{l-1} \leq j < 4^l} \frac{\xi_j}{j}, l \geq 1$$

Par le lemme,  $\forall l < m$ ,

$$\begin{aligned}
E[Z_l Z_m] &= \sum_{4^{l-1} \leq j < 4^l} \sum_{4^{m-1} \leq k < 4^m} \frac{1}{jk} E[\xi_j \xi_k] \\
&\leq \sum_{4^{l-1} \leq j < 4^l} \sum_{4^{m-1} \leq k < 4^m} \frac{C}{jk} \sqrt{\frac{j}{k}} \\
&\leq \frac{C}{\sqrt{4^{l-1}} (\sqrt{4^{m-1}})^3} (3 \cdot 4^{l-1} \cdot 3 \cdot 4^{m-1}) \\
&\leq \frac{C'}{2^{m-l}}
\end{aligned}$$

où  $C'$  est indépendant de  $l$  et  $m$ . Par les mêmes calculs, on voit que  $(\xi_n)$  et  $(Z_n)$  sont bornées dans  $L^2$ , donc en appliquant la loi forte des grands nombres pour les v.a. quasi-orthogonales, on a

$$\frac{1}{n} \sum_{l \leq n} Z_l \longrightarrow 0, p.s..$$

On a ainsi montré que *p.s.*,

$$\frac{1}{\log 4^n} \sum_{k \leq 4^n} \frac{\xi_k}{k} \longrightarrow 0.$$

On pose maintenant

$$D_n = \sup_{4^n \leq k < 4^{n+1}} \left| \sum_{i=1}^k \frac{\xi_i}{i} - (Z_1 + \dots + Z_n) \right|,$$

et on utilise encore une fois le fait que les  $(\xi_n)$  sont faiblement dépendantes pour obtenir que  $(D_n)$  est bornée dans  $L^2$ , donc par l'inégalité de Chebychev,

$$P(D_n \geq \varepsilon n) \leq \frac{E[D_n^2]}{\varepsilon^2 n^2} \leq \frac{A}{\varepsilon^2 n^2},$$

et donc

$$\sum P(D_n \geq \varepsilon n) < \infty.$$

Par le lemme de Borel-Cantelli,

$$\frac{D_n}{n} \longrightarrow 0, p.s..$$

Enfin  $\forall k$  tel que  $4^n \leq k < 4^{n+1}$ , on écrit

$$\begin{aligned} \frac{1}{\log k} \left| \sum_{i=1}^k \frac{\xi_i}{i} \right| &\leq \frac{|Z_1 + \dots + Z_n|}{\log k} + \frac{|D_n|}{\log k} \\ &\leq \frac{|Z_1 + \dots + Z_n|}{\log 4^n} + \frac{|D_n|}{\log 4^n}, \end{aligned}$$

le terme à droite tend vers 0. Ceci termine la preuve.

## APPENDICE A Loi des grands nombres pour v.a quasi-orthogonales

La loi des grands nombres peut être étendue à un cas des variables aléatoires non pas indépendantes et identiquement distribuées, mais avec une corrélation assez faible. Ici on cherche pas l'hypothèse la plus faible, le théorème ci-dessous va nous suffire à établir le théorème dans le chapitre précédent.

**Théorème** Soit  $(X_n)_n$  une suite de v.a. de moyennes nulles et bornées dans  $L^2$ , qui vérifie

$$\exists C > 0, \lambda < 1, \forall i < j, E[X_i X_j] \leq C \lambda^{j-i}$$

en notant  $S_n = \sum_{i \leq n} X_i$ , on a

$$\frac{S_n}{n} \longrightarrow 0, p.s..$$

*Démonstration* Cette démonstration illustre une méthode importante, qui s'agit de considérer d'abord une sous-suite puis étendre à la suite toute entière. On pose  $\forall i, E[X_i^2] \leq M$ .

$$E[S_n^2] = \sum_{i=1}^n E[X_i^2] + 2 \sum_{i < j} E[X_i X_j] \leq nM + \frac{2n}{\lambda - 1},$$

donc d'après l'inégalité de Chebychev,  $\forall \varepsilon$ ,

$$P(|S_n| \geq n\varepsilon) \leq \frac{n \left( M + \frac{2}{\lambda - 1} \right)}{n^2 \varepsilon^2} = \frac{M + \frac{2}{\lambda - 1}}{n \varepsilon^2}.$$

On se restreint sur la sous suite  $S_{n^2}$ , ce qui rend la sommation des termes à droite convergente:

$$\sum_{n \geq 0} P \left( |S_{n^2}| \geq n^2 \varepsilon \right) < \infty.$$

Donc d'après lemme de Borel-Cantelli, *p.s.*,  $\frac{|S_{n^2}|}{n^2} \geq \varepsilon$  pour un nombre fini de  $n$ , d'où

$$\frac{S_{n^2}}{n^2} \longrightarrow 0, p.s..$$

Puis on essaye de comparer les termes généraux de  $(S_n)$  et ceux de la sous-suite  $(S_{n^2})$  pour voir que la différence est assez petite. On pose

$$D_n := \max_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}|,$$

alors

$$D_n^2 \leq \sum_{k=n^2}^{(n+1)^2} |S_k - S_{n^2}|^2,$$

en tenant compte que

$$\begin{aligned} E \left[ |S_{n^2+k} - S_{n^2}|^2 \right] &= E \left[ |X_{n^2+1} + \dots + X_{n^2+k}|^2 \right] \\ &\leq kM + 2k \frac{1}{1-\lambda} \\ &\leq 2nM + \frac{2n}{1-\lambda} \end{aligned}$$

On a

$$E \left[ D_n^2 \right] \leq E \left[ \sum_{k=n^2}^{n^2+2n} (S_k - S_{n^2})^2 \right] \leq 2n \left( 2nM + \frac{2n}{1-\lambda} \right).$$

Donc

$$P \left( D_n \geq n^2 \varepsilon \right) \leq \frac{4M + \frac{4}{1-\lambda}}{n^2 \varepsilon^2},$$

et *p.s.*,  $\frac{D_n}{n} \longrightarrow 0$ . Or  $\forall k$  tel que  $n^2 \leq k \leq (n+1)^2$ ,

$$\frac{|S_k|}{k} \leq \frac{|S_{n^2}| + D_n}{k} \leq \frac{|S_{n^2}|}{n^2} + \frac{D_n}{n^2}.$$

On conclut que *p.s.*,

$$\frac{S_n}{n} \longrightarrow 0.$$

Ceci termine la preuve.



## APPENDICE B Théorème ergodique et suites stationnaires

Dans cette appendice on discute le théorème ergodique, qui est utilisé à plusieurs reprises dans la démonstration du Brossamer. On prépare d'abord quelques définitions.

**Définition B.1.** Soit  $(\Omega, \mathcal{F}, P)$  un espace de probabilité, et  $T : \Omega \rightarrow \Omega$  une fonction mesurable. On dit que  $T$  préserve la mesure si, pour tout  $A$  dans  $\mathcal{F}$ , on a  $P(T^{-1}A) = P(A)$ . En d'autres termes,  $T$  préserve la mesure si et seulement si la mesure image et la mesure de départ coïncident. On dit qu'un événement  $A$  est invariant pour  $T$  si  $T^{-1}A = A$ , *p.s.*. On dit enfin que  $T$  est ergodique si tout événement  $T$ -invariant a pour probabilité 0 ou 1.

On montre facilement

**Proposition B.2.** Pour  $T : \Omega \rightarrow \Omega$  mesurable, l'ensemble  $\mathcal{I}$  des événements  $T$ -invariants est une sous-tribu de  $\mathcal{F}$ .

**Théorème ergodique (de Birkhoff) B.3.** On se donne un espace de probabilité  $(\Omega, \mathcal{F}, P)$ , et une transformation  $T$  préservant  $P$ . Soit  $X$  une v.a. dans  $L^1$ , alors *p.s.* et dans  $L^1$ ,

$$\frac{X + X \circ T + \dots + X \circ T^{n-1}}{n} \longrightarrow E[X \mid \mathcal{I}].$$

si de plus  $T$  est ergodique, alors

$$\frac{X + X \circ T + \dots + X \circ T^{n-1}}{n} \longrightarrow E[X].$$

Soit  $(X_n)$  une suite de v.a. sur  $\Omega$ . Alors ceci définit une v.a.  $X$  à valeur dans  $\Omega^{\mathbb{N}}$  muni de la tribu produit. Sur  $\Omega^{\mathbb{N}}$  on définit l'opérateur de décalage

$$\theta : (x_0, x_1, \dots) \longmapsto (x_1, x_2, \dots).$$

On dit que  $(X_n)$  est une suite stationnaire si  $\theta$  préserve la mesure  $PX^{-1}$  sur  $\Omega^{\mathbb{N}}$ . De façon équivalente,  $(X_n)$  est stationnaire si et seulement si  $\forall k \in \mathbb{N}$ ,

$$(X_0, X_1, \dots) = (X_k, X_{k+1}, \dots)$$

en loi.

Donnons un exemple de suite stationnaire. Soit  $X$  une v.a. sur  $\Omega$  et  $T$  une transformation préservant la mesure, on vérifie immédiatement que  $(X \circ T^n)$  définit une suite stationnaire. En fait cet exemple est un prototype: à partir de chaque suite stationnaire, on peut construire un espace de probabilité, une v.a. et une transformation préservant la mesure sur cet espace, tels que la suite définie ci-dessus coïncide avec notre suite stationnaire en loi. Plus précisément, il suffit de considérer l'espace  $\Omega^{\mathbb{N}}$  et la v.a.  $X = (X_n)$ , la transformation préservant la mesure est donnée par l'opérateur de décalage. On conclut donc que l'étude des suites stationnaires se ramène à celle des transformations qui préservent la mesure, et on va faire des translations pour des notions introduites au début de l'appendice aux suites stationnaires.

**Définition B.4.** On dit qu'un événement  $A$  est invariant par rapport à  $(X_n)$  s'il existe un  $B$  dans  $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$  tel que  $\forall n \in \mathbb{N}$ ,

$$A = \{\omega \mid (X_n, X_{n+1}, \dots) \in B\}.$$

Encore une fois, on vérifie que l'ensemble des  $A$  forme une sous-tribu de  $\mathcal{F}$ , qu'on note  $\mathcal{I}_X$ . On dit que  $(X_n)$  est ergodique si  $\mathcal{I}_X$  est grossière. En faisant une modification de langage du théorème ergodique de Birkhoff, on obtient

**Théorème ergodique, bis B.5.** Soit  $(X_n)$  une suite stationnaire avec  $X_1 \in L^1$ , alors *p.s.* et dans  $L^1$ , on a

$$\frac{X_1 + \dots + X_n}{n} \longrightarrow E[X_1 \mid \mathcal{I}_X].$$

si de plus  $(X_n)$  est ergodique, on a

$$\frac{X_1 + \dots + X_n}{n} \longrightarrow E[X_1].$$

Evidemment ce théorème entraîne la loi forte des grands nombres.

## References

- [1] Patrick BILLINGSLEY. *Convergence of probability measures*. Wiley Series in Probability and Mathematical Statistics, John Wiley, New York, 1968.
- [2] Gunnar A. BROSAMLER. *An almost everywhere central limit theorem*. *Math. Proc. Cambridge Philos. Soc.* 104(1988), no.3, 561-574.
- [3] Richard M. DUDLEY. *Real analysis and probability*. Cambridge University Press, 2002.
- [4] David FREEDMAN. *Brownian motion and diffusion*. Holden-Day, 1971.
- [5] Ulrich KRENGEL. *Ergodic theorems*. Walter de Gruyter, 1985.
- [6] Michael T. LACEY and Walter PHILIPP. *A note on the almost sure central limit theorem*. *Statist. Probab. Lett.* 9(1990), no. 3, 201-205.
- [7] Péter MAJOR. *The approximation of partial sums of independent RV's*. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 35 (1976), no. 3, 213-220.
- [8] Albert N. SHIRYAEV. *Probability*. Springer-Verlag, 1996.