N° d'ordre : 664

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE FLANDRES ARTOIS

pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

par

Paulo Eduardo de OLIVEIRA

INFINIE DIVISIBILITÉ, PRINCIPES D'INVARIANCE ET ESTIMATION DE NOYAUX DE TRANSITION EN THÉORIE DES MESURES ALÉATOIRES

Soutenue le 23 janvier 1991 devant la Commission d'Examen :

Président et Rapporteur :	J. DELPORTE, Université Catholique de Lille
Rapporteur :	N. MENDES LOPES, Université de Coimbra
Examinateurs :	M. HALLIN, Université Libre de Bruxelles
	P. JACOB, Université de Lille
	R. MOCHÉ, Université de Lille

Je tiens à exprimer ma respectueuse gratitude à Monsieur le Professeur Jean Delporte, qui a bien voulu me faire l'honneur de présider le Jury de cette thèse.

Je veux remercier Monsieur le Professeur Pierre Jacob qui m'a guidé dans cette initiation à la recherche. Ses suggestions, ses questions, ses critiques et sa patience sont à la base de la plupart des résultats obtenus. Sans son précieux support ce travail n'aurait pas été possible.

Je veux aussi exprimer ma gratitude à Madame le Professeur Maria Nazaré Mendes Lopes, qui m'a fait découvrir la théorie des probabilités et m'a guidé dans mes premières études en théorie des mesures aléatoires.

Je remercie Messieurs les Professeurs Marc Hallin et Raymond Moché qui me font l'honneur de participer au Jury.

Je remercie l'INIC - Instituto Nacional de Investigação Científica - dont j'ai été boursier.

Je remercie la JNICT - Junta Nacional para a Investigação Científica e Technológica.

Table des Matières

Infinie divisibilité, principes d'invariance et estimation de noyaux de transition en théorie de mesures aléatoires

Paulo Eduardo Oliveira

INTRODUCTION

Comme beaucoup d'autres problèmes de la théorie des probabilités l'étude des mesures aléatoires est née à partir d'applications pratiques de la mathématique. Evidemment, les mathématiciens en étudiant ces applications ont cherché à formaliser les situations réelles dans un cadre théorique, et on est arrivé aux mesures aléatoires comme formalisation satisfaisante. Les situations pratiques qui ont motivé cette étude apparaissent dans ce cadre comme des mesures aléatoires particulières: les processus ponctuels, où toutes les masses ont une valeur entière. De par la nature des problèmes envisagés on était naturellement amené à considérer les valeurs possibles comme non négatives. La généralisation aux mesures aléatoires suggèrait d'approfondir l'étude de topologies sur des espaces de mesures, pour permettre l'étude des convergences stochastiques des mesures aléatoires, par exemple. Des caractérisations des convergences de mesures non aléatoires, et d'autres notions topologiques liées, ont été obtenues dans les années 1940-1943, par Alexandroff [1], et plus tard, 1956, par Prokhorov [41]. L'étude des mesures aléatoires posait des problèmes intéressants du point de vue mathématique en ce qui concerne l'identification de mesures aléatoires, Prokhorov [42] pour les premiers résultats, la caractérisation des mesures aléatoires à partir des lois de probabilité des boréliens, Nawrotzki [33], Harris [17], [18], ou la caractérisation de la convergence de suites de mesures aléatoires, Matthes, Kerstan, Mecke [31] et d'autres auteurs. Evidemment, ces questions résolues, les mathématiciens pouvaient s'intéresser à d'autres problèmes concernant les mesures aléatoires. Bien sûr, beaucoup d'entre les problèmes étudiés n'ont pas été posés par des situations pratiques, mais ont été suggerés par la formalisation mathématique trouvée. Une référence classique pour la théorie des mesures aléatoires est le livre de Kallenberg [26].

Cet exposé prétend aller un peu plus loin dans ce désir de généralisation, en laissant les mesures prendre des valeurs réelles. Le fait de supposer les mesures à valeurs réelles pas forcément non négatives, s'il nous éloigne des situations pratiques, pose des questions intéressantes du point de vue mathématique. D'abord, l'espace de mesures sur lequel nous devrons travailler change de propriétés, notamment il n'est plus métrisable, ce qui, évidemment, pose des problèmes plus délicats de caractérisation de convergence et de compacité relative, pour en parler des notions topologiques qui nous intéressent le plus. Ce type de problèmes a déjà été l'objet d'étude des mathématiciens, parmi lesquels nous référons Varadarajan [47], Marle [30] et Tortrat [46]. Si nous nous intéressons au cas aléatoire les références sont plus rares, par exemple Jacob [21]. Ce changement des propriétés de l'espace de mesures pose des problèmes déjà en ce qui concerne la définition de mesure aléatoire, car l'introduction de la tribu convenable sur l'espace de mesures n'est plus évidente. Ce problème est étudié dans le chapitre 2. Une fois définie une tribu sur l'espace de mesures nous pouvons nous intéresser aux problèmes d'identification de mesures aléatoires, de convergence de mesures aléatoires, de caractérisation de mesures aléatoires à partir des masses des boréliens, en suivant le plan proposé par Kallenberg dans son livre [26]. Ce chapitre termine en faisant une incursion dans d'autres types de convergences de mesures aléatoires.

Dans le troisième chapitre nous proposons l'étude d'un problème classique de la théorie des probabilités: l'infinie divisibilité. Nous cherchons des caractérisations de quelques classes particulières de mesures aléatoires. Remarquons que l'étude du cas à accroissements indépendants suit de près l'étude des processus stochastiques à accroissements indépendants indexés par un intervalle de la droite, selon Gihman, Skorohod [12].

Les deux derniers chapitres concernent des applications des mesures aléatoires à d'autres situations, ce qui nous permet de faire des mathématiques appliquées. Dans un premier temps nous particularisons l'espace de base, ce qui nous permet de démontrer des propriétés de compacité relative pour des variables aléatoires à valeurs dans $L^2[0,1]$, et des versions faibles du théorème de Donsker. En ce qui concerne le dernier chapitre nous nous contentons de considerer les mesures aléatoires non négatives pour trouver un regard général sur diverses problèmes classiques d'estimation. Notre approche des problèmes d'estimation envisagés permet de clarifier quelques conditions de convergence, car elles nous apparaissent de façon plus naturelle. De plus, l'approche que nous utilisons nous permet d'étudier la convergence des processus empiriques associés aux estimateurs définis.

Chapitre 1

Espaces de mesures

Le but de ce premier chapitre est d'étudier la topologie vague sur l'espace des mesures de Radon à signe sur un espace métrique séparable, complet et localement compact. Ce chapitre servira de base pour l'étude du cas aléatoire, notamment pour les problèmes de convergence en loi de mesures aléatoires par rapport à la topologie vague. L'introduction de la topologie vague peut être faite de deux façons différentes: la façon classique qui consiste à utiliser directement l'ensemble des fonctions continues à support compact, ou, en utilisant une idée de Matthes, Kersten, Mecke [31], on peut introduire la topologie vague à partir de la topologie faible sur l'espace des mesures à signe bornées. Cette deuxième définition pour la topologie vague se révèlera très utile pour la caractérisation des ensembles vaguement relativement compacts. De plus, à partir de la deuxième définition de la topologie vague, nous généraliserons au cas des mesures signées quelques résultats de convergence connus pour le cas des mesures non négatives. Bien sûr, cette définition alternative suggère l'étude de la topologie faible sur l'espace des mesures à signe bornées. En ce qui nous concerne pour cette étude nous nous appuyerons sur Varadarajan [47] et développerons un peu les résultats qui peuvent nous intéresser le plus.

Dans tout ce travail (S,d) sera un espace métrique séparable, complet et localement compact (dans la plupart des cas la métrique d n'a aucune importance, le seul fait qui compte c'est que la topologie soit métrisable). Par $\mathcal{S}$ nous désignerons la tribu borélienne de S et par $\mathcal{B}$ l'anneau des boréliens relativement compacts. Considérons encore les espaces de Banach $\mathbf{C(S)}$ et $\mathrm{C}_c(\mathrm{S})$ des fonctions f: $\mathrm{S} \longrightarrow \mathbb{R}$ continues et bornées, continues à support compact, respectivement, avec la norme

$$||f|| = \sup_{s \in \mathbf{S}} |f(s)|.$$

Désignons encore par $\mathcal{M}_b$ l'espace des mesures boréliennes à signe bornées sur $\mathbf{S}$ ($\mu \in \mathcal{M}_b$ si et seulement si $\sup_{\mathrm{B} \in \mathcal{S}} |\mu(\mathrm{B})| < \infty$), et par $\mathcal{M}_c$ l'espace des mesures de Radon à signe sur $\mathbf{S}$. Une mesure de Borel à signe sur $\mathbf{S}$, μ, est une mesure de Radon à signe si et seulement si elle est finie pour tout élément $\mathrm{B} \in \mathcal{B}$. De plus, étant donné une mesure à signe (bornée ou de Radon), μ, sur $\mathbf{S}$ nous pouvons écrire cette mesure comme la différence de deux mesures non négatives de la même espèce que μ. Cela veut dire que si μ est bornée elle est la différence de deux mesures non négatives bornées et, si μ est de Radon elle est la différence de deux mesures de Radon non négatives. Ces deux mesures non négatives, en géneral, ne sont pas uniques. En posant

$$\mu^+(\mathrm{B}) = \sup\{\mu(\mathrm{C}), \mathrm{C} \subset \mathrm{B}\} \text{ et } \mu^-(\mathrm{B}) = -\inf\{\mu(\mathrm{C}), \mathrm{C} \subset \mathrm{B}\}$$

nous obtenons une décomposition de la mesure μ comme la différence de deux mesures non négatives concentrées sur des boréliens disjoints. Cette décomposition est appelée la décomposition de Hahn-Jordan. La mesure μ^+ sera appelée la variation positive de μ, μ^- la variation négative de μ, et $|\mu| = \mu^+ + \mu^-$ la variation totale de μ. Remarquons que les mesures μ^+ et μ^- sont une décomposition minimale de μ comme différence de deux mesures non négatives. Cela veut dire que si $\mu = \nu_1 - \nu_2$, où les mesures ν_1 et ν_2 sont non négatives alors $\mu^+ \leq \nu_1$ et $\mu^- \leq \nu_2$. A ce moment notons que d'après Dinculeanu [9], I, §3, prop. 7, si $|\mu(B)| < \infty$ alors aussi $|\mu|(B) < \infty$. L'implication réciproque étant évidente nous pouvons donc caractériser les mesures de Radon à signe comme les mesures boréliennes à signe qui sont à variation totale finie sur tous les boreliens relativement compacts. Aussi, nous pouvons caractériser les éléments de $\mathcal{M}_b$ comme les mesures boréliennes sur $\mathbf{S}$ telles que la variation totale soit une mesure non négative finie sur $\mathbf{S}$.

Finalement, pour toute mesure μ, appartenant à $\mathcal{M}_b$ ou à $\mathcal{M}_c$, et pour toute fonction f μ-intégrable on désignera l'intégrale de f par raport à μ par un des symboles suivants:

$$\mu f, \int f d\mu, \text{ou} \int f(s)\mu(ds).$$

1.1 La topologie faible sur $\mathcal{M}_b$

L'étude de la topologie faible dans l'espace des mesures à signe bornées est faite dans l'article classique de Varadarajan [47]. Ici nous citerons les résultats de Varadarajan qui nous seront utiles pour la poursuite de notre exposé et nous développerons quelques-uns d'entre eux pour leur donner une forme qui sera plus convenable pour nous.

1.1.1 Propriétés générales et définitions

Avant d'entreprendre l'étude de la topologie faible remarquons quelques propriétés de regularité des mesures à signe bornées. Pour cela définissons la notions de régularité pour les mesures à signe.

Définition 1.1.1 *1. Un borélien $C\in \mathcal{S}$ est régulier relativement à une mesure $\mu \in\mathcal{M}_b$ si pour tout $\varepsilon > 0$ il existe un compact $K\subset C$ et un ouvert $G\supset C$ tels que pour tout $C'\in \mathcal{S}$ tel que $K\subset C'\subset G$ on ait*

$$|\mu(C) - \mu(C')| < \varepsilon.$$

2. Une mesure $\mu \in\mathcal{M}_b$ est régulière si tous les boréliens sont réguliers relativement à μ.

3. Une mesure μ est tendue si, pour tout $\varepsilon > 0$ il existe un ensemble compact K, dans **S** *tel que $|\mu|(K^c)< \varepsilon$, où K^c représente le complémentaire de K dans* **S**.

Remarquons que si la mesure μ est non négative les définitions de mesure régulière et tendue coïncident avec les définitions classiques, qui sont présentées dans Billingsley [5] par exemple (en ce qui concerne la régularité voir Dinculeanu [9], III, §15, prop. 1).

Théorème 1.1.2 1. [Billingsley [5], th. 1.1] *Soit $\mu \in\mathcal{M}_b$, $C\in \mathcal{S}$, $\varepsilon > 0$. Il existe un ensemble F fermé et un ensemble G ouvert tels que $F\subset C\subset G$ et $|\mu|(G\setminus F) < \varepsilon$.*

2. [Dinculeanu [9], III, §15, prop. 24] *Toute mesure $\mu \in\mathcal{M}_b$ est régulière.*

3. [Billingsley [5], th. 1.2] *Toute mesure $\mu \in\mathcal{M}_b$ est tendue.*

Donc, d'après ce théorème les mesures à signe bornées vérifient les mêmes propriétés de régularité que les mesures non négatives bornées. On remarquera que le fait que les mesures à signe ont les mêmes propriétés que les mesures non négatives se produit quand on s'intéresse à l'étude de certains problèmes concernant la topologie faible sur $\mathcal{M}_b$ que nous introduisons tout de suite.

Définition 1.1.3 *Une suite $\{\mu_n\}$ de mesures à signe bornées dans* **S** *converge faiblement vers $\mu \in \mathcal{M}_b$ si pour toute fonction $f \in$C(S) se vérifie $\mu_n f \longrightarrow \mu f$. On écrira $\mu_n \xrightarrow{w} \mu$ pour traduire le fait que μ_n converge faiblement vers μ.*

Remarquons que cette définition correspond à la définition de la topologie faible étoile, selon la terminologie de l'analyse fonctionnelle, sauf que dans ce cas l'espace $\mathcal{M}_b$ n'est pas le dual de C(S).

Evidemment, pour la topologie faible une base de voisinages d'une mesure μ est la famille des ensembles de la forme

$$\mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon) = \{\nu \in \mathcal{M}_b : |\mu f_i - \nu f_i| < \varepsilon, i = 1, \ldots, n\}$$

où $\varepsilon > 0, f_1, \ldots, f_n \in$C(S), $n \in \mathbb{N}$.

Au vu de cette définition il est évident qu'il suffit de définir une base de voisinages de la mesure nulle, les bases de voisinages d'une mesure quelconque s'obtenant alors par translation associée à la mesure donnée. Ce fait nous sera utile par exemple quand on voudra établir la continuité d'applications linéaires car il suffira d'établir la continuité à l'origine de l'espace (la mesure nulle).

Considérons une fonction $f \in$C(S) et définissons l'application

$$\begin{array}{rcl} \Pi_f : \mathcal{M}_b & \longrightarrow & \mathbb{R} \\ \mu & \hookrightarrow & \mu f. \end{array}$$

Les applications Π_f, pour $f \in$C(S), jouent évidemment un rôle important dans la topologie faible. Il est facile de vérifier que Π_f, pour chaque fonction $f \in$C(S), est faiblement continue. De plus, on pourrait même définir la topologie faible dans $\mathcal{M}_b$ comme la topologie engendrée par la famille d'applications $\{\Pi_f,\ f \in$C(S)$\}$.

Théorème 1.1.4 (Varadarajan [47], II.1, th. 1, page 181) *L'espace de mesures $\mathcal{M}_b$ avec la topologie faible est un espace vectoriel topologique de Hausdorff complètement régulier.*

<u>**Démonstration**</u> : Le produit d'une mesure par un réel est évidemment une application continue car, si a$\in$IR, $(a\mu)f = a(\mu f)$. Etablissons donc la continuité de l'application $s(\mu, \nu) = \mu + \nu$. Par des raisons déjà expliquées il suffit de vérifier la continuité à l'origine $(0,0)$. Un élément de la base de voisinages de $s(0,0) = 0$ est de la forme $\{\mu : |\mu f_i| < \varepsilon, i = 1, \ldots, n\}$, pour $\varepsilon > 0, n \in \mathbb{N}, f_1, \ldots, f_n \in \mathrm{C(S)}$. Les ensembles $\mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon/2)$ sont des voisinages faibles de la mesure nulle dans $\mathcal{M}_b$, donc le produit cartésien

$$\mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon/2) \times \mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon/2)$$

est un voisinage faible de $(0,0)$ dans $\mathcal{M}_b \times \mathcal{M}_b$ et

$$s(\mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon/2) \times \mathrm{W}(\mu, f_1, \ldots, f_n, \varepsilon/2)) \subset \{\mu : |\mu f_i| < \varepsilon, i = 1, \ldots, n\}.$$

L'espace est évidemment de Hausdorff car si $\mu \neq \nu$ il existe une fonction $f \in \mathrm{C(S)}$ telle que $\mu f \neq \nu f$.

Démontrons finalement que l'espace est complètement régulier. Soit μ_0 une mesure bornée à signe et considérons un voisinage de μ_0:

$$\mathrm{W}(\mu_0, f_1, \ldots, f_n, \varepsilon).$$

Définissons les fonctions

$$h_i(\mu) = \frac{1}{\varepsilon} \min \left(|\int f_i d\mu - \int f_i d\mu_0|, \varepsilon \right).$$

Les fonctions h_i sont continues et à valeurs dans [0,1]. Evidemment $h_i(\mu_0) = 0$ et h_i prend la valeur 1 dans $(\mathrm{W}(\mu_0, f_i, \varepsilon))^c$. En posant

$$h = \max\ (h_1, \ldots, h_n)$$

nous obtenons une fonction continue telle que $h(\mu_0) = 0$ et h prend la valeur 1 dans $(\mathrm{W}(\mu_0, f_1, \ldots, f_n, \varepsilon))^c$, ce qui démontre le théorème. ∎

Les résultats qui nous intéressent le plus sont les caractérisations des ensembles relativement compacts et de la convergence faible de suites de mesures.

1.1.2 Compacité relative faible

Les premières caractérisations de compacité relative sont déduites du fait que la topologie faible est définie comme une topologie faible au sens de l'analyse fonctionnelle.

Théorème 1.1.5 *Soit* $M \subset \mathcal{M}_b$ *faiblement relativement compact. Alors les conditions suivantes sont vérifiées*

1. *pour toute fonction* $f \in \mathrm{C}(\mathbf{S})$, $\sup_{\mu \in \mathrm{M}} |\int_{\mathbf{S}} f d\mu| < \infty$;
2. $\sup_{\mu \in \mathrm{M}} |\mu|(\mathbf{S}) < \infty$.

Ce théorème est une adaptation d'un théorème de Varadarajan [47]. Dans son théorème, Varadarajan démontre encore que les conditions 1 et 2 impliquent la compacité relative de l'ensemble M. Mais, pour obtenir ce résultat il faut que l'espace des mesures soit l'espace dual de $\mathrm{C}(\mathbf{S})$, ce qui se passe si nous supposons que l'espace métrique **S** est, en plus des propriétés déjà supposées, compact. Dans ce cas nous pouvons énoncer le théorème suivant, qui est en fait une traduction pour notre cadre du théorème de Banach-Steinhaus.

Théorème 1.1.6 (Varadarajan [47], II.7, th. 2.1, page 200) *Supposons que l'espace métrique* S *soit compact et soit* $M \subset \mathcal{M}_b$. *Les conditions suivantes sont équivalentes.*

1. *M est faiblement relativement compact;*
2. *pour toute fonction* $f \in \mathrm{C}(\mathbf{S})$, $\sup_{\mu \in \mathrm{M}} |\int_{\mathbf{S}} f d\mu| < \infty$;
3. $\sup_{\mu \in \mathrm{M}} |\mu|(\mathbf{S}) < \infty$.

Désignons par $\mathcal{M}_b^+$ le sous-espace de $\mathcal{M}_b$ des éléments non négatifs. La topologie trace dans $\mathcal{M}_b^+$ est égale à la topologie faible classiquement définie sur cet espace (voir Billingsley [5], ch. 1, ou Marle [30], 9.9.1). Il est facile de vérifier que $\mathcal{M}_b^+$ est fermé pour la topologie faible de $\mathcal{M}_b$. Dans l'espace $\mathcal{M}_b^+$ nous disposons d'un critère très utile pour établir la compacité relative d'un ensemble de mesures: le critère de Prokhorov. Nous allons récupérer ce critère dans l'espace $\mathcal{M}_b$. D'abord remarquons qu'il y a une relation étroite entre la compacité relative d'un ensemble M dans la topologie faible de $\mathcal{M}_b$ et la compacité relative d'ensembles liés à M dans la topologie faible de $\mathcal{M}_b^+$.

Théorème 1.1.7 (Varadarajan [47], II.7, th. 26, page 201) *Soit M un sous-ensemble de $\mathcal{M}_b$. Les conditions suivantes sont équivalentes.*

1. *M est faiblement relativement compact dans $\mathcal{M}_b$;*
2. *M est faiblement relativement sequentiellement compact dans $\mathcal{M}_b$;*
3. *$|M| = \{|\mu| : \mu \in M\}$ est faiblement relativement compact dans $\mathcal{M}_b^+$;*
4. *$M^+ = \{\mu^+ : \mu \in M\}$ et $M^- = \{\mu^- : \mu \in M\}$ sont faiblement relativement compacts dans $\mathcal{M}_b^+$.*

Remarquons que Varadarajan a démontré que les équivalences 1$\Leftrightarrow$3$\Leftrightarrow$4 restent vraies même dans le cas où l'espace $\mathbf{S}$ est seulement un espace topologique géneral (Varadarajan [47], II.7, th. 28, page 203).

Avant d'énoncer la version du critère de Prokhorov pour les mesures à signe définissons la notion d'ensemble de mesures équitendu.

Définition 1.1.8 *Un ensemble $M \subset \mathcal{M}_b$ est équitendu si*

$$\forall \varepsilon > 0\ \exists K \text{ compact de } \mathbf{S} : sup_{\mu \in \mathrm{M}} |\mu|(K^c) < \varepsilon,$$

et

$$\sup_{\mu \in \mathrm{M}} |\mu|(\mathbf{S}) < \infty.$$

Il est évident que si $\mathrm{M} \subset \mathcal{M}_b^+$ cette définition coïncide avec la définition classique d'équitension. Le critère de Prokhorov dit alors qu'un ensemble $\mathrm{M} \subset \mathcal{M}_b^+$ est faiblement relativement compact si et seulement si l'ensemble M est équitendu, car l'espace $\mathbf{S}$ est métrique séparable et complet (Billingsley [5], ths. 6.1 et 6.2). Alors comme corollaire du théorème 1.1.7 nous avons

Corollaire 1.1.9 (critère de Prokhorov) *Un ensemble $M \subset \mathcal{M}_b$ est relativement compact si et seulement s'il est équitendu.*

1.1.3 Convergence faible de suites de mesures dans $\mathcal{M}_b$

Nous commençons l'étude de conditions de convergence faible dans $\mathcal{M}_b$ par le rappel d'un théorème fondamental dans l'étude de la convergence faible dans $\mathcal{M}_b^+$: le théorème du portmanteau.

Théorème 1.1.10 (Alexandroff [1]) *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b^+$ et μ un élément de $\mathcal{M}_b^+$. Les conditions suivantes sont équivalentes.*

1. $\mu_n \xrightarrow{w} \mu$ *dans* $\mathcal{M}_b^+$;
2. $\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$, *pour tout* $F \subset \mathbf{S}$ *fermé, et* $\mu_n(\mathbf{S}) \longrightarrow \mu(\mathbf{S})$;
3. $\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$, *pour tout* $G \subset \mathbf{S}$ *ouvert, et* $\mu_n(\mathbf{S}) \longrightarrow \mu(\mathbf{S})$;
4. $\lim_{n\to\infty} \mu_n(A) = \mu(A)$, *pour tout ensemble* $A \in \mathcal{S}$ *tel que* $\mu(fr(A)) = 0$, *où* $fr(A)$ *désigne la frontière de l'ensemble* A.

Ce théorème nous donne des conditions très générales pour établir la convergence faible dans $\mathcal{M}_b^+$. Par la suite de cet exposé on remarquera qu'il ne sera pas possible de récupérer une version de ce théorème dans $\mathcal{M}_b$. En effet, on verra que, si les conditions de relative compacité faible étaient facilement récupérables dans $\mathcal{M}_b$, quand on veut identifier la mesure limite on commence à trouver des difficultés qui en général nous empêchent de trouver des conditions suffisantes utilisables. On remarquera encore que l'identification de la mesure limite n'est obtenue que d'une façon très indirecte.

Au cours de cette section on remarquera l'importance des variations positive, négative et surtout de la variation totale dans les conditions à obtenir. En fait, la variation totale avait déjà un rôle important dans les caractérisations de la compacité relative faible. Tout de suite nous pouvons dire, d'après le théorème 1.1.7, que la compacité relative faible de la suite des variations totales est une condition nécessaire pour la convergence faible d'une suite de mesures. Mais il est possible de dire plus.

Théorème 1.1.11 (Varadarajan [47], II.2, th. 3, page 183) *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b$ et $\mu \in \mathcal{M}_b$. Supposons que $\mu_n \xrightarrow{w} \mu$. Alors*

$$\liminf_{n\longrightarrow\infty} |\mu_n|(G) \geq |\mu|(G),$$

pour tout $G \subset \mathbf{S}$ ouvert.

Ce théorème nous donne presque la condition 3 du théorème 1.1.10. Cela veut dire que la suite des variations totales $\{|\mu_n|\}$ est "presque" faiblement convergente vers la variation totale de la limite faible de la suite initiale $\{\mu_n\}$. En effet, il suffirait de démontrer que $|\mu_n|(\mathbf{S}) \longrightarrow |\mu|(\mathbf{S})$ pour avoir la convergence faible $|\mu_n| \xrightarrow{w} |\mu|$. Pourtant, il est facile de vérifier que la convergence $|\mu_n|(\mathrm{S}) \longrightarrow |\mu|(\mathrm{S})$ en général n'est pas vraie. Pour ceci il suffit de prendre $\mu_n = \delta_{\frac{1}{n}} - \delta_{-\frac{1}{n}}$ dans $\mathbb{R}$, où δ_x répresente la mesure de Dirac avec une masse au point x. Evidemment $\mu_n \xrightarrow{w} 0$ (mesure nulle) et $\mu_n^+ = \delta_{\frac{1}{n}} \xrightarrow{w} \delta_0$, $\mu_n^- = \delta_{-\frac{1}{n}} \xrightarrow{w} \delta_0$, donc $|\mu_n| = \mu_n^+ + \mu_n^- \xrightarrow{w} 2\delta_0$.

Du fait que $\mu f = \mu^+ f - \mu^- f$ et $|\mu| f = \mu^+ f + \mu^- f$, pour toute fonction $f \in \mathrm{C}(\mathrm{S})$, il est évident que si pour une suite de mesures $\{\mu_n\}$ on a la convergence faible des suites des variations positives et négatives on aura aussi la convergence faible de la suite $\{\mu_n\}$ et de la suite des variations totales $\{|\mu_n|\}$. De plus, si $\mu_n^+ \xrightarrow{w} \nu_1$ et $\mu_n^- \xrightarrow{w} \nu_2$, où ν_1, $\nu_2 \in \mathcal{M}_b^+$ et ne sont pas forcément concentrées dans des boréliens disjoints, alors $\mu_n \xrightarrow{w} \nu_1 - \nu_2$ et $|\mu_n| \xrightarrow{w} \nu_1 + \nu_2$. Dans le cas où nous avons la convergence faible simultanée de la suite de mesures et de la suite des variations totales nous pouvons retrouver une adaptation du théorème 1.1.10. Démontrons d'abord un résultat auxiliaire.

Théorème 1.1.12 *Soit $\mu \in \mathcal{M}_b$ telle que $\mu(A) \geq 0$ pour tout A tel que $\mu(fr(A)) = 0$. Alors μ est une mesure non négative.*

<u>**Démonstration**</u> : Soit $\mathrm{F} \subset \mathrm{S}$ fermé et $\{\varepsilon_n\}$ une suite décroissante vers zéro. Posons $\mathrm{F}^{\varepsilon_n} = \{x \in \mathrm{S} : d(x, \mathrm{F}) \leq \varepsilon_n\}$. Comme $\mu \in \mathcal{M}_b$ nous pouvons choisir la suite $\{\varepsilon_n\}$ telle que

$$\mu^+ \{x \in \mathrm{S} : d(x, \mathrm{F}) = \varepsilon_n\} = \mu^- \{x \in \mathrm{S} : d(x, \mathrm{F}) = \varepsilon_n\} = 0,$$

puisque la mesure μ a au plus une infinité dénombrable d'atomes. Comme $\mu^+(\mathrm{F}^{\varepsilon_n})$ est décroissante vers $\mu^+(\mathrm{F})$ et $\mu^-(\mathrm{F}^{\varepsilon_n})$ est décroissante vers $\mu^-(\mathrm{F})$ il s'ensuit

$$\mu(\mathrm{F}^{\varepsilon_n}) \longrightarrow \mu(\mathrm{F}).$$

Du choix de la suite $\{\varepsilon_n\}$ et du fait $\mathrm{fr}(\mathrm{F}^{\varepsilon_n}) \subset \{x \in \mathrm{S} : d(x, \mathrm{F}) = \varepsilon_n\}$ nous avons $\mu^+(\mathrm{fr}(\mathrm{F}^{\varepsilon_n})) = \mu^-(\mathrm{fr}(\mathrm{F}^{\varepsilon_n})) = 0$, donc $\mu(\mathrm{fr}(\mathrm{F}^{\varepsilon_n})) = 0$. Alors, par hypothèse $\mu(\mathrm{F}^{\varepsilon_n}) \geq 0$, d'où il en résulte $\mu(\mathrm{F}) \geq 0$.

Soit $\mathrm{B} \in \mathcal{S}$. Alors

$$\mu^+(\mathrm{B}) = \sup \left\{\mu^+(\mathrm{F}), \mathrm{F} \subset \mathrm{B} \text{ fermé}\right\}$$

$$\mu^-(\mathrm{B}) = \sup \left\{\mu^-(\mathrm{F}), \mathrm{F} \subset \mathrm{B} \text{ fermé}\right\}.$$

Donc il existe deux suite de fermés $\{\mathrm{F}_n\}$ et $\{\mathrm{F}_n^*\}$ telles que

$$\mu^+(\mathrm{F}_n) > \mu^+(\mathrm{B}) - \frac{1}{n}$$

$$\mu^-(F_n^*) > \mu^-(B) - \frac{1}{n}$$

Posons $F_n^{**} = F_n \cup F_n^*$. Les ensembles F_n^{**} sont fermés, inclus dans B et

$$\mu^+(F_n^{**}) > \mu^+(B) - \frac{1}{n}$$

$$\mu^-(F_n^{**}) > \mu^-(B) - \frac{1}{n}$$

d'où il s'ensuit

$$\mu(F_n^{**}) = \mu^+(F_n^{**}) - \mu^-(F_n^{**}) \longrightarrow \mu^+(B) - \mu^-(B) = \mu(B),$$

et donc $\mu(B) \geq 0$. ■

Théorème 1.1.13 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b$, $\mu \in \mathcal{M}_b$ et $\nu \in \mathcal{M}_b^+$.*

1. *Si, pour tout ensemble $A \in \mathcal{S}$ tel que $\nu(fr(A)) = |\mu(fr(A))|$, se vérifient*
   (a) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$;
   (b) $\lim_{n\to\infty} |\mu_n|(A) = \nu(A)$,
   alors $\mu_n \xrightarrow{w} \mu$ et $|\mu_n| \xrightarrow{w} \nu$.
2. *Réciproquement, si $\mu_n \xrightarrow{w} \mu$ et $|\mu_n| \xrightarrow{w} \nu$ alors, pour tout ensemble $A \in \mathcal{S}$ tel que $\nu(fr(A)) = |\mu(fr(A))| = 0$*
   (a) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$;
   (b) $\lim_{n\to\infty} |\mu_n|(A) = \nu(A)$.

<u>**Démonstration**</u> : 1. D'après les hypothèses, remarquant que $\mu_n^+ = \frac{1}{2}(|\mu_n| + \mu_n)$ et $\mu_n^- = \frac{1}{2}(|\mu_n| - \mu_n)$, pour tout ensemble $A \in \mathcal{S}$ tel que $\nu(\mathrm{fr}(A)) = |\mu(\mathrm{fr}(A))|$, nous déduisons

$$\lim_{n\to\infty} \mu_n^+(A) = \frac{1}{2}(\nu(A) + \mu(A))$$

et

$$\lim_{n\to\infty} \mu_n^-(A) = \frac{1}{2}(\nu(A) - \mu(A)).$$

Si $\nu(\mathrm{fr}(A))$ - $\mu(\mathrm{fr}(A)) = 0$ ou $\nu(\mathrm{fr}(A)) + \mu(\mathrm{fr}(A)) = 0$ alors $\nu(\mathrm{fr}(A)) = |\mu(\mathrm{fr}(A))|$. Les mesures $\nu + \mu$ et $\nu - \mu$ sont non négatives d'après le théorème

1.1.12. Donc, d'après la condition 4 du théorème 1.1.10 on a $\mu_n^+ \xrightarrow{w} \frac{1}{2}(\nu + \mu)$ et $\mu_n^- \xrightarrow{w} \frac{1}{2}(\nu - \mu)$. D'après le commentaire précédent le théorème 1.1.12 il s'ensuit $\mu_n \xrightarrow{w} \mu$ et $|\mu_n| \xrightarrow{w} \nu$.

2. Dans ce cas on déduit aisément les convergences $\mu_n^+ \xrightarrow{w} \frac{1}{2}(\nu + \mu)$ et $\mu_n^- \xrightarrow{w} \frac{1}{2}(\nu - \mu)$. Donc, d'après la condition 4 du théorème 1.1.10, pour les ensembles A$\in \mathcal{S}$ tels que $\nu(\text{fr}(A)) = \mu(\text{fr}(A)) = 0$ il s'ensuit $\mu_n^+(A) \xrightarrow{w} \frac{1}{2}(\nu(A) + \mu(A))$ et $\mu_n^-(A) \xrightarrow{w} \frac{1}{2}(\nu(A) - \mu(A))$, d'où l'on déduit les convergences annoncées. ■

Il faut remarquer à nouveau que le fait d'avoir la convergence faible d'une suite de mesures n'implique pas la convergence faible de la suite des variations totales comme on peut le vérifier par l'exemple

$$\mu_n = \begin{cases} 0 & \text{n pair} \\ \delta_{\frac{1}{n}} - \delta_{-\frac{1}{n}} & \text{n impair} \end{cases} \qquad |\mu_n| = \begin{cases} 0 & \text{n pair} \\ \delta_{\frac{1}{n}} + \delta_{-\frac{1}{n}} & \text{n impair.} \end{cases}$$

Evidemment, $\mu_n \xrightarrow{w} 0$ et pour la suite des variations totales on vérifie que $|\mu_{2n}| \xrightarrow{w} 0$, $|\mu_{2n+1}| \xrightarrow{w} 2\delta_0$, donc la suite $\{\{|\mu_n|\}\}$ n'est pas faiblement convergente. Néanmoins, d'après le théorème 1.1.7 si une suite $\{\mu_n\}$ est faiblement convergente alors l'ensemble des variations totales $\{|\mu_n|\}$ est faiblement relativement compact. Donc, il existe au moins une sous-suite $\{|\mu_{n_k}|\}$ faiblement convergente dans $\mathcal{M}_b^+$. Encore d'après le théorème 1.1.7 on déduit que les ensembles $\{\mu_n^+\}$ et $\{\mu_n^-\}$ sont faiblement relativement compacts dans $\mathcal{M}_b^+$, donc on peut trouver des sous-suites faiblement convergentes aussi dans ce cas. Evidemment il est possible de trouver une suite croissante d'indices $\{n_k\}$ telle que les deux sous-suites $\{\mu_{n_k}^+\}$ et $\{\mu_{n_k}^-\}$ soient faiblement convergentes. Dans ce cas il existe des relations simples entre les diverses limites obtenues, qui peuvent être résumées de la façon suivante.

Théorème 1.1.14 *Considérons une suite $\{\mu_n\}$ dans $\mathcal{M}_b$ et $\mu \in \mathcal{M}_b$ telles que $\mu_n \xrightarrow{w} \mu$. Soit $\{n_k\}$ une suite croissante de nombres entiers tels que les suites $\{\mu_{n_k}^+\}$ et $\{\mu_{n_k}^-\}$ soient faiblement convergentes. Posons*

$$\nu_1 = \lim \mu_{n_k}^+, \ \nu_2 = \lim \mu_{n_k}^-.$$

Alors $\mu = \nu_1 - \nu_2$ et $|\mu_{n_k}| \xrightarrow{w} \nu_1 + \nu_2$.

Donc, si une suite $\{\mu_n\}$ est faiblement convergente vers une mesure μ, l'ensemble des points d'accumulation de $\{|\mu_n|\}$ dans la topologie faible est contenu dans $\{\nu_1 + \nu_2 : \mu = \nu_1 - \nu_2\}$. Pourtant, la mesure $|\mu|$ n'est pas forcément un point d'accumulation de l'ensemble $\{|\mu_n|\}$, mais nous disposons d'un critère d'identification.

Corollaire 1.1.15 *Supposons vérifiées les conditions du théorème précédent. S'il existe une sous-suite $\{|\mu_{n_k}|\}$ telle que $|\mu_{n_k}|(\mathbf{S}) \longrightarrow |\mu|(\mathbf{S})$ alors $|\mu|$ est point d'accumulation faible de l'ensemble $\{|\mu_n|\}$.*

Démonstration : Comme la suite $\{\mu_n\}$ est convergente il s'ensuit qu'il existe une sous-suite $\{|\mu_{n_{k_j}}|\}$, de la suite $\{|\mu_{n_k}|\}$, faiblement convergente. Evidemment, pour cette sous-suite $|\mu_{n_{k_j}}|(\mathbf{S}) \longrightarrow |\mu|(\mathbf{S})$. D'après le théorème 1.1.11 nous déduisons que cette sous-suite vérifie $\liminf |\mu_{n_{k_j}}|(\mathrm{G}) \geq |\mu|(\mathrm{G})$, pour tout $\mathrm{G} \subset \mathbf{S}$ ouvert, donc d'après la condition 3 du théorème 1.1.10, il s'ensuit $|\mu_{n_{k_j}}| \xrightarrow{w} |\mu|$, donc $|\mu|$ est point d'accumulation faible de l'ensemble $\{|\mu_n|\}$. ■

Nous pouvons caractériser d'une façon un peu plus précise les mesures qui sont des points d'accumulation de la suite des variations totales (toujours en supposant que la suite est convergente).

Théorème 1.1.16 *Soit $\{\mu_n\}$ une suite de mesures faiblement convergente vers une mesure $\mu \in \mathcal{M}_b$. Tout point d'accumulation ν de l'ensemble $\{|\mu_n|\}$ vérifie $|\mu| \leq \nu$.*

Démonstration : Soit ν un point d'accumulation de l'ensemble $\{|\mu_n|\}$. Alors il existe une sous-suite $\{|\mu_{n_k}|\}$ faiblement convergente vers ν. Il s'ensuit que les suites $\{\mu^+_{n_k}\}$ et $\{\mu^-_{n_k}\}$ sont faiblement convergentes. Soient $\nu_1 = \lim \mu^+_{n_k}$, et $\nu_2 = \lim \mu^-_{n_k}$. Nous avons déjà remarqué, théorème 1.1.14, que ces mesures vérifient $\mu = \nu_1 - \nu_2$ et $\nu = \nu_1 + \nu_2$. Evidemment les mesures ν_1 et ν_2, étant des limites faibles de suites de mesures non négatives, sont des mesures non négatives. Alors, d'après les propriétés minimales de la décomposition de Hahn-Jordan (page 2) les mesures ν_1 et ν_2 vérifient $\mu^+ \leq \nu_1$ et $\mu^- \leq \nu_2$, donc $|\mu| = \mu^+ + \mu^- \leq \nu_1 + \nu_2 = \nu$, ce qui démontre le théorème. ■

Nous avons déjà démontré des propriétés nécessaires et des propriétés suffisantes de convergence faible en utilisant la convergence des masses de certains boréliens, théorème 1.1.13. Dans ce cas on étudiait la convergence faible simultanée de la suite de mesures et de la suite des variations totales. Nous pouvons démontrer une condition nécessaire avec la convergence des masses de certains boréliens plus générale. Dans cette caractérisation l'ensemble des points d'accumulation a un rôle central.

Théorème 1.1.17 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b$ faiblement convergente vers $\mu \in \mathcal{M}_b$. Si $A \in S$ est tel que, pour tout point d'accumulation ν de l'ensemble $\{|\mu_n|\}$, se vérifie $\nu(fr(A)) = 0$, alors $\lim_{n\to\infty} \mu_n(A) = \mu(A)$.*

<u>**Démonstration**</u> : Soit ν un point d'accumulation de l'ensemble $\{|\mu_n|\}$ et soit $\{|\mu_{n_k}|\}$ une sous-suite faiblement convergente vers ν. Alors les sous-suites $\{\mu_{n_k}^+\}$ et $\{\mu_{n_k}^-\}$ sont aussi faiblement convergentes. Soient ν_1 et ν_2 les mesures limites, respectivement. Comme $\nu = \nu_1 + \nu_2$ et les mesures sont toutes non négatives si $\nu(\mathrm{fr(A)}) = 0$ alors $\nu_1(\mathrm{fr(A)}) = \nu_2(\mathrm{fr(A)}) = 0$, donc $\lim \mu_{n_k}^+(\mathrm{A}) = \nu_1(\mathrm{A})$ et $\lim \mu_{n_k}^-(\mathrm{A}) = \nu_2(\mathrm{A})$, d'où on déduit $\lim \mu_{n_k}(\mathrm{A}) = \mu(\mathrm{A})$, parce que $\mu = \nu_1 + \nu_2$. ■

Dans le cas général il ne semble pas possible de trouver une condition suffisante de convergence faible en utilisant la convergence des masses de certains boréliens. Remarquons que même pour cette condition nécessaire les ensembles pour lesquelles nous pouvons assurer la convergence des masses doivent vérifier une condition qui peut être forte car l'ensemble des points d'accumulation peut être grand, donc la frontière de l'ensemble doit avoir masse nulle par rapport à beaucoup de mesures.

Un résultat qui nous donne une condition suffisante en utilisant des conditions sur les masses de certains boréliens est le théorème suivant, du à Alexandroff (voir [10], page 316). Remarquons que le théorème n'est pas géneral car il faut imposer que la suite soit uniformement bornée, ce qui n'est pas une restriction très sérieuse d'après le théorème 1.1.5.

Théorème 1.1.18 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b$ telle que*

$$\sup_{n\in\mathbf{N}} |\mu_n|(\mathbf{S}) < \infty.$$

Soit $\mu \in \mathcal{M}_b$ et supposons que, pour tout ouvert $G \subset \mathbf{S}$ tel que $\mu(\overline{G}) = \mu(G)$ (où $\overline{G}$ représente l'adhérence de G), se vérifie la convergence $\mu_n(G) \longrightarrow \mu(G)$. Alors $\mu_n \xrightarrow{w} \mu$.

1.1.4 Convergence faible de mesures induites

Cette section n'est pas essentielle à la poursuite de notre exposé. Elle est présentée ici pour compléter l'étude de la topologie faible car on utilisera les mêmes méthodes que dans l'exposé précédent.

Considérons $\mathbf{S}'$ un autre espace métrique séparable, complet et localement compact, et soit $h:\mathbf{S}\longrightarrow\mathbf{S}'$ une fonction mesurable. Alors, étant donné une mesure μ sur S, on note μh^{-1} la mesure induite par h dans l'espace $\mathbf{S}'$. Nous nous intéressons à l'étude des fonctions h pour lesquelles la convergence faible de mesures est préservée. Evidemment, si h est une fonction continue cette fonction préserve la convergence faible. Comme dans le cas des mesures non négatives nous pouvons affaiblir l'hypothèse de continuité sur h. Comme dans le théorème 1.1.17 l'ensemble des points d'accumulation de la suite des variations totales continue à avoir un rôle important. Désignons par D_h l'ensemble des points de discontinuité de la fonction h.

Théorème 1.1.19 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b(\mathbf{S})$ faiblement convergente vers $\mu \in\mathcal{M}_b(\mathbf{S})$. Si, pour tout point d'accumulation ν de l'ensemble $\{|\mu_n|\}$, se vérifie $\nu(D_h)= 0$, alors $\mu_n h^{-1}\xrightarrow{w}\mu h^{-1}$ dans $\mathcal{M}_b(\mathbf{S}')$.*

__Démonstration__ : Soit ν un point d'accumulation de $\{|\mu_n|\}$ et $\{n_k\}$ la suite d'indices tels que $|\mu_{n_k}|\xrightarrow{w}\nu$. Soient encore $\nu_1 = \lim\mu^+_{n_k}$ et $\nu_2 = \lim\mu^-_{n_k}$. Du fait que $\nu = \nu_1 + \nu_2$ et que les mesures sont non négatives il s'ensuit que $\nu_1(\mathrm{D}_h) = \nu_2(\mathrm{D}_h) = 0$. Donc, d'après Billingsley [5], th. 5.1, nous déduisons $\mu^+_{n_k}h^{-1}\xrightarrow{w}\nu_1 h^{-1}$ et $\mu^-_{n_k}h^{-1}\xrightarrow{w}\nu_2 h^{-1}$, ce qui implique $\mu_{n_k}h^{-1}\xrightarrow{w}\nu h^{-1}$.

Donc, de toute sous-suite $\{\mu_{n_k}h^{-1}\}$ nous pouvons extraire une sous-suite $\{\mu_{n_{k_j}}h^{-1}\}$ faiblement convergente vers μh^{-1}, ce qui démontre le théorème. ■

De la même façon, s'appuyant cette fois sur le théorème 5.5 de Billingsley [5], on démontre le théorème suivant qui généralise l'enoncé precedent. Considérons $h_n:\mathbf{S}\longrightarrow\mathbf{S}'$ des fonctions mesurables, pour chaque $n \in\mathbb{N}$, et définissons l'ensemble

$$\mathrm{E} = \{\mathrm{s}\in\mathbf{S}|\ \exists\{\mathrm{s}_n\}:\mathrm{s}_n \longrightarrow \mathrm{s} \text{ et } h_n(\mathrm{s}_n)\not\longrightarrow h(\mathrm{s})\}.$$

Théorème 1.1.20 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_b(\mathbf{S})$ faiblement convergente vers $\mu \in\mathcal{M}_b(\mathbf{S})$. Si, pour tout point d'accumulation ν de $\{|\mu_n|\}$, se vérifie $\nu(E) = 0$, alors $\mu_n h_n^{-1}\xrightarrow{w}\mu h^{-1}$ dans $\mathcal{M}_b(\mathbf{S}')$.*

Remarquons que en ce qui concerne les résultats présentés ici il suffit de supposer que S est un espace métrique géneral, sauf dans le cas du critère de Prokhorov, Corollaire 1.1.9, où il faut supposer que S est séparable et complet, et dans le cas du théorème précédent où il faut supposer que l'espace S est séparable.

1.2 La topologie vague dans $\mathcal{M}_c$

Nous avons déjà dit que la topologie vague dans $\mathcal{M}_c$ peut être introduite de deux façons. Nous utiliserons la définition classique, qui est plus générale, en utilisant les fonctions continues à support compact, pour déduire les premières caractérisations des ensembles vaguement relativement compacts. Pour ceci nous suivrons Marle [30], ch. 9.8. Avec la seconde définition de la topologie vague nous allons démontrer des liens avec la topologie faible dans $\mathcal{M}_b$, développer un peu les caractérisations des ensembles vaguement relativement compacts, et enfin établir des liens entre la convergence des masses de certains boréliens et la convergence vague d'une suite de mesures de Radon, comme dans le théorème 1.1.17.

Remarquons que pour étudier les mesures de Radon il est convenable de considérer l'espace métrique $\mathbf{S}$ séparable, complet et localement compact. Ces propriétés nous permettent de choisir une suite croissante de compacts dans $\mathbf{S}$, $\{\mathrm{K}_n\}$ qui vérifie les propriétés $\mathrm{K}_n \subset \mathrm{int}(\mathrm{K}_{n+1})$, $\mathrm{K}_n \uparrow \mathbf{S}$ (Lima [29], Prop. 4, page 278), où $\mathrm{int}(\mathrm{K}_n)$ désigne l'intérieur de l'ensemble K_n . Dans tout ce texte nous considérons cette suite croissante de compacts fixée.

1.2.1 Définitions

Classiquement la topologie vague est définie de façon semblable à celle de la topologie faible dans $\mathcal{M}_b$, sauf que nous utilisons des fonctions continues et à support compact.

Définition 1.2.1 *Une suite de mesures $\{\mu_n\}$ dans $\mathcal{M}_c$ converge vaguement vers $\mu \in \mathcal{M}_c$ si, pour toute fonction $f \in \mathrm{C}_c(\mathbf{S})$, se vérifie la convergence $\mu_n j \longrightarrow \mu f$. On écrira $\mu_n \xrightarrow{v} \mu$.*

Une base de voisinages dans la topologie vague d'une mesure de Radon μ est la famille des ensembles

$$\mathrm{V}(\mu, f_1, \ldots, f_n, \varepsilon) = \{\nu \in \mathcal{M}_c : |\mu f_i - \nu f_i| < \varepsilon, i = 1, \ldots, n\}$$

où $\varepsilon > 0, f_1, \ldots, f_n \in \mathrm{C}_c(\mathbf{S})$, $n \in \mathbb{N}$. Pour cette topologie sont vraies des propriétés semblables aux propriétés que la topologie faible vérifiait:

- Il suffit de définir la base de voisinages pour l'origine (mesure nulle). Pour les autres mesures on obtient la base de voisinages par translation;

- la topologie vague est la topologie engendrée par la famille d'applications $\{\Pi_f, f \in \mathrm{C}_c(\mathrm{S})\}$;
- l'espace $\mathcal{M}_c$ avec la topologie vague est un espace vectoriel topologique de Hausdorff complètement régulier.

Considérons maintenant $\mu \in \mathcal{M}_c$ et $f \in \mathrm{C}_c(\mathrm{S})$. Représentons par $f\mu$ la mesure de densité f par rapport à μ. Alors, pour tout borélien $\mathrm{A} \in \mathcal{S}$

$$|(f\mu)(\mathrm{A})| = |\int_{\mathrm{A}} f d\mu| \leq ||f|| \, |\mu|(\mathrm{supp}(f)),$$

où $\mathrm{supp}(f)$ désigne le support de la fonction f. Donc la mesure $f\mu$ est une mesure bornée sur $\mathbf{S}$. Pour chaque fonction $f \in \mathrm{C}_c(\mathrm{S})$ définissons l'application

$$\begin{aligned} \psi_f : \mathcal{M}_c &\longrightarrow \mathcal{M}_b \\ \mu &\hookrightarrow f\mu. \end{aligned}$$

Dans $\mathcal{M}_b$ nous avons défini la topologie faible donc, nous pouvons définir sur l'espace $\mathcal{M}_c$ la topologie engendrée par les applications $\{\psi_f, f \in \mathrm{C}_c(\mathrm{S})\}$.

Théorème 1.2.2 *La topologie engendrée sur l'espace $\mathcal{M}_c$ par la famille d'applications $\{\psi_f, f \in \mathrm{C}_c(\mathrm{S})\}$ coïncide avec la topologie vague.*

Démonstration : La démonstration suit les idées de Bonkian [6], page 9. Nous allons démontrer que les voisinages d'une mesure μ fixée sont les mêmes dans les deux topologies.

Une base de voisinages de μ pour la topologie engendrée par la famille $\{\psi_f, f \in \mathrm{C}_c(\mathrm{S})\}$ est constituée par les ensembles de la forme

$$\mathrm{U}(\mu, f_i, \mathrm{W}_i, i = 1, \ldots, n) = \bigcap_{i=1}^{n} \psi_{f_i}^{-1}(\mathrm{W}_i)$$

où $n \in \mathbb{N}$, $f_1, \ldots, f_n \in \mathrm{C}_c(\mathrm{S})$ et $\mathrm{W}_1, \ldots, \mathrm{W}_n$ sont des voisinages de $\psi_{f_i}(\mu) = f_i\mu$ dans la topologie faible de $\mathcal{M}_b$. Il suffit de prendre les W_i de la forme $\mathrm{W}_i = \mathrm{W}(f_i\mu, g_1, \ldots, g_{k_i}, \varepsilon_i)$, où les fonctions $g_j \in \mathrm{C}(\mathrm{S})$. Alors

$$\mathrm{U}(\mu, f_i, \mathrm{W}_i, i = 1, \ldots, n) =$$

$$= \bigcap_{i=1}^{n} \psi_{f_i}^{-1} \left\{ \nu \in \mathcal{M}_b : |\int g_k d\nu - \int g_k d(f_i\mu)| < \varepsilon_i, k = 1, \ldots, k_i \right\} =$$

$$= \bigcap_{i=1}^{n} \left\{ \lambda \in \mathcal{M}_c : |\int g_k d(f_i\lambda) - \int g_k d(f_i\mu)| < \varepsilon_i, k = 1, \ldots, k_i \right\} =$$

$$= \bigcap_{i=1}^{n} \left\{ \lambda \in \mathcal{M}_c : |\int g_k f_i d\lambda - \int g_k f_i d\mu| < \varepsilon_i, k = 1, \ldots, k_i \right\}.$$

Les fonctions $g_k f_i$ sont évidemment continues et à support compact car $\mathrm{supp}(g_k f_i) \subset \mathrm{supp}(f_i)$. Donc, les ensembles $\mathrm{U}(\mu, f_i, \mathrm{W}_i, i = 1, \ldots, n)$ sont des voisinages de la mesure μ pour la topologie vague.

Réciproquement, considérons un voisinage vague de $\mu, \mathrm{V}(\mu, f_1, \ldots, f_n, \varepsilon)$, où $\varepsilon > 0, n \in \mathbb{N}$, $f_1, \ldots, f_n \in \mathrm{C}_c(\mathbf{S})$. Pour chaque $i = 1, \ldots, n$ définissons une fonction g_i, continue à support compact et à valeurs dans $[0,1]$ telle que

$$g_i(s) = \begin{cases} 1 & \text{si } s \in \mathrm{supp}(f_i) \\ 0 & \text{si } s \in \{t \in \mathbf{S} : d(t, \mathrm{supp}(f_i)) \geq 1\}. \end{cases}$$

Alors

$$\mathrm{V}(\mu, f_1, \ldots, f_n, \varepsilon) =$$

$$= \bigcap_{i=1}^{n} \left\{ \nu \in \mathcal{M}_c : |\int g_i f_i d\nu - \int g_i f_i d\mu| < \varepsilon \right\} =$$

$$= \bigcap_{i=1}^{n} \left\{ \nu \in \mathcal{M}_c : |\int f_i d(g_i\nu) - \int f_k d(g_i\mu)| < \varepsilon \right\} =$$

$$= \bigcap_{i=1}^{n} \psi_{g_i}^{-1} \left(\mathrm{W}(g_i\mu, f_i, \varepsilon)\right).$$

Par conséquent, tout voisinage vague de μ est un voisinage de μ dans la topologie engendrée par la famille d'applications $\{\psi_f, f \in \mathrm{C}_c(\mathbf{S})\}$. ■

1.2.2 Compacité relative vague

Nous rappelons ici quelques résultats de caractérisation de compacité relative vague qui peuvent être déduits de la définition classique 1.2.1. Remarquons que la topologie vague dans $\mathcal{M}_c$ peut être interprétée comme une limite inductive de topologies faibles sur des espaces duaux des espaces de Banach $\mathrm{C}(\mathrm{K}_n)$, l'espace des fonctions continues à support contenu dans K_n (voir Marle [30], ch. 9). De cette façon on peut s'attendre à des caractérisations des ensembles vaguement relativement compacts qui ressemble aux caractérisations pour les topologies faibles au sens de l'analyse fonctionnelle. En effet,

Théorème 1.2.3 (Marle [30], prop. 9.8.7) *Soit $M \subset \mathcal{M}_c$. Les propriétés suivantes sont équivalentes.*

1. *M est vaguement relativement compact;*
2. *Pour toute fonction $f \in \mathrm{C}_c(\mathbf{S})$, $\sup_{\mu \in \mathrm{M}} |\mu f| < \infty$;*
3. *Pour tout borelien relativement compact $B \in \mathcal{B}, \sup_{\mu \in \mathrm{M}} |\mu|(B) < \infty$.*

Notons par $\mathcal{M}_c^+$ le sous-espace de $\mathcal{M}_c$ des mesures de Radon non négatives. Alors d'après la condition 3 le résultat suivant est immédiat .

Corollaire 1.2.4 *Un ensemble $M \subset \mathcal{M}_c$ est vaguement relativement compact si et seulement si $|M|$ est vaguement relativement compact dans $\mathcal{M}_c^+$.*

Liées aux propriétés de compacité relative sont les propriétés de métrisabilité. En effet, en général la topologie vague dans $\mathcal{M}_c$ n'est pas métrisable. Varadarajan ([47], th. 16, page 193), a démontré que la topologie faible dans $\mathcal{M}_b$ n'est métrisable que si $\mathbf{S}$ a un nombre fini d'éléments. Donc, d'après le théorème 1.2.2, la même propriété est valable en ce qui concerne la topologie vague de $\mathcal{M}_c$. Pourtant, il est possible de démontrer que certains sous-ensembles de $\mathcal{M}_c$ sont métrisables.

Théorème 1.2.5 (Marle [30], prop. 9.8.10) *Tout sous-ensemble M de $\mathcal{M}_c$ relativement vaguement compact est métrisable pour la topologie vague.*

Pour certains ensembles relativement vaguement compact il est même possible de démontrer la séparabilité. D'abord remarquons que l'espace $\mathrm{C}_c(\mathbf{S})$ est séparable. En effet, comme chaque K_n est un compact, l'espace $\mathrm{C}(\mathrm{K}_n)$ est séparable. Donc, pour chaque $n \in \mathbb{N}$ fixé, il existe une suite $\{f_{k,n}\}_{k \in \mathbb{N}}$ qui est dense dans $\mathrm{C}(\mathrm{K}_n)$. Alors l'ensemble $\{f_{k,n}\}_{k,n \in \mathbb{N}}$ est dense dans $\mathrm{C}_c(\mathbf{S})$. Pour le vérifier il suffit de remarquer que si $f \in \mathrm{C}_c(\mathbf{S})$, alors il existe un entier $n \in \mathbb{N}$ tel que $\mathrm{supp}(f) \subset \mathrm{K}_n$.

Théorème 1.2.6 *Les ensembles $\mathcal{M}_c^k = \{\mu \in \mathcal{M}_c : |\mu|(\mathbf{S}) \leq k\}$ sont métrisables séparables, pour chaque $k > 0$.*

<u>**Démonstration**</u> : Choisissons $\{g_1, \ldots, g_r, \ldots\}$ un ensemble de fonctions continues à support compact, dense dans $\mathrm{C}_c(\mathbf{S})$. Définissons la fonction

$$\begin{array}{rccl} \mathbf{T}: & \mathcal{M}_c & \longrightarrow & \mathbb{R}^\infty \\ & \mu & \hookrightarrow & (\mu g_1, \ldots, \mu g_r, \ldots). \end{array}$$

La fonction $\mathbf{T}$ est injective: en fait, si $\mathbf{T}(\mu) = \mathbf{T}(\nu)$, où $\mu, \nu \in \mathcal{M}_c$, on déduit $\mu g_r = \nu g_r, r \in \mathbb{N}$. Etant donné $f \in \mathrm{C}_c(\mathrm{S})$, f est la limite dans $\mathrm{C}_c(\mathrm{S})$ d'une sous-suite de la suite de fonctions g_r. Donc $\mu f = \lim \mu g_{r_j} = \lim \nu g_{r_j} = \nu f$, et alors $\mu = \nu$. La continuité de $\mathbf{T}$ est évident d'après la définition de la topologie vague de $\mathcal{M}_c$. Prouvons maintenant que la fonction $\mathbf{T}^{-1}$ est aussi continue sur $\mathcal{M}_c^k$. Soit $\{\mu_\alpha\}$ une suite généralisée dans $\mathcal{M}_c^k$ telle que $\mathbf{T}(\mu_\alpha) \longrightarrow \mathbf{T}(\mu)$. Alors, pour tout $r \in \mathbb{N}$, $\mu_\alpha g_r \longrightarrow \mu g_r$. Soit $g \in \mathrm{C}_c(\mathrm{S})$,

$$|\mu_\alpha g - \mu g| \leq |\mu_\alpha g - \mu_\alpha g_r| + |\mu_\alpha g_r - \mu g_r| + |\mu g_r - \mu g| \leq$$

$$\leq |\mu_\alpha|(\mathrm{S})||g - g_r|| + |\mu_\alpha g_r - \mu g_r| + |\mu|(\mathrm{S})||g - g_r|| \leq$$

$$\leq 2k||g - g_r|| + |\mu_\alpha g_r - \mu g_r|.$$

Donc $\lim_\alpha |\mu_\alpha g - \mu g| \leq 2k||g - g_r||$, $r \in \mathbb{N}$, d'où nous pouvons conclure $\lim_\alpha |\mu_\alpha g - \mu g| = 0$. Nous avons donc démontré que $\mathcal{M}_c^k$ est homéomorphe à une partie de $\mathbb{R}^\infty$, donc $\mathcal{M}_c^k$, muni de la topologie vague trace, est métrisable et séparable. ■

En utilisant le théorème 1.2.2 nous pouvons démontrer une condition de compacité relative vague en utilisant la topologie faible dans $\mathcal{M}_b$.

Théorème 1.2.7 *Un sous-ensemble M de $\mathcal{M}_c$ est vaguement relativement compact si et seulement si, pour toute fonction $f \in \mathrm{C}_c(\mathrm{S})$, l'ensemble $\psi_f(M)$ est faiblement relativement compact dans $\mathcal{M}_b$.*

Démonstration : Nous allons suivre le plan de démonstration utilisé par Bonkian [6], page 11. D'après le théorème 1.2.2 les applications ψ_f sont continues, donc si M est vaguement relativement compact, $\psi_f(\overline{M})$ est un compact faible. Comme $\psi_f(\mathrm{M}) \subset \psi_f(\overline{M})$ il en résulte que $\psi_f(\mathrm{M})$ est faiblement relativement compact.

Pour démontrer la réciproque posons, pour toute fonction $f \in \mathrm{C}_c(\mathrm{S})$, $\mathcal{M}_f = \mathcal{M}_b$, et plongeons $\mathcal{M}_c$ dans $\prod_{f \in \mathrm{C}_c(\mathrm{S})} \mathcal{M}_f$. On définit une application ψ à partir des applications $\psi_f, f \in \mathrm{C}_c(\mathrm{S})$ par,

$$\psi(\mu) = (\psi_f(\mu))_{f \in \mathrm{C}_c(\mathrm{S})}.$$

Alors $\psi(\mathcal{M}_c) \subset \prod_{f \in \mathrm{C}_c(\mathrm{S})} \mathcal{M}_f$, et ψ est surjective de $\mathcal{M}_c$ dans l'image de $\mathcal{M}_c$. Vérifions maintenant que ψ est injective. Si $\mu, \nu \in \mathcal{M}_c$ et $\mu \neq \nu$, il existe une fonction $f \in C_c(S)$ telle que $\mu f \neq \nu f$ et donc il existe un ensemble $B \in \mathcal{B}$ tel que $(f\mu)(B) \neq (f\nu)(B)$ (il suffit de prendre $B = \text{supp}(f)$), ce qui montre que ψ est injective. Comme ψ est bijective sur $\psi(\mathcal{M}_c)$ nous pouvons identifier les éléments de $\psi(\mathcal{M}_c)$ avec les éléments de $\mathcal{M}_c$ de façon évidente.

Soit $\mathcal{T}$ la topologie engendrée dans $\mathcal{M}_c$ par l'application ψ, en considérant la topologie produit dans $\prod_{f \in C_c(S)} \mathcal{M}_f$. Un système fondamental de voisinages pour la topologie produit est de la forme $\prod_{f \in C_c(S)} W_f$, où W_f est un voisinage de la mesure correspondante à la coordonnée associée à f dans $\mathcal{M}_b$ pour la topologie faible, et où les W_f sont égaux à $\mathcal{M}_b$, sauf un nombre fini. Alors un système fondamental de voisinages de μ dans $\mathcal{M}_c$ pour la topologie $\mathcal{T}$ est de la forme

$$\bigcap_{f \in F} \{\nu \in \mathcal{M}_c : \ f\nu \in W_f\}$$

où F est une partie finie de $C_c(S)$ et, pour tout $f \in F$, W_f est un voisinage de $f\mu$ dans la topologie faible de $\mathcal{M}_b$. Mais ce système fondamental n'est autre qu'un système fondamental de voisinages pour la topologie engendrée par $\{\psi_f, f \in C_c(S)\}$, ce qui montre que ψ est un homéomorphisme de $\mathcal{M}_c$, avec la topologie vague, sur son image.

Finalement, les inclusions

$$M = \prod_{f \in C_c(S)} \psi_f(M) \subset \prod_{f \in C_c(S)} \overline{\psi_f(M)}$$

et le théorème de Tychonov montrent que, si $\psi_f(M)$ est faiblement relativement compact pout tout $f \in C_c(S)$, alors M est vaguement relativement compact. ■

Nous allons encore démontrer un théorème pour la topologie vague équivalent au théorème 1.1.7, mais avant cela nous avons besoin d'établir quelques résultats concernant les suites vaguement convergentes.

1.2.3 Convergence vague de suites de mesures dans $\mathcal{M}_c$

D'après le théorème 1.2.2 nous savons déjà qu'une suite de mesures de Radon $\{\mu_n\}$ converge vaguement vers $\mu \in \mathcal{M}_c$ si et seulement si, pour tout $f \in C_c(S)$, $f\mu_n \xrightarrow{w} f\mu$. Nous allons démontrer qu'il n'est pas nécessaire de considérer toutes les fonctions $f \in C_c(S)$, mais seulement une quantité dénombrable de fonctions convenablement choisies.

D'abord, choisissons une suite croissante $\{G_k\}$ d'ouverts dans l'anneau $\mathcal{B}$, telle que $G_k \uparrow S$. Pour chaque $k \in \mathbb{N}$ soit g_k un élément de $C_c(S)$ tel que $\mathbb{1}_{G_k} \leq g_k \leq 1 = \mathbb{1}_S$, où $\mathbb{1}_B$ représente la fonction indicatrice de l'ensemble B (pour l'existence de ces fonctions voir Kallenberg [26], A.6.1). Dans le reste de cette section nous considérerons les suites $\{G_k\}$ et $\{g_k\}$ fixées.

Théorème 1.2.8 *Considérons une suite de mesures de Radon* $\{\mu_n\}$ *et* μ *un élément de* $\mathcal{M}_c$. $\mu_n \xrightarrow{v} \mu$ *si et seulement si, pour tout* $k \in \mathbb{N}$, $g_k\mu_n \xrightarrow{w} \mu'_k$, *et dans ce cas* $\mu'_k = g_k\mu$.

<u>**Démonstration**</u> : Si $\mu_n \xrightarrow{v} \mu$, l'implication énoncée est évidente d'après le théorème 1.2.2. Supposons donc que $g_k\mu_n \xrightarrow{w} \mu'_k$, pour tout $k \in \mathbb{N}$. La difficulté est la construction d'une mesure de Radon telle que $\mu'_k = g_k\mu$. Cela étant résolu la convergence vague se déduit aisément. Fixons $k \in \mathbb{N}$. Il est évident que pour toute fonction $f \in C_c(S)$ telle que $\mathrm{supp}(f) \subset G_k$ se vérifient les égalités $\mu'_k f = \mu'_{k+1} f = \cdots$. Donc, pour tout ouvert $G \subset G_k$ nous avons $\mu'_k(G) = \mu'_{k+1}(G) = \cdots$, et ces égalités de mesures peuvent encore s'étendre à tout ensemble de la forme $G = B \cap G_k$, $B \in \mathcal{S}$. On définit donc bien une mesure de Radon en posant, pour tout borélien borné B contenu dans un G_k : $\mu(B) = \mu'_k(B)$. La mesure μ vérifie $\mu'_k = g_k\mu, k \in \mathbb{N}$. Il nous reste à vérifier que $\mu_n \xrightarrow{v} \mu$. Soit $f \in C_c(S)$, alors il existe un $k \in \mathbb{N}$ tel que $\mathrm{supp}(f) \subset G_k$, donc

$$\mu_n f = \mu_n(fg_k) = (g_k\mu_n)f \longrightarrow (g_k\mu)f = \mu_n(fg_k) = \mu_n f. \blacksquare$$

Maintenant nous pouvons démontrer un théorème analogue à 1.1.17 pour les mesures de Radon.

Théorème 1.2.9 *Soit* $\{\mu_n\}$ *une suite dans* $\mathcal{M}_c$ *vaguement convergente vers* $\mu \in \mathcal{M}_c$. *Si* $B \in \mathcal{B}$ *est tel que, pour tout point d'accumulation* ν *de l'ensemble* $\{|\mu_n|\}$, *se vérifie* $\nu(fr(B)) = 0$, *alors* $\lim_{n\to\infty} \mu_n(B) = \mu(B)$.

<u>**Démonstration**</u> : Comme $B \in \mathcal{B}$, il existe un $k \in \mathbb{N}$ tel que $B \subset G_k$. Considérons ce k fixé. D'après le théorème précédent $g_k\mu_n \xrightarrow{w} g_k\mu$. Du fait que la fonction g_k est positive on déduit $|g_k\mu_n| = g_k|\mu_n|$, et, de la même façon que dans la démonstration du théorème précédent, il s'ensuit que tout point d'accumulation faible de $\{g_k|\mu_n|\}$ est de la forme $g_k\nu$, où ν est un point d'accumulation vague de $\{|\mu_n|\}$. D'après les hypothèses sur

l'ensemble B on vérifie que $(g_k\nu)(\mathrm{fr}(B)) = 0$, donc, d'après le théorème 1.1.17 $(g_k\mu_n)(B) \longrightarrow (g_k\mu)(B)$ ce qui est équivalent à $\mu_n(B) \longrightarrow \mu(B)$, d'après les propriétés de la fonction g_k. ■

Nous pouvons maintenant démontrer la version pour la topologie vague de théorème 1.1.7.

Théorème 1.2.10 *Soit $M \subset \mathcal{M}_c$. Les conditions suivantes sont équivalentes.*

1. *M est vaguement relativement compact dans $\mathcal{M}_c$;*
2. *M est vaguement relativement sequentiellement compact dans $\mathcal{M}_c$;*
3. *$|M|$ est vaguement relativement compact dans $\mathcal{M}_c^+$;*
4. *M^+ et M^- sont vaguement relativement compacts dans $\mathcal{M}_c^+$.*

<u>Démonstration</u> : L'équivalence $1 \Leftrightarrow 3$ s'ensuit d'après 3 du théorème 1.2.3. Les conditions 1 et 3 ensemble entrainent 4, et la condition 4 implique 1 et 3. Il reste donc à démontrer l'équivalence $1 \Leftrightarrow 2$.

Supposons que M soit vaguement relativement compact. D'après le théorème 1.2.7, pour toute fonction $f \in C_c(S)$, l'ensemble

$$\psi_f(M) = fM = \{f\mu : \mu \in C_c(S)\}$$

est faiblement relativement compact dans $\mathcal{M}_b$. Reprenons les fonctions g_k que nous avions fixées. Donc, pour tout $k \in \mathbb{N}$, l'ensemble $g_k M$ est faiblement relativement compact, et, d'après le théorème 1.1.7, il est aussi faiblement relativement sequentiellement compact. Considérons une suite $\{\mu_n\}$ d'éléments de M. De cette suite nous pouvons extraire une sous-suite $\{\mu_{1,n}\}$ telle que la suite $\{g_1\mu_{1,n}\}$ soit faiblement convergente vers une mesure $\lambda_1 \in \mathcal{M}_b$. De la même façon, de la suite $\{\mu_{1,n}\}$ nous pouvons extraire une sous-suite $\{\mu_{2,n}\}$ telle que la suite $\{g_2\mu_{2,n}\}$ soit faiblement convergente vers une mesure $\lambda_2 \in \mathcal{M}_b$. Evidemment, nous avons aussi $g_1\mu_{2,n} \xrightarrow{w} \lambda_1$. Poursuivant de cette façon nous obtenons une collection de sous-suites $\{\mu_{k,n}\}$ telle que

- pour chaque $k \in \mathbb{N}$, $\{\mu_{k+1,n}\}_{n\in\mathbb{N}}$ est une sous-suite de $\{\mu_{k,n}\}_{n\in\mathbb{N}}$;
- pour chaque $k \in \mathbb{N}$, $j = 1, \ldots, k$, $g_j\mu_{k,n} \xrightarrow{w} \lambda_j$, quand $n \longrightarrow \infty$.

En posant $\nu_n = \mu_{n,n}$, nous obtenons une sous-suite de la suite initiale $\{\mu_n\}$ telle que $g_k\nu_n \xrightarrow{w} \lambda_k$, quand $n \longrightarrow \infty$, pour chaque $k \in \mathbb{N}$. Donc, d'après le théorème 1.2.8, $\nu_n \xrightarrow{v} \nu \in \mathcal{M}_c$ telle que $\lambda_k = g_k\nu$ ce qui démontre que M est vaguement relativement sequentiellement compact.

Réciproquement, si M est vaguement relativement sequentiellement compact on déduit immédiatement du théorème 1.2.2 que, pour toute fonction $f \in \mathbf{C}_c(\mathbf{S})$, l'ensemble $\psi_f(\mathrm{M})$ est faiblement relativement sequentiellement compact. Donc, d'après le théorème 1.1.7, pour tout $f \in \mathbf{C}_c(\mathbf{S})$, $\psi_f(\mathrm{M})$ est faiblement relativement compact, d'où il s'ensuit, d'après le théorème 1.2.7, que M est vaguement relativement compact. ■

Finalement, le théorème 1.1.13 a un analogue pour la convergence vague qui se démontre de la même façon que le théorème 1.1.13, donc nous ne présentons pas la démonstration.

Théorème 1.2.11 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_c$, $\mu \in \mathcal{M}_c$ et $\nu \in \mathcal{M}_c^+$.*

1. *Si, pour tout ensemble $A \in \mathcal{B}$ tel que $\nu(fr(A)) = |\mu(fr(A))|$, se vérifient*
   (a) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$;
   (b) $\lim_{n\to\infty} |\mu_n|(A) = \nu(A)$,

   alors $\mu_n \xrightarrow{w} \mu$ et $|\mu_n| \xrightarrow{w} \nu$.
2. *Réciproquement, si $\mu_n \xrightarrow{w} \mu$ et $|\mu_n| \xrightarrow{w} \nu$ alors, pour tout ensemble $A \in \mathcal{B}$ tel que $\nu(fr(A)) = |\mu(fr(A))| = 0$*
   (a) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$;
   (b) $\lim_{n\to\infty} |\mu_n|(A) = \nu(A)$.

Chapitre 2

Les mesures aléatoires

Dans ce deuxième chapitre nous allons nous intéresser à l'étude des mesures aléatoires de Radon à signe. La première partie de ce chapitre sera consacrée aux problèmes de la définition d'une mesure aléatoire. Ce type de problème ne se pose pas quand on s'intéresse à l'étude des mesures aléatoires de Radon non négatives. Les problèmes qu'on doit résoudre sont liés aux propriétés de non séparabilité et de non métrisabilité de l'espace $\mathcal{M}_c$ avec la topologie vague. Quand on travaille sur $\mathcal{M}_c^+$ avec la topologie vague, on travaille sur un espace métrique séparable et complet ([7], page 61), et dans ce cas les questions d'égalité de tribus qui se poseront pour l'étude sur $\mathcal{M}_c$ ne se posent pas. Après avoir résolu ces questions nous nous intéresserons aux problèmes habituels: caractérisation de l'égalité en loi de mesures aléatoires, l'étude de conditions nécessaires et suffisantes de convergence en loi et la caractérisation des lois de dimension finie. Pour terminer ce chapitre général sur les mesures aléatoire on s'intéresse à l'étude de la convergence en moyenne quadratique de mesures aléatoires. Relativement à ce type de convergence nous pouvons démontrer quelques conditions suffisantes de convergence récupèrant partiellement les caractérisations que nous démontrerons pour le cas des mesures non négatives.

2.1 Les tribus sur $\mathcal{M}_c$

Nous considérerons l'espace des mesures de Radon à signe $\mathcal{M}_c$ toujours avec la topologie vague. Dans le premier chapitre nous avons fait l'étude de cet espace et nous savons déjà que l'espace n'est pas métrisable ni séparable.

Cela pose des problèmes quand on veut identifier des classes de fonctions qui engendrent les tribus sur $\mathcal{M}_c$. Soit $\mathcal{T}_c$ la tribu borélienne associée à la topologie vague sur $\mathcal{M}_c$. Nous savons que la topologie vague est engendrée par la famille d'applications $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$. Comme l'espace $\mathcal{M}_c$ avec cette topologie n'est pas séparable nous ne pouvons pas affirmer que la tribu $\mathcal{T}_c$ soit engendrée par la famille d'applications $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$. Pour l'instant, tout ce qu'on peut dire c'est que les applications π_f sont mesurables, pour toute fonction $f \in \mathrm{C}_c(\mathbf{S})$.

Introduisons les projections associées aux boréliens relativement compacts, $\mathrm{B} \in \mathcal{B}$

$$\begin{array}{rccl} \pi_{\mathrm{B}}: & \mathcal{M}_c & \longrightarrow & \mathbb{R} \\ & \mu & \hookrightarrow & \mu(\mathrm{B}). \end{array}$$

Les applications π_{B}, $\mathrm{B} \in \mathcal{B}$, sont aussi mesurables par rapport à la tribu $\mathcal{T}_c$. Quand on s'intéresse aux mesures non négatives, soit à l'étude de cette tribu sur le sous-espace $\mathcal{M}_c^+$, on peut dire tout de suite que la tribu borélienne $\mathcal{T}_c^+$ est engendrée par la famille $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$, car la topologie vague sur $\mathcal{M}_c^+$ est séparable. On démontre encore que la tribu $\mathcal{T}_c^+$ est la tribu engendrée par la famille d'applications $\{\pi_{\mathrm{B}},\ \mathrm{B} \in \mathcal{B}\}$ ([26], th. 1.3). On verra plus tard que l'on peut récupérer ce résultat pour la tribu $\mathcal{T}_c$ sur $\mathcal{M}_c$.

Définition 2.1.1 *1. $\mathcal{T}_1$ est la tribu sur $\mathcal{M}_c$ engendrée par la famille d'applications $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$.*

2. $\mathcal{T}_2$ est la tribu sur $\mathcal{M}_c$ engendrée par la famille d'applications $\{\pi_{\mathrm{B}},\ \mathrm{B} \in \mathcal{B}\}$.

Rappelons maintenant que les ensembles $\mathcal{M}_c^k = \{\mu \in \mathcal{M}_c :\ |\mu|(\mathbf{S}) \leq k\}$, $k > 0$, sont métrisables, séparables (théorème 1.2.6) et compacts (théorème 1.2.3) pour la topologie vague, donc à base dénombrable.

Définition 2.1.2 *On appelera fonction simple toute fonction mesurable qui ne prend qu'un nombre fini de valeurs.*

Théorème 2.1.3 *1. Les tribus $\mathcal{T}_1$ et $\mathcal{T}_2$ sur $\mathcal{M}_c$ coïncident.*

2. Les tribus $\mathcal{T}_1$, $\mathcal{T}_2$ et $\mathcal{T}_c$ sur $\mathcal{M}_c^k$ coïncident, pour tout $k > 0$.

<u>**Démonstration**</u> : 1. a) Si f est une fonction simple à support compact, l'application π_f est $\mathcal{T}_2$-mesurable. Si f est une fonction continue à support

compact positive, il existe une suite $\{f_n\}$ de fonctions simples à support compact telle que $0 \leq f_n \uparrow f$. Donc, pour toute mesure $\mu \in \mathcal{M}_c$

$$\mu f = \mu^+ f - \mu^- f = \lim_{n\to\infty} \mu^+ f_n - \lim_{n\to\infty} \mu^- f_n = \lim_{n\to\infty} \mu f_n.$$

Alors l'application π_f est $\mathcal{T}_2$-mesurable, pour $f \in \mathbf{C}_c(\mathbf{S})$ positive.

Finalement, si $f \in \mathbf{C}_c(\mathbf{S})$, la fonction peut s'écrire comme la différence de deux fonctions continues à support compact positives f^+ et f^- : $f = f^+ - f^-$. Donc l'application π_f est $\mathcal{T}_2$-mesurable.

b) Soit C un ensemble compact dans $\mathbf{S}$, alors il existe un entier n_0 tel que $\mathrm{C} \subset \mathrm{int}(\mathrm{K}_{n_0}) \subset \mathrm{K}_{n_0}$, et donc, pour $\delta > 0$ assez petit, $\mathrm{C} \subset \mathrm{C}^\delta \subset \mathrm{K}_{n_0}$, où $\mathrm{C}^\delta = \{s \in \mathbf{S} : d(s, \mathrm{C}) < \delta\}$. D'après le théorème de Tietze-Urysohn ([30], 7.3.8), il existe une fonction f à valeurs dans [0,1], égale à 0 sur le complémentaire de C^δ et égale à 1 sur C. D'après l'inclusion $\mathrm{C}^\delta \subset \mathrm{K}_{n_0}$ il s'ensuit que f est à support compact. On peut donc construire une suite de fonctions continues à support compact $\{f_n\}$ telle que $f_n \downarrow \mathbb{1}_{\mathrm{C}}$. Comme la fonction f_1 est μ^+ et μ^- intégrable pour toute mesure $\mu \in \mathcal{M}_c$, le théorème de la convergence dominée implique

$$\mu(\mathrm{C}) = \mu^+(\mathrm{C}) - \mu^-(\mathrm{C}) = \lim_{n\to\infty} \mu^+ f_n - \lim_{n\to\infty} \mu^+ f_n = \lim_{n\to\infty} \mu f_n,$$

ce qui montre que l'application π_{C} est $\mathcal{T}_1$-mesurable, pour tout C compact.

Alors, C étant un compact fixé, la classe $\mathcal{D}$ des boréliens D contenus dans C tels que l'application π_{D} soit $\mathcal{T}_1$-mesurable est un λ-système de parties de C qui contient C et le π-système des compacts de C. Par conséquent $\mathcal{D}$ contient la tribu engendrée par les compacts de C, c'est à dire, la tribu borélienne du sous-espace C de $\mathbf{S}$. Il en résulte que, pour tout borélien borné D, l'application π_{D} est $\mathcal{T}_1$-mesurable.

2. Tout ouvert de la topologie vague sur $\mathcal{M}_c$ est réunion d'ensembles ouverts de la forme $\{\mu : |\mu f_i| < \varepsilon_i, i = 1, \ldots, n\}$, où $f_i \in \mathbf{C}_c(\mathbf{S})$, $i = 1, \ldots, n$. Fixons $k > 0$. On sait que $\mathcal{M}_c^k$ est métrisable compact, donc à base dénombrable. Tout ouvert vague de $\mathcal{M}_c^k$ est donc réunion dénombrable d'ensembles ouverts de la forme $\{\mu : |\mu f_i| < \varepsilon_i, i = 1, \ldots, n, |\mu|(\mathbf{S}) \leq k\}$. Ces ensembles engendrent la tribu $\mathcal{T}_1^k$, trace de $\mathcal{T}_1$ sur $\mathcal{M}_c^k$, puisque $\mathcal{T}_1$ est engendrée par les ensembles $\{\mu : |\mu f| \leq \varepsilon\}$, $f \in \mathbf{C}_c(\mathbf{S})$. D'autre part, comme $\mathcal{M}_c^k$ est borélien de $\mathcal{M}_c$ ($\mathcal{M}_c^k$ est vaguement compact), la trace de $\mathcal{T}_c$ sur $\mathcal{M}_c^k$ n'est autre que la tribu borélienne de $\mathcal{M}_c^k$, $\mathcal{T}_c^k$. On voit ainsi que, pour tout $k > 0$, $\mathcal{T}_c^k = \mathcal{T}_1^k$. ∎

Pour l'instant nous savons seulement que $\mathcal{T}_1=\mathcal{T}_2\subset\mathcal{T}_c$, donc nous pouvons travailler seulement avec $\mathcal{T}_1$ et $\mathcal{T}_c$. Etudions encore la mesurabilité de quelques applications qui nous serons utiles plus tard quand on voudra établir l'égalité de tribus $\mathcal{T}_1$ et $\mathcal{T}_c$.

Théorème 2.1.4 *Soit $B\in\mathcal{B}$. Définissons les applications de $\mathcal{M}_c$ dans $\mathbb{R}$:*

$$\begin{aligned} \pi_{\mathrm{B}}^-(\mu) &= \mu^-(B), \\ \pi_{\mathrm{B}}^+(\mu) &= \mu^+(B), \\ |\pi|_{\mathrm{B}}(\mu) &= |\mu|(B). \end{aligned}$$

Les applications π_{B}^+, π_{B}^- et $|\pi|_{\mathrm{B}}$ sont $\mathcal{T}_c$-mesurables.

<u>**Démonstration**</u> : D'après les relations entre les mesures μ, μ^+, μ^- et $|\mu|$ et le fait que π_{B} est $\mathcal{T}_c$-mesurable, il suffit de démontrer que $|\pi|_{\mathrm{B}}$ est $\mathcal{T}_c$-mesurable. De plus, il suffit de vérifier la mesurabilité pour B ouvert. Pour cela démontrons que $|\pi|_{\mathrm{B}}$ est semi-continue inférieurement pour la topologie vague sur $\mathcal{M}_c$. Comme $\mathcal{T}_c$ est la tribu borélienne associée, la mesurabilité s'ensuit ([7], page 181).

Fixons $\varepsilon > 0$ et $\mu\in\mathcal{M}_c$. Il existe une fonction continue à valeurs dans [0,1], h, nulle sur B^c telle que

$$\int_{\mathbf{S}} h d|\mu| \geq |\mu|(\mathrm{B}) - \frac{\varepsilon}{3}.$$

Soit $\mathbf{S}=\mathrm{D}^+\cup\mathrm{D}^-$, où D^+ et D^- sont les boréliens où μ^+ et μ^- sont concentrées, respectivement. Posons $g=\min(h,\mathbb{1}_{\mathrm{D}^+})+\max(-h,-\mathbb{1}_{\mathrm{D}^-})$. Il est clair que g est à valeurs dans [-1,+1], et que, pour tout borélien E

$$\int_{\mathrm{E}} h\ d|\mu| = \int_{\mathrm{E}} g\ d\mu.$$

Les mesures μ^+ et μ^- sont finies sur le sous-espace métrique $\mathrm{B}\subset\mathbf{S}$. Alors, d'après [38], page 31, il existe un compact K de $\mathbf{S}$, contenu dans B, tel que

$$\mu^+(\mathrm{B}\setminus\mathrm{K}) < \frac{\varepsilon}{12},\ \ \mu^-(\mathrm{B}\setminus\mathrm{K}) < \frac{\varepsilon}{12},$$

et tel que la restriction de g au compact K soit continue. La fonction g est aussi continue dans le fermé $(\mathbf{S}\setminus\mathrm{B})\cup\mathrm{K}$, car la fonction h est continue en $\mathbf{S}$.

Comme l'espace S est normal la fonction g peut se prolonger en une fonction f continue sur S, à valeurs dans [-1,+1] et nulle en dehors de B ([30], 7.2.9). Alors

$$\begin{aligned} \left|\int_S f\, d\mu - \int_S h\, d|\mu|\, \right| &= \left|\int_{B\setminus K} f\, d\mu - \int_{B\setminus K} h\, d|\mu|\, \right| \leq \\ &\leq \left|\int_{B\setminus K} f\, d\mu\right| + \left|\int_{B\setminus K} h\, d|\mu|\, \right| \leq\ 2|\mu|(B\setminus K) < \tfrac{\varepsilon}{3}. \end{aligned}$$

Par conséquent

$$\left|\int_S f\, d\mu\right| > \int_S h\, d|\mu| - \frac{\varepsilon}{3} > |\mu|(B) - \frac{2\varepsilon}{3}.$$

Remarquons que supp(f)$\subset \overline{B}$, donc supp(f) est compact, ci qui permet de conclure $f \in C_c(S)$. Ainsi, l'ensemble

$$V(\mu, f, \frac{\varepsilon}{3}) = \{\nu \in \mathcal{M}_c : \left|\int f\, d\mu - \int f\, \nu\right| < \frac{\varepsilon}{3}\}$$

est un voisinage de μ dans la topologie vague. Pour toute mesure $\nu \in V(\mu, f, \frac{\varepsilon}{3})$ il se vérifie

$$|\nu|(B) \geq \left|\int f\, d\nu\right| > \left|\int f\, d\mu\right| - \frac{\varepsilon}{3} > |\mu|(B) - \varepsilon,$$

c'est-à-dire $|\pi|_B(\nu) \geq |\pi|_B(\mu) - \varepsilon$, donc $|\pi|_B$ est semi-continue inférieurement sur $\mathcal{M}_c$ avec la topologie vague, d'où il s'ensuit que $|\pi|_B$ est mesurable par rapport à la tribu $\mathcal{T}_c$. ■

Ensuite nous allons démontrer le même résultat en utilisant la tribu $\mathcal{T}_1$, mais cette fois il faut procéder d'autre façon. Nous démontrerons les mêmes propriétés de mesurabilité pour des projections associées à des fonctions $f \in C_c(S)$. La mesurabilité des projections associées aux boréliens relativement compacts se déduit alors par la même procédure que dans la condition 1 du théorème 2.1.3.

Théorème 2.1.5 *Soit $f \in C_c(S)$. Définissons les applications de $\mathcal{M}_c$ dans $\mathbb{R}$:*

$$\begin{aligned} \pi_f^-(\mu) &= \mu^- f, \\ \pi_f^+(\mu) &= \mu^+ f, \\ |\pi|_f(\mu) &= |\mu| f. \end{aligned}$$

Les applications π_f^-, π_f^+ et $|\pi|_f$ sont $\mathcal{T}_1$-mesurables.

Démonstration : Démontrons la $\mathcal{T}_1$-mesurabilité de $|\pi|_f$, pour $f\in \mathrm{C}_c(\mathrm{S})$ positive. Nous pouvons écrire

$$|\pi|_f(\mu) = |\mu|f = \sup\{|\mu h|, 0 \leq |h| \leq f, h \in \mathrm{C}_c(\mathrm{S})\}$$

([30], 9.5.3). Comme $f \in \mathrm{C}_c(\mathrm{S})$, il existe $n_0 \in \mathbb{N}$ tel que $\mathrm{supp}(f) \subset \mathrm{int}(\mathrm{K}_{n_0})$. Posons

$$\mathrm{C}_f = \{h \in \mathrm{C}_c(\mathrm{S}) :\ 0 \leq |h| \leq f\}.$$

L'ensemble C_f est fermé, donc un sous-espace séparable de $\mathrm{C}_c(\mathrm{S})$ (puisque $\mathrm{C}_c(\mathrm{S})$ est lui même séparable). Remarquons que si $0 \leq |h| \leq f$ et $h \in \mathrm{C}_c(\mathrm{S})$ alors $h \in \mathrm{C}_f$, donc

$$|\mu|f = \sup\{|\mu h|, h \in \mathrm{C}_f\}.$$

De plus, si h, $h' \in \mathrm{C}_f$ et $\|h - h'\| < \varepsilon$ on déduit

$$|\mu h - \mu h'| \leq \|h - h'\|\ |\mu|(\mathrm{K}_{n_0}) < \varepsilon|\mu|(\mathrm{K}_{n_0}).$$

Soit $\{h_n\}_{n\in\mathbb{N}}$ un sous-ensemble dense dans C_f. D'après la définition de $|\mu|f$, pour tout $n \in\mathbb{N}$, il existe $g_n \in \mathrm{C}_f$ telle que $|\mu g_n - |\mu|f| < \frac{1}{n}$. Aussi, il existe $r_n \in \mathbb{N}$ tel que $|\mu g_n - \mu h_{r_n}| < \frac{1}{n}$. Donc $|\ |\mu|f - \mu h_{r_n}| < \frac{2}{n}$ et alors

$$|\mu|f = \sup_{n\in\mathbb{N}}\{|\mu h_n|, 0 \leq h_n \leq f\},$$

où la suite $\{h_n\}$ est fixée et ne dépend pas de la fonction f ni de la mesure μ. Donc il s'ensuit que l'application $|\pi|_f(\mu) = |\mu|f$ est $\mathcal{T}_1$-mesurable.

La mesurabilité de π_f^- et π_f^+ s'ensuit par composition des applications π_f et $|\pi|_f$. ∎

Maintenant nous pouvons démontrer la mesurabilité de la décomposition de Hahn-Jordan d'une mesure à signe.

Corollaire 2.1.6 *Les applications de $\mathcal{M}_c$ dans $\mathcal{M}_c^+$*

$$\begin{aligned} HJ^+(\mu) &= \mu^+, \\ HJ^-(\mu) &= \mu^-, \\ HJ^{||}(\mu) &= |\mu|. \end{aligned}$$

sont $\mathcal{T}_c$-mesurables et $\mathcal{T}_1$-mesurables.

<u>**Démonstration**</u> : Pour chaque B$\in \mathcal{B}$, nous pouvons décomposer l'application π_{B}^{+} de la façon suivante:

$$\begin{tikzcd}
(\mathcal{M}_c,\mathcal{T}_c) \arrow[rr, "\mathrm{HJ}^+"] & & (\mathcal{M}_c^+,\mathcal{T}_c^+) \arrow[dd, "\pi_{\mathrm{B}}^{+}|_{\mathcal{M}_c^+}"] \\
\mu \arrow[r] \arrow[dr, "\pi_{\mathrm{B}}^{+}"'] & \mu^+ \arrow[d] & \\
& \mu^+(\mathrm{B}) & (\mathbb{R},\mathcal{B}(\mathbb{R}))
\end{tikzcd}$$

D'après [26], page 4, l'application $\pi_{\mathrm{B}}^{+}|_{\mathcal{M}_c^+}$ est mesurable, donc, de la commutativité du diagramme il s'ensuit la mesurabilité de HJ^+ par rapport à la tribu $\mathcal{T}_c$. La mesurabilité des applications HJ^- et $\mathrm{HJ}^{||}$ se démontre de la même façon. En ce qui concerne de $\mathcal{T}_1$-mesurabilité il suffit de remarquer que $\mathcal{T}_c^+=\mathcal{T}_1^+$ ([26], page 4 et lemme 4.1). ■

Remarquons que, pour l'instant, nous disposons de deux tribus sur $\mathcal{M}_c$ qui se ressemblent beaucoup. En effet nous avons démontré que les tribus $\mathcal{T}_1$ et $\mathcal{T}_c$ ont exactement les mêmes propriétés en ce qui concerne la mesurabilité des possibles décompositions de mesures. Toutes ces coïncidences nous laissent deviner que les tribus sont en fait égales, mais, pour l'instant, nous ne disposons encore des moyens nécessaires pour démontrer cette égalité.

2.2 Mesure aléatoire

Comme nous disposons de deux tribus sur $\mathcal{M}_c$ nous pouvons définir deux types de mesures aléatoires selon la tribu qui nous intéresse.

Définition 2.2.1 *1. Une $\mathcal{T}_c$-mesure aléatoire est une application mesurable définie sur un espace de probabilité $(\Omega,\mathcal{F},\mathbf{P})$ presque sûrement à valeurs dans $(\mathcal{M}_c,\mathcal{T}_c)$.*

2. Une $\mathcal{T}_1$-mesure aléatoire est une application mesurable définie sur un espace de probabilité $(\Omega,\mathcal{F},\mathbf{P})$ presque sûrement à valeurs dans $(\mathcal{M}_c,\mathcal{T}_1)$.

Il est difficile de choisir quelle définition on doit privilégier. En fait les deux tribus sur $\mathcal{M}_c$ sont naturelles: $\mathcal{T}_c$ parce qu'il s'agit de la tribu borélienne de la topologie vague, $\mathcal{T}_1$ parce que c'est la tribu engendrée par la famille $\{\pi_f,\ f \in C_c(S)\}$ ayant donc des propriétés très convenables. Heureusement nous démontrerons que les deux définitions coïncident ce qui nous sauve du problème du choix de la tribu et, en même temps, nous facilite le travail avec la tribu $\mathcal{T}_c$. Remarquons encore que, comme $\mathcal{T}_1 \subset \mathcal{T}_c$, toute $\mathcal{T}_c$-mesure aléatoire est aussi une $\mathcal{T}_1$-mesure aléatoire. On démontrera, encore avant de démontrer l'égalité des tribus, que la réciproque est aussi vraie.

Théorème 2.2.2 *1. Toute $\mathcal{T}_c$-mesure aléatoire ξ est la différence de deux $\mathcal{T}_c^+$-mesures aléatoires positives, ξ^+ et ξ^-.*

2. Toute $\mathcal{T}_1$-mesure aléatoire ξ est la différence de deux $\mathcal{T}_1^+$-mesures aléatoires positives, ξ^+ et ξ^-.

<u>**Démonstration**</u> : Du corollaire 2.1.6 nous déduisons que l'application

$$\begin{array}{ccccc} (\Omega,\mathcal{F},\mathbf{P}) & \xrightarrow{\xi} & (\mathcal{M}_c,\mathcal{T}_c) & \xrightarrow{\mathrm{HJ}^+} & (\mathcal{M}_c^+,\mathcal{T}_c^+) \\ \omega & \hookrightarrow & \xi(\omega) & \hookrightarrow & \xi^+(\omega) \end{array}$$

est mesurable. Donc ξ^+ est une $\mathcal{T}_c^+$-mesure aléatoire. De même pour ξ^-. Evidemment nous avons $\xi = \xi^+ - \xi^-$.

De la même façon on démontre le resultat pour les $\mathcal{T}_1$-mesures aléatoires. ■

Remarquons que $\mathcal{T}_1^+ = \mathcal{T}_c^+$, donc il n'est pas nécessaire de dire qu'il s'agit de $\mathcal{T}_c^+$-mesures aléatoires positives ou de $\mathcal{T}_1^+$-mesures aléatoires positives car les deux notions sont équivalentes.

Théorème 2.2.3 *1. La différence de deux mesures aléatoires positives est une $\mathcal{T}_c$-mesure aléatoire.*

2. La différence de deux mesures aléatoires positives est une $\mathcal{T}_1$-mesure aléatoire.

<u>**Démonstration**</u> : Dans l'espace produit $\mathcal{M}_c^+ \times \mathcal{M}_c^+$ on considère la topologie produit des topologies vagues sur $\mathcal{M}_c^+$ et la tribu $\mathcal{T}_c^+ \otimes \mathcal{T}_c^+$. Comme $\mathcal{M}_c^+$ est, pour la topologie vague, un espace métrisable séparable, $\mathcal{T}_c^+ \otimes \mathcal{T}_c^+$ est la tribu borélienne de $\mathcal{M}_c^+ \times \mathcal{M}_c^+$. L'application

$$\begin{array}{cccc} \mathrm{d}: & (\mathcal{M}_c^+ \times \mathcal{M}_c^+, \mathcal{T}_c^+ \otimes \mathcal{T}_c^+) & \longrightarrow & (\mathcal{M}_c,\mathcal{T}_c) \\ & (\mu_1,\mu_2) & \hookrightarrow & \mu_1 - \mu_2 \end{array}$$

est mesurable, parce que l'application d est continue par rapport aux topologies définies. Soient ξ et η deux mesures aléatoires positives. Alors, la fonction

$$\begin{array}{rcl}(\xi,\eta) : (\Omega,\mathcal{F},\mathbf{P}) & \longrightarrow & (\mathcal{M}_c^+ \times \mathcal{M}_c^+, \mathcal{T}_c^+ \otimes \mathcal{T}_c^+) \\ \omega & \hookrightarrow & (\xi(\omega),\eta(\omega))\end{array}$$

est aussi mesurable. Donc, il s'ensuit que $\nu = \mathrm{d}(\xi,\eta) = \xi - \eta$ est une application mesurable à valeurs dans $(\mathcal{M}_c,\mathcal{T}_c)$, cela veut dire, une $\mathcal{T}_c$-mesure aléatoire. Pour terminer la démonstration il suffit de remarquer que toute $\mathcal{T}_c$-mesure aléatoire est une $\mathcal{T}_1$-mesure aléatoire. ■

Corollaire 2.2.4 *Toute $\mathcal{T}_1$-mesure aléatoire est aussi $\mathcal{T}_c$-mesure aléatoire.*

<u>**Démonstration**</u> : En fait, soit ξ une $\mathcal{T}_1$-mesure aléatoire. D'après le théorème 2.2.2, ξ est la différence de deux mesures aléatoires positives ξ^+ et ξ^-. Donc, d'après le théorème précédent la différence $\xi^+ - \xi^-$ est une $\mathcal{T}_1$-mesure aléatoire. Comme $\xi = \xi^+ - \xi^-$, le corollaire est démontré. ■

Nous avons donc démontré qu'en fait il n'existe qu'un type de mesures aléatoires. Alors nous n'aurons plus besoin de parler de $\mathcal{T}_c$-mesures aléatoires ou de $\mathcal{T}_1$-mesures aléatoires, mais seulement de mesures aléatoires.

En plus, nous pouvons maintenant démontrer l'égalité de tribus $\mathcal{T}_c$ et $\mathcal{T}_1$.

Théorème 2.2.5 $\mathcal{T}_1 = \mathcal{T}_c$.

<u>**Démonstration**</u> : Nous savons déjà que $\mathcal{T}_1 \subset \mathcal{T}_c$, donc il suffit de démontrer $\mathcal{T}_c \subset \mathcal{T}_1$.

Définissons la $\mathcal{T}_c$-mesure aléatoire

$$\begin{array}{rcl}(\mathcal{M}_c,\mathcal{T}_1) & \longrightarrow & (\mathcal{M}_c,\mathcal{T}_c) \\ \mu & \hookrightarrow & \mu\end{array}$$

Il s'agit en effet d'une $\mathcal{T}_c$-mesure aléatoire parce que l'application

$$\begin{array}{rcl}(\mathcal{M}_c,\mathcal{T}_1) & \longrightarrow & (\mathcal{M}_c,\mathcal{T}_1) \\ \mu & \hookrightarrow & \mu\end{array}$$

est une $\mathcal{T}_1$-mesure aléatoire. Or, d'après le corollaire 2.2.4 cette application est aussi une $\mathcal{T}_c$-mesure aléatoire. Alors $\mathcal{T}_c \subset \mathcal{T}_1$. ■

Nous sommes finalement arrivés à la conclusion que la tribu $\mathcal{T}_c$ est engendrée par la famille $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$ ou par la famille $\{\pi_{\mathrm{B}},\ \mathrm{B} \in \mathcal{B}\}$. Ce fait est remarquable car la topologie vague sur $\mathcal{M}_c$ n'a pas des propriétés de séparabilité telle que la topologie vague sur $\mathcal{M}_c^+$. La possibilité d'identifier des familles d'applications engendrant la tribu $\mathcal{T}_c$ va nous permettre d'établir les resultats d'égalité en loi et de convergence en loi correspondant aux résultats connus pour les mesures aléatoires positives.

2.3 Egalité en loi de mesures aléatoires à signe

Dans le cas des mesures aléatoires positives on démontre des conditions nécessaires et suffisantes d'égalité en loi en utilisant les projections associées à des fonctions $f \in \mathrm{C}_c(\mathrm{S})$ et à des boréliens relativement compacts ([26], th. 3.1). La démonstration est basée sur le fait que la tribu $\mathcal{T}_c^+$ est engendrée par chacune des familles d'applications $\{\pi_f,\ f \in \mathrm{C}_c(\mathrm{S})\}$ ou $\{\pi_{\mathrm{B}},\ \mathrm{B} \in \mathcal{B}\}$. Nous avons démontré que ces propriétés restent vraies dans le cas des mesures aléatoires à signe. Nous allons donc pouvoir récupérer le même résultat d'égalité en loi.

D'abord définissons les fonctions suivantes sur $\mathrm{C}_c(\mathrm{S})$.

Définition 2.3.1 *Soit ξ une mesure aléatoire.*

1. *La fonction caractéristique ou transformée de Fourier de ξ est la fonction $F_\xi(f) = \mathbf{E}(e^{i\xi f})$.*
2. *La transformée de Laplace de ξ est la fonction $L_\xi(f) = \mathbf{E}(e^{-\xi f})$.*

La transformée de Laplace est habituellement utilisée dans les caractérisations des mesures aléatoires positives. Pour les mesures aléatoires à signe nous utiliserons la transformée de Fourier. Ce choix est lié au fait que la transformée de Laplace d'une mesure à signe peut ne pas être définie parce que l'intégrale ξf peut être négative, et donc il y a le risque que l'espérance n'existe pas.

Avant de démontrer le théorème de caractérisation d'égalité en loi remarquons que, d'après les liens démontrés entre les tribus $\mathcal{T}_c$, $\mathcal{T}_1$ et $\mathcal{T}_2$ une application $\xi : (\Omega, \mathcal{F}, \mathbf{P}) \longrightarrow (\mathcal{M}_c, \mathcal{T}_c)$ est une mesure aléatoire si et seulement si ξf est une variable aléatoire réelle, pour tout $f \in \mathrm{C}_c(\mathrm{S})$, ou encore si et seulement si $\xi(\mathrm{B})$ est une variable aléatoire réelle, pour tout $\mathrm{B} \in \mathcal{B}$.

Définition 2.3.2 *Deux mesures aléatoires ξ et η sont égales en loi si les lois de probabilité sur $\mathcal{M}_c$, $\mathbf{P}\xi^{-1}$ et $\mathbf{P}\eta^{-1}$ sont égales. On écrira $\xi \overset{d}{=} \eta$.*

Théorème 2.3.3 *Soient ξ et η deux mesures aléatoires. Les conditions suivantes sont équivalentes.*

1. $\xi \overset{d}{=} \eta$;
2. $\xi f \overset{d}{=} \eta f$, $f \in \mathrm{C}_c(\mathbf{S})$,
3. $F_\xi(f) = F_\eta(f)$, $f \in \mathrm{C}_c(\mathbf{S})$;
4. $(\xi(B_1), \ldots, \xi(B_n)) \overset{d}{=} (\eta(B_1), \ldots, \eta(B_n)), n \in \mathbb{N}$, $B_1, \ldots, B_n \in \mathcal{B}$.

<u>**Démonstration**</u> : Les implications $1 \Rightarrow 2$, $1 \Rightarrow 3$, $1 \Rightarrow 4$ et $2 \Rightarrow 3$ sont évidentes. Alors, pour démontrer le théorème il suffit de démontrer $4 \Rightarrow 1$ et $3 \Rightarrow 1$.

$4 \Rightarrow 1$: L'ensemble

$$\mathcal{D} = \{\mathrm{M} \in \mathcal{T}_c : \mathbf{P}\{\xi \in \mathrm{M}\} = \mathbf{P}\{\eta \in \mathrm{M}\}\}$$

est un λ-système qui contient $\mathcal{M}_c$ et le π-système des ensembles de la forme $\{\mu \in \mathcal{M}_c : \mu(\mathrm{B}_1) \leq t_1, \ldots, \mu(\mathrm{B}_n) \leq t_n\}$, $n \in \mathbb{N}$, $\mathrm{B}_1, \ldots, \mathrm{B}_n \in \mathcal{B}$ et $t_1, \ldots, t_n \in \mathbb{R}$. Comme $\mathcal{T}_c$ est engendrée par ce π-système (car $\mathcal{T}_c = \mathcal{T}_2$) il s'ensuit $\xi \overset{d}{=} \eta$.

$3 \Rightarrow 1$: Pour tous $t_1, \ldots, t_n \in \mathbb{R}$, $f_1, \ldots, f_n \in \mathrm{C}_c(\mathbf{S})$ il se vérifie

$$\mathrm{F}_{\xi f_1, \ldots, \xi f_n}(t_1, \ldots, t_n) = \mathrm{F}_\xi\Big(\sum_{j=1}^{n} t_j f_j\Big) = \mathrm{F}_\eta\Big(\sum_{j=1}^{n} t_j f_j\Big) = \mathrm{F}_{\eta f_1, \ldots, \xi f_n}(t_1, \ldots, t_n),$$

donc $(\xi f_1, \ldots, \xi f_n) \overset{d}{=} (\eta f_1, \ldots, \eta f_n)$. Comme la tribu $\mathcal{T}_c$ est aussi engendrée par la famille d'applications $\{\pi_f,\ f \in \mathrm{C}_c(\mathbf{S})\}$ nous pouvons terminer la démonstration comme dans le cas précédent. ■

Remarquons que dans le cas des mesures aléatoires positives on utilise, dans la condition 3, la transformée de Laplace au lieu de la fonction caractéristique.

Cette caractérisation ne dépendait pas directement des propriétés topologiques de l'espace $\mathcal{M}_c$, mais seulement de l'identification d'ensembles engendrant la tribu $\mathcal{T}_c$. Pour cette raison on a pu récupérer le résultat d'égalité. On trouvera que, pour ce qui concerne les résultats directement liés aux propriétés topologiques on ne pourra récupérer les résultats que partiellement.

2.4 Convergence en loi de mesures aléatoires

Définition 2.4.1 *Etant donné une suite de mesures aléatoires $\{\xi_n\}$ et une mesure aléatoire ξ, on dira que ξ_n converge en loi vers ξ si la suite de mesures de probabilités $\mathbf{P}\xi_n^{-1}$ converge faiblement vers la mesure de probabilité $\mathbf{P}\xi^{-1}$, dans l'espace des mesures de probabilité sur $\mathcal{M}_c$. On écrira $\xi_n \xrightarrow{d} \xi$.*

Quand on étudie les caractérisations de convergence en loi on commence à s'apercevoir des différences entre l'étude des mesures aléatoires positives et les mesures aléatoires à signe. En fait, il faut tenir compte des différences en ce qui concerne la topologie vague sur $\mathcal{M}_c$. Comme on l'a remarqué dans le premier chapitre, quand on prétend identifier les mesures limites de suites vaguement convergentes on a des difficultés qui ne se posent pas dans l'étude des mesures aléatoires positives.

Avant d'établir le résultat de convergence en loi nous allons étudier des caractérisations d'équitension pour des suites de mesures aléatoires. Les propriétés d'équitension sont liées aux propriétés de compacité relative, donc utiles pour établir le résultat de convergence en loi. Remarquons que le critère de Prokhorov, corollaire 1.1.9, ne peut pas s'appliquer dans ce cas car on étudie des mesures de probabilité sur $\mathcal{M}_c$, et $\mathcal{M}_c$ n'est même pas un espace métrique, ce qui est exigé par le critère de Prokhorov. Par contre, quand on étudiait des mesures aléatoires positives on s'intéressait à des mesures de probabilité sur $\mathcal{M}_c^+$, et $\mathcal{M}_c^+$, avec la topologie vague, est un espace métrique séparable et complet ([26], A.7.7), donc on pouvait utiliser le critère de Prokhorov pour caractériser complètement les suites relativement compactes en loi. C'est en utilisant ce fait qu'on a démontré la caractérisation de compacité relative en loi suivante.

Théorème 2.4.2 (Kallenberg [26], lemme 4.5) *Une suite $\{\xi_n\}$ de mesures aléatoires positives est relativement compacte en loi si et seulement si*

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{\xi_n(B) > t\} = 0, \quad B \in \mathcal{B}.$$

Cette caractérisation peut se généraliser aux mesures aléatoires à signe, mais comme caractérisation d'équitension seulement. On ne retrouvera pas le résultat de compacité relative en loi. Néanmoins on pourra démontrer un résultat de compacité relative séquentielle en loi, ce qui nous sera suffisant pour la démonstration des caractérisations de convergence en loi.

Théorème 2.4.3 *Une suite $\{\xi_n\}$ de mesures aléatoires est équitendue si et seulement si*

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n|(B) > t\} = 0, \quad B \in \mathcal{B}.$$

<u>**Démonstration**</u> : Supposons la suite $\{\xi_n\}$ équitendue. D'après la définition, 1.1.8, pour tout $\varepsilon > 0$, il existe un compact K_ε de $\mathcal{M}_c$, tel que, pour tout $n \in \mathbb{N}$,

$$\mathbf{P}\{\xi_n \notin \mathrm{K}_\varepsilon\} < \varepsilon,$$

donc, pour tout $\mathrm{B} \in \mathcal{B}$

$$\mathbf{P}\{|\xi_n|(\mathrm{B}) > \sup_{\mu \in \mathrm{K}_\varepsilon} |\mu|(\mathrm{B})\} \leq \mathbf{P}\{\xi_n \notin \mathrm{K}_\varepsilon\} < \varepsilon.$$

Comme K_ε est un compact vague il se vérifie $\sup_{\mu \in \mathrm{K}_\varepsilon} |\mu|(\mathrm{B}) < \infty$, $\mathrm{B} \in \mathcal{B}$ (théorème 1.2.3). Alors, pour t assez grand, nous avons $t > \sup_{\mu \in \mathrm{K}_\varepsilon} |\mu|(\mathrm{B})$, donc

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n|(\mathrm{B}) > t\} < \varepsilon,$$

d'où, comme ε est arbitraire, il s'ensuit

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n|(\mathrm{B}) > t\} = 0.$$

Réciproquement, si nous supposons vérifiée la condition

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n|(\mathrm{B}) > t\} = 0, \quad \mathrm{B} \in \mathcal{B},$$

alors d'après Kallenberg [26], lemme 4.5, la suite $\{|\xi_n|\}$ est équitendue, et les ensembles $\{\xi_n^+\}$ et $\{\xi_n^-\}$ sont aussi équitendus dans $\mathcal{M}_c^+$, puisque $\xi_n^+(\mathrm{B})$, $\xi_n^-(\mathrm{B}) \leq |\xi_n|(\mathrm{B})$. Donc, étant donné $\varepsilon > 0$, il existe deux compacts vagues K_1, K_2 dans $\mathcal{M}_c^+$ tels que

$$\mathbf{P}\{\xi_n^+ \in \mathrm{K}_1\} > 1 - \frac{\varepsilon}{2}, \quad \mathbf{P}\{\xi_n^- \in \mathrm{K}_2\} > 1 - \frac{\varepsilon}{2}.$$

Alors $\mathbf{P}\{\xi_n \in \mathrm{K}_1 - \mathrm{K}_2\} > 1 - \varepsilon$, et $\mathrm{K}_1 - \mathrm{K}_2$ est un compact vague de $\mathcal{M}_c$, car l'application $d(\mu, \nu) = \mu - \nu$ est continue de $\mathcal{M}_c^+ \times \mathcal{M}_c^+$, avec la topologie produit, dans $\mathcal{M}_c$, avec la topologie vague. ∎

Corollaire 2.4.4 *Si la suite de mesures aléatoires $\{\xi_n\}$ est équitendue alors la suite est relativement séquentiellement compacte en loi.*

Démonstration : En fait, d'après le théorème précédent, si $\{\xi_n\}$ est équitendue, les suites $\{\xi_n^+\}$ et $\{\xi_n^-\}$ sont équitendues dans $\mathcal{M}_c^+$, donc relativement compactes dans $\mathcal{M}_c^+$, puisque $\mathcal{M}_c^+$ avec la topologie vague est un espace métrique séparable et complet. Alors, étant donné $\varepsilon > 0$, il existe deux compacts M_1 et M_2 dans $\mathcal{M}_c^+$ tels que, pour tout $n \in \mathbb{N}$

$$\mathbf{P}\{\xi_n^+ \notin M_1\} < \frac{\varepsilon}{2}, \quad \mathbf{P}\{\xi_n^- \notin M_2\} < \frac{\varepsilon}{2},$$

donc

$$\mathbf{P}\{(\xi_n^+, \xi_n^-) \notin M_1 \times M_2\} < \varepsilon.$$

Comme $M_1 \times M_2$ est un compact de $\mathcal{M}_c^+ \times \mathcal{M}_c^+$ avec la topologie vague produit, il s'ensuit que la suite $\{(\xi_n^+, \xi_n^-)\}$ est relativement compacte dans l'espace métrique séparable et complet $\mathcal{M}_c^+ \times \mathcal{M}_c^+$. Alors il existe une suite croissante de nombres entiers $\{n_k\}$ telle que la suite $\{(\xi_{n_k}^+, \xi_{n_k}^-)\}$ est convergente en loi. Comme l'application $d(\mu, \nu) = \mu - \nu$, définie sur l'espace $\mathcal{M}_c^+ \times \mathcal{M}_c^+$ avec la topologie vague produit, et à valeurs dans $\mathcal{M}_c$ avec la topologie vague, est continue il s'ensuit que la suite $\{\xi_{n_k}\}$ est convergente en loi. ∎

Remarquons que dans notre cadre il est plus utile d'avoir la compacité relative séquentielle que la compacité relative. En fait, comme $\mathcal{M}_c$ avec la topologie vague n'est pas un espace métrique c'est la première propriété qui nous permet d'extraire des sous-suites convergentes, ce qui nous sera utile pour démontrer le théorème de caractérisation de convergence en loi de suites de mesures aléatoires.

Théorème 2.4.5 *La suite $\{\xi_n\}$ est équitendue si et seulement si*

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n f| > t\} = 0, \quad f \in \mathrm{C}_c(\mathbf{S}).$$

Démonstration : Supposons $\{\xi_n\}$ équitendue. Alors

$$\forall_{\varepsilon>0}; \exists_{\mathrm{K}_\varepsilon \text{ compact de } \mathcal{M}_c}, \forall_{n\in\mathbb{N}} \ \mathbf{P}\{\xi_n \notin \mathrm{K}_\varepsilon\} < \varepsilon.$$

Soit $f \in \mathbf{C}_c(\mathbf{S})$, évidemment

$$\mathbf{P}\{|\xi_n f| > \sup_{\mu\in \mathrm{K}_\varepsilon} |\mu f|\} \leq \mathbf{P}\{\xi_n \notin \mathrm{K}_\varepsilon\} < \varepsilon.$$

D'après le théorème 1.2.3, $\sup_{\mu \in K_\varepsilon} |\mu f| < \infty$, donc pour t assez grand $t > \sup_{\mu \in K_\varepsilon} |\mu f|$, et alors

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n f| > t\} < \varepsilon.$$

ε étant arbitraire, il s'ensuit

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n f| > t\} = 0.$$

Supposons maintenant

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n f| > t\} = 0, \quad f \in \mathrm{C}_c(\mathbf{S}),$$

et soit $\varepsilon > 0$ fixé. D'après [30], 9.2.2

$$|\mu|(\mathrm{K}_n) = \sup\{|\mu f|, 0 \leq |f| \leq \mathbb{1}_{\mathrm{K}_n}, \ f \in \mathrm{C}_c(\mathbf{S})\}.$$

Comme $|f| \leq \mathbb{1}_{\mathrm{K}_n}$, nous pouvons supposer que $f \in \mathrm{C}(\mathrm{K}_n)$, le sous-espace de $\mathrm{C}_c(\mathbf{S})$ des fonctions à support contenu dans K_n. Le sous-espace $\mathrm{C}(\mathrm{K}_n)$ est séparable car K_n est compact. Soit $\{g_{r,n}\}_{r\in\mathbb{N}}$ un ensemble dénombrable dense dans $\mathrm{C}(\mathrm{K}_n)$. Alors, en raisonnant comme dans la démonstration du théorème 2.1.5, nous pouvons écrire

$$|\mu|(\mathrm{K}_n) = \sup_r\{|\mu g_{r,n}|, 0 \leq |g_{r,n}| \leq \mathbb{1}_{\mathrm{K}_n}\}.$$

Par hypothèse la suite $\{\xi_k g_{r,n}\}_{k\in\mathbb{N}}$ est équitendue, pour chaque $r, n \in \mathbb{N}$. Alors, il existe une constante $c_{r,n} > 0$ telle que

$$\mathbf{P}\ \{|\xi_k g_{r,n}| \leq c_{r,n}\} \geq 1 - \varepsilon 2^{-n-r}, \quad r, n \in \mathbb{N}.$$

Soit $\mathrm{M}_{r,n} = \{\mu \in \mathcal{M}_c : \ |\mu g_{r,n}| \leq c_{r,n}\}$, alors $\mathbf{P}\ \{\xi_k \in \mathrm{M}_{r,n}\} \geq 1 - \varepsilon 2^{-n-r}$. Donc, $\mathbf{P}\ \{\xi_k \in \bigcap_r \mathrm{M}_{r,n}\} \geq 1 - \varepsilon 2^{-n}$ et $\mathbf{P}\ \{\xi_k \in \bigcap_n \bigcap_r \mathrm{M}_{r,n}\} \geq 1 - \varepsilon$. Posons $\mathrm{M} = \bigcap_n \bigcap_r \mathrm{M}_{r,n}$. Vérifions maintenant que M est vaguement compact dans $\mathcal{M}_c$. En effet, M est évidemment fermé et

$$\sup_{\mu\in\mathrm{M}} |\mu g_{r,n}| < \infty,$$

donc, d'après le théorème de Banach-Steinhaus, il s'ensuit

$$\sup_{\mu\in\mathrm{M}} |\mu|(\mathrm{K}_n) < \infty.$$

Alors, quelque soit B$\in \mathcal{B}$, il existe un $n \in \mathbb{N}$ tel que B$\subset$K$_n$, d'où il s'ensuit

$$\sup_{\mu \in \mathrm{M}} |\mu(B)| \leq \sup_{\mu \in \mathrm{M}} |\mu(\mathrm{K}_n)| < \infty,$$

ce qui démontre que M est vaguement compact, d'après le théorème 1.2.3. ■

Nous sommes maintenant en condition d'établir le résultat de caractérisation de la convergence en loi de suites de mesures aléatoires.

Théorème 2.4.6 *Soit $\{\xi_n\}$ une suite de mesures aléatoires et ξ une mesure aléatoire. Les conditions suivantes sont équivalentes.*

1. $\xi_n \xrightarrow{d} \xi$;
2. $\xi_n f \xrightarrow{d} \xi f$, $f \in \mathrm{C}_c(\mathbf{S})$;
3. $F_{\xi_n}(f) \xrightarrow{d} F_\xi(f)$, $f \in \mathrm{C}_c(\mathbf{S})$.

<u>**Démonstration**</u> : Les conditions 2 et 3 sont évidemment équivalentes car il s'agit de la convergence en loi de variables aléatoires réelles. L'implication 1$\Rightarrow$2 est évidente car les applications $\pi_f(\mu) = \mu f$, $f \in \mathrm{C}_c(\mathbf{S})$, sont continues par rapport à la topologie vague. Il reste à démontrer que 2$\Rightarrow$1.

Supposons donc vérifiée la condition 2, alors, pour tout $f \in \mathrm{C}_c(\mathbf{S})$, la suite $\{\xi_n f\}$ est équitendue, donc vérifie la condition

$$\lim_{t\to\infty} \limsup_{n\to\infty} \mathbf{P}\{|\xi_n f| > t\} = 0.$$

Donc, d'après le théorème 2.4.5, la suite $\{\xi_n\}$ est équitendue, et, d'après le corollaire 2.4.4, $\{\xi_n\}$ est relativement séquentiellement compacte. Alors de toute sous-suite $\{\xi_{n_k}\}$ nous pouvons extraire une sous-suite $\{\xi_{n_{k_j}}\}$ convergente en loi vers une mesure aléatoire η. D'après l'implication 1$\Rightarrow$2 il s'ensuit $\xi_n f \xrightarrow{d} \eta f$, $f \in \mathrm{C}_c(\mathbf{S})$. Donc, comme nous supposons la condition 2, nous avons $\eta f \stackrel{d}{=} \xi f$, $f \in \mathrm{C}_c(\mathbf{S})$. Alors, d'après le théorème d'égalité en loi de mesures aléatoires, 2.3.3, on déduit $\eta \stackrel{d}{=} \xi$. Donc $\xi_n \xrightarrow{d} \xi$. ■

Remarquons que cet énoncé est seulement une récupération partielle des caractérisations possibles pour la convergence en loi de suites de mesures aléatoires non négatives. En effet, pour les mesures aléatoires non négatives nous avons encore l'équivalence avec une condition utilisant les masses de certains boréliens.

Théorème 2.4.7 (Kallenberg [26], th. 4.2) *Soit $\{\xi_n\}$ une suite de mesures aléatoires non négatives et ξ une mesure aléatoire non négative. Les conditions suivantes sont équivalentes.*

1. $\xi_n \xrightarrow{d} \xi$;
2. $\xi_n f \xrightarrow{d} \xi f$, $f \in \mathrm{C}_c(\mathrm{S})$;
3. $L_{\xi_n}(f) \xrightarrow{d} L_\xi(f)$, $f \in \mathrm{C}_c(\mathrm{S})$,
4. $(\xi_n(B_1), \ldots, \xi_n(B_n)) \xrightarrow{d} (\eta_n(B_1), \ldots, \eta_n(B_n))$, $k \in \mathbb{N}$, $B_1, \ldots, B_n \in \{B \in \mathcal{B} : \xi(fr(B)) = 0 \ p.s.\}$.

La possibilité de démontrer l'équivalence avec la condition 4 vient du fait que la convergence des masses de certains boréliens implique la convergence vague, ce qui ne se vérifie pas quand on étudie la topologie vague sur $\mathcal{M}_c$.

2.5 Lois marginales d'une mesure aléatoire

Une des conditions du théorème 2.3.3 caractérise la loi d'une mesure aléatoire en utilisant les lois marginales de dimension finie de la mesure aléatoire, soit, en utilisant les lois des vecteurs aléatoires $(\xi_n(\mathrm{B}_1), \ldots, \xi_n(\mathrm{B}_n)), n \in \mathbb{N}$, $\mathrm{B}_1, \ldots, \mathrm{B}_n \in \mathcal{B}$. Alors il est naturel de se demander si, étant donné les lois des vecteurs aléatoires, nous pouvons construire une mesure aléatoire ayant comme lois de dimension finie les lois données. Evidemment, la question posée comme ci-dessus, la réponse est négative, car il faut au moins que le système de lois données soit compatible, au sens du théorème de Kolmogorov. Il est évident que l'hypothèse de compatibilité au sens du théorème de Kolmogorov d'un système de lois de probabilités indexé par $\mathcal{B}$, $\{\mathbf{P}_{\mathrm{B}_1,\ldots,\mathrm{B}_k}\}$, où $\mathbf{P}_{\mathrm{B}_1,\ldots,\mathrm{B}_k}$ est une loi de probabilité sur $\mathbb{R}^k$, n'est pas suffisante pour qu'on puisse construire une mesure aléatoire ξ telle que ses lois marginales $\mathbf{P}(\xi_n(\mathrm{B}_1), \ldots, \xi_n(\mathrm{B}_k))^{-1} = \mathbf{P}_{\mathrm{B}_1,\ldots,\mathrm{B}_k}$, $k \in \mathbb{N}$, $\mathrm{B}_1, \ldots, \mathrm{B}_k \in \mathcal{B}$. Le théorème de Kolmogorov nous assure de la possibilité de construire un processus stochastique indexé par $\mathcal{B}$ avec les lois marginales donnés. Evidemment, pour que le processus soit une mesure aléatoire il faut que le processus soit additif et σ-additif. Donc il nous faudra supposer des conditions supplementaires pour assurer l'additivité et la σ-additivité du processus. Aussi, pour ce problème les conditions à imposer ne sont pas les mêmes quand on considère mesures aléatoires à signe ou seulement les mesures aléatoires non négatives. La caractérisation complète du système des lois marginales d'une mesure aléatoire non négative a été démontré par Nawrotzki [33] et Harris [17], [18].

Théorème 2.5.1 *Soit* $\mathbf{P}_{B_1,\dots,B_k}$, $k \in \mathbb{N}$, $B_1,\dots,B_k \in \mathcal{B}$ *un système compatible de lois de probabilité sur* $(0,\infty)^k$. *Il existe une mesure aléatoire non négative* ξ *telle que* $\mathbf{P}(\xi_n(B_1),\dots,\xi_n(B_k))^{-1} = \mathbf{P}_{B_1,\dots,B_k}$, $k \in \mathbb{N}$, $B_1,\dots,B_k \in \mathcal{B}$ *si et seulement si*

1. $\mathbf{P}_{B,C,B\cup C}\ \{(x,y,z):\ z = x+y\} = 1$, *pour tous* $B,C \in \mathcal{B}$ *disjoints;*
2. $\mathbf{P}_{B_n} \xrightarrow{w} \delta_0$ *pour toutes les suites telles que* $B_n \downarrow \emptyset$.

La démonstration de ce théorème utilise le fait que la mesure aléatoire est non négative. En effet il n'est pas possible de démontrer que les conditions du théorème précédent sont suffisantes car la condition 2 est dans une forme trop faible quand on travaille avec les mesures aléatoires à signe. Dans ce dernier cas il nous faut avoir un contrôle sur la variation totale de la mesure et non pas seulement sur la mesure elle même. Evidemment, quand on s'intéresse seulement aux mesures aléatoires non négatives ce problème ne se pose pas car la mesure est égale à sa variation totale. Remarquons encore que c'est la condition 2 qui nous assure la σ-additivité de la mesure ξ. C'est donc cette condition que nous aurons besoin de modifier.

Dans l'espace $\mathbf{S}$ considérons une base dénombrable pour la topologie $\mathcal{D}$, que nous supposerons fermée pour les unions et intersections finies, et $\mathcal{D} \subset \mathcal{B}$. Soit $\mathcal{A}$ l'anneau engendré par $\mathcal{D}$. L'anneau $\mathcal{A}$ est aussi dénombrable et $\mathcal{A} \subset \mathcal{B}$. Etant donné un système de lois de probabilité $\{\mathbf{P}_{B_1,\dots,B_k}\}$, $k \in \mathbb{N}$, $B_1,\dots,B_k \in \mathcal{A}$, introduisons les conditions

(M1) $\forall_{A,B\in\mathcal{A},\,A\cap B=\emptyset}\quad \mathbf{P}_{A,B,A\cup B}\ \{(x,y,z):\ z = x+y\} = 1$;

(M2) $\forall_{\{A_n\}\subset\mathcal{A},\,A_n\downarrow\emptyset},\ \forall_{t>0}\ \lim_{n\to\infty} \mathbf{P}\{\sup_{A\subset A_n} |\xi(\omega)(A)| \in [0,t]\ \} = 1$.

Il est évident que si ξ est une mesure aléatoire non négative la condition (M2) se réduit à la condition 2 du théorème précédent.

Remarquons que si le système de lois de probabilité a été construit à partir d'une mesure aléatoire, il est évidemment compatible. De plus, le système vérifie la condition (M1) parce que la mesure ξ est additive, et la condition (M2) parce que la variation totale de la mesure est σ-additive. Nous allons démontrer que les conditions (M1) et (M2) sont aussi suffisantes pour l'existence d'une mesure aléatoire.

Théorème 2.5.2 *Supposons* $\mathbf{S}$ *compact et* $\mathbf{S} \in \mathcal{D}$. *Etant donné un système de lois de probabilité* $\{\mathbf{P}_{B_1,\ldots,B_k}\}$, $k \in \mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{A}$, *compatible au sens du théorème de Kolmogorov, et qui vérifie les conditions (M1) et (M2), il existe une mesure aléatoire* ξ *telle que* $\mathbf{P}(\xi_n(B_1),\ldots,\xi_n(B_k))^{-1} = \mathbf{P}_{B_1,\ldots,B_k}$, $k \in \mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{A}$.

<u>**Démonstration**</u> : Cette démonstration suit le plan de la démonstration de la proposition 1.3 de Jagers [25]. Soit $\mathrm{G} \subset \mathbf{S}$ un ouvert. Alors le complémentaire G^c de G est compact. Comme $\mathbf{S}$ est de Hausdorff, pour chaque $x \in \mathrm{G}$ il existe deux ensembles disjoints $U_x, V_x \in \mathcal{D}$ tels que $x \in U_x$ et $\mathrm{G}^c \subset V_x$ ([27], page 140). Alors

$$\bigcup_{x\in \mathrm{G}} V_x^c \subset \mathrm{G} \subset \bigcup_{x\in \mathrm{G}} U_x \subset \bigcup_{x\in \mathrm{G}} V_x^c,$$

donc

$$\mathrm{G} = \bigcup_{x\in \mathrm{G}} V_x^c.$$

Comme chaque $V_x \in \mathcal{D}$, et $\mathcal{D}$ est une classe dénombrable, dans l'union il existe seulement une quantité dénombrable d'ensembles V_x^c. Alors, on peut écrire

$$\mathrm{G} = \bigcup_{n=1}^{\infty} \mathrm{G}_n$$

où $\mathrm{G}_n \in \mathcal{A}$ et G_n est fermé, pour tout $n \in \mathbb{N}$. Si G est la différence propre de deux éléments de $\mathcal{D}$, $\mathrm{G}=\mathrm{G}_1 \setminus \mathrm{G}_2$, on obtient la même représentation en posant $\mathrm{G}_n = \mathrm{G}_{1,n} \cap \mathrm{G}_2^c$. Les ensembles G_n ainsi obtenus sont dans l'algèbre $\mathcal{A}$ et sont fermés. L'algèbre $\mathcal{A}$ étant fermée pour les unions finies des différences propres d'ensembles de $\mathcal{D}$, vérifie la propriété

$$\forall_{\mathrm{A}\in\mathcal{A}}\ \exists_{\{\mathrm{A}_n\}\subset\mathcal{A},\mathrm{A}_n \text{fermé}} : \ \mathrm{A} = \bigcup_{n=1}^{\infty} \mathrm{A}_n.$$

De plus, $\mathcal{A}$ étant une algèbre nous pouvons choisir la suite $\{\mathrm{A}_n\}$ croissante. Nous avons donc associé, à chaque élément de l'algèbre $\mathcal{A}$, une suite croissante d'éléments de $\mathcal{A}$.

D'après la compatibilité du système de lois de probabilité donné nous pouvons construire une variable aléatoire ξ à valeurs dans $\mathbb{R}^{\mathcal{A}}$. De la condition (M1) on déduit que ce processus indexé par $\mathcal{A}$ est presque sûrement concentré dans le sous-ensemble des processus indexés par $\mathcal{A}$ qui sont additifs. Tenant compte de la condition (M2) il existe $\Omega_0 \subset \Omega$, $\mathbf{P}(\Omega_0) = 1$ tel que $\omega \in \Omega_0$ implique que $\xi(\omega)$ soit additif et que l'on ait:

$$\lim_{n\to\infty} \sup_{\mathrm{B}\subset \mathrm{A}\setminus \mathrm{A}_n} |\xi(\omega)(\mathrm{B})| = 0, \quad \mathrm{A} \in \mathcal{A}. \tag{2.1}$$

En effet, remarquons que cette propriété est vraie, d'après la condition (M2), si nous considérons A$\in \mathcal{A}$ fixé. Donc, pour chaque A$\in \mathcal{A}$, il existe un sous-ensemble $\Omega_{\mathrm{A}} \subset \Omega$ tel que $\mathbf{P}(\Omega_{\mathrm{A}}) = 1$, où se vérifie la convergence (2.1). Comme l'algèbre $\mathcal{A}$ est dénombrable, en prenant l'intersection des Ω_{A}, nous obtenons l'ensemble Ω_0 avec les propriétés référées.

Nous allons maintenant démontrer que $\xi(\omega), \omega \in \Omega_0$, est en fait σ-additive. Fixons $\omega \in \Omega_0$, et des ensembles $\mathrm{A}_k \in \mathcal{A}$ décroissants tels que $|\xi(\omega)(\mathrm{A}_k)| \geq \alpha > 0$, α fixé. Alors d'après (2.1)

$$\forall_{k\in\mathbb{N}} \exists_{n_k\in\mathbb{N}} : \sup_{\mathrm{B}\subset \mathrm{A}_k\setminus \mathrm{A}_{k,n_k}} |\xi(\omega)(\mathrm{B})| < \alpha 2^{-k-1}.$$

Posons

$$\mathrm{F}_j = \bigcap_{k=1}^{j} \mathrm{A}_{k,n_k} = \mathrm{A}_j \setminus \bigcup_{k=1}^{j} (\mathrm{A}_k \setminus \mathrm{A}_{k,n_k}).$$

Pour tout $j \in \mathbb{N}$, F_j est fermé, car les A_{k,n_k} sont fermés, $\mathrm{F}_j \in \mathcal{A}$, et

$$\mathrm{F}_{j+1} = \mathrm{A}_{j+1,n_{j+1}} \cap \mathrm{F}_j \subset \mathrm{F}_j.$$

En général nous avons

$$\mathrm{A}_j \subset \mathrm{F}_j \cup \bigcup_{k=1}^{j} (\mathrm{A}_k \setminus \mathrm{A}_{k,n_k}),$$

donc

$$\begin{aligned}
|\xi(\omega)(\mathrm{A}_j)| &\leq \sup_{\mathrm{B}\subset \mathrm{F}_j \cup \bigcup_{k=1}^{j}(\mathrm{A}_k\setminus \mathrm{A}_{k,n_k})} |\xi(\omega)(\mathrm{B})| \leq \\
&\leq \sup_{\mathrm{B}\subset \mathrm{F}_j} |\xi(\omega)(\mathrm{B})| + \sup_{\mathrm{B}\subset \bigcup_{k=1}^{j}(\mathrm{A}_k\setminus \mathrm{A}_{k,n_k})} |\xi(\omega)(\mathrm{B})| \leq \\
&\leq \sup_{\mathrm{B}\subset \mathrm{F}_j} |\xi(\omega)(\mathrm{B})| + \sum_{k=1}^{j} \sup_{\mathrm{B}\subset \mathrm{A}_k\setminus \mathrm{A}_{k,n_k}} |\xi(\omega)(\mathrm{B})|,
\end{aligned}$$

d'où on déduit

$$\sup_{B\subset F_j} |\xi(\omega)(B)| \geq |\xi(\omega)(A_j)| - \sum_{k=1}^{j} \sup_{B\subset A_k\setminus A_{k,n_k}} |\xi(\omega)(B)| \geq \frac{\alpha}{2}.$$

De cette inégalité nous déduisons $F_j \neq \emptyset$, pour tout $j \in\mathbb{N}$. Comme l'espace **S** est supposé compact, il s'ensuit $\bigcap_{j\in\mathbb{N}} F_j \neq \emptyset$. Donc, comme $\bigcap_{j\in\mathbb{N}} A_j \supset \bigcap_{j\in\mathbb{N}} F_j$, il s'ensuit $\bigcap_{j\in\mathbb{N}} A_j \neq \emptyset$.

On a donc démontré que la variable aléatoire ξ est concentrée dans le sous-espace de $\mathbb{R}^{\mathcal{A}}$ des processus σ-additifs. Maintenant, en tenant compte du choix de la classe $\mathcal{D}$ (et donc de $\mathcal{A}$) nous pouvons faire l'extension de $\xi(\omega)$, $\omega \in \Omega_0$, à une mesure dans **S**. Donc la variable aléatoire ξ est en fait une mesure aléatoire. ■

Remarquons que définir une mesure aléatoire est équivalent à définir une mesure de probabilité sur $(\mathcal{M}_c,\mathcal{T}_c)$.

Pour la résolution du cas général nous utiliserons la suite croissante de compacts $\{K_n\}$, telle que $\bigcup_{n\in\mathbb{N}} K_n = \mathbf{S}$.

Théorème 2.5.3 *Soit $\mathcal{D}' \subset \mathcal{B}$ une base dénombrable pour la topologie de* **S**, *$\mathcal{D} = \mathcal{D}' \cup \{K_n\}$, et soit $\mathcal{A}$ le plus petit anneau sur* **S** *qui contienne $\mathcal{D}$. Etant donné un système compatible au sans du théorème de Kolmogorov de lois de probabilité $\{\mathbf{P}_{B_1,\ldots,B_k}\}$, $k \in\mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{A}$, vérifiant les conditions (M1) et (M2). Alors, il existe une mesure aléatoire ξ telle que $\mathbf{P}(\xi_n(B_1),\ldots,\xi_n(B_k))^{-1} = \mathbf{P}_{B_1,\ldots,B_k}$, $k \in\mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{A}$.*

<u>**Démonstration**</u> : Pour chaque $n \in\mathbb{N}$ fixé le système $\{\mathbf{P}_{B_1,\ldots,B_k}\}$, $k \in\mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{A}\cap K_n$ est compatible au sens du théorème de Kolmogorov et vérifie les conditions (M1) et (M2). Alors, d'après le théorème précédent, il existe une mesure de probabilité $\mathbf{P}_n$ sur $\mathcal{M}_c(K_n)$, l'espace des mesures de Radon sur K_n. On peut interpréter l'espace $\mathcal{M}_c(K_n)$ comme l'ensemble des restrictions à K_n d'une mesure de Radon sur **S**:

$$\mathcal{M}_c(K_n) = \{\mu_{|K_n} : \mu \in \mathcal{M}_c\}.$$

Comme K_n est compact l'ensemble $C_c(K_n)$ est l'ensemble des restrictions à K_n d'une fonction de $C_c(\mathbf{S})$.

Soit $f \in C_c(\mathbf{S})$ et $M=\{\mu \in \mathcal{M}_c : |\mu f| < \varepsilon\}$, $\varepsilon > 0$. Il existe $n \in\mathbb{N}$ tel que $\mathrm{supp}(f)\subset K_n$. Alors

$$M_n = \{\mu \in \mathcal{M}_c(K_n) : |\mu f| < \varepsilon\} = \{\nu_{|K_n} : \nu \in \mathcal{M}_c \wedge |\nu f| < \varepsilon\} = M_{|K_n}.$$

On définit alors $\mathbf{P}(\mathrm{M}) = \mathbf{P}_n(\mathrm{M}_n)$. A cause de la compatibilité du système de lois de probabilité $\{\mathbf{P}_{\mathrm{B}_1,\ldots,\mathrm{B}_k}\}$, $k \in \mathbb{N}$, $\mathrm{B}_1,\ldots,\mathrm{B}_k \in \mathcal{A}$, les quantités $\mathbf{P}(\mathrm{M})$ sont bien définies. De plus, nous pouvons étendre la définition de $\mathbf{P}$ aux ensembles obtenus par des intersections et unions finies d'ensembles de la forme de M. Donc, nous avons la fonction $\mathbf{P}$ définie sur un anneau. Alors nous pouvons faire l'extension de $\mathbf{P}$ à la tribu engendrée par cet anneau. Evidemment cette tribu n'est autre que la tribu $\mathcal{T}_c$, ce qui définit une mesure de probabilité sur $(\mathcal{M}_c,\mathcal{T}_c)$, donc il existe une mesure aléatoire avec les lois marginales données.
■

Dans les énoncés des deux théorèmes précédents nous avons supposé que les classes $\mathcal{D}$ et $\mathcal{A}$ était incluses dans $\mathcal{B}$. Cette condition n'est pas obligatoire mais, si on ne fait pas cette supposition, il faut ajouter une condition supplémentaire pour assurer que les variables aléatoires correspondantes aux boréliens bornés sont presque sûrement finies:

(M3) $\forall_{\mathrm{A}\in\mathcal{A}\cap\mathcal{B}}\ \mathbf{P}_{\mathrm{A}}(\mathbb{R}) = 1.$

Remarquons que c'est la condition (M2) qui nous permet de démontrer la σ-additivité du processus indexé par $\mathcal{A}$. Si nous ne supposons pas cette condition, mais seulement (M1), on pourrait démontrer, de façon analogue, l'existence d'une mesure aléatoire finiment additive avec les lois marginales données.

2.6 Convergence en moyenne quadratique

Avant d'étudier la convergence en moyenne quadratique de suites de mesures aléatoires nous allons démontrer deux résultats d'identification de mesures (non aléatoires) qui nous permettrons d'établir la convergence en moyenne quadratique. Pour cela introduisons la classe de fonctions

$$\mathrm{H} = \{f \otimes g(s,t) = f(s)g(t) \colon f,g \in \mathrm{C}_c(\mathrm{S})\}.$$

Evidemment $\mathrm{H} \subset \mathrm{C}_c(\mathrm{S}\times\mathrm{S})$. Ce que nous allons démontrer c'est qu'il est suffisant de s'intéresser à cette classe de fonctions pour identifier une mesure de Radon à signe sur $\mathrm{S}\times\mathrm{S}$ et aussi que cette classe détermine la convergence vague sur $\mathcal{M}_c^+(\mathrm{S}\times\mathrm{S})$, le sous-espace des mesures de Radon non négatives sur $\mathrm{S}\times\mathrm{S}$.

Théorème 2.6.1 *La classe H détermine les mesures de $\mathcal{M}_c(\mathbf{S}\times\mathbf{S})$, c'est à dire que, si $\mu \in\mathcal{M}_c(\mathbf{S}\times\mathbf{S})$ et $\mu h = 0$, pour tout $h \in H$, alors $\mu = 0$.*

<u>**Démonstration**</u> : Soit A×B un pavé ouvert borné. Comme **S** est un espace normal, il existe une suite $\{h_k\}$ d'éléments de H qui converge, en croissant, vers l'indicatrice de A×B. Ensuite, tout ouvert borné de **S**×**S** est réunion dénombrable de pavés ouverts, et la classe des pavés ouverts est stable pour les intersections finies. Alors, si μ^+ et μ^- coïncident sur H, elles coïncident aussi sur la classe des ouverts bornés. Comme μ^+ et μ^- sont régulières, il s'ensuit $\mu^+ = \mu^-$, donc $\mu = 0$. ■

Si on se restreint à $\mathcal{M}_c^+(\mathbf{S}\times\mathbf{S})$ nous n'avons pas besoin de vérifier la convergence de l'intégrale par rapport à toutes fonctions dans $\mathbf{C}_c(\mathbf{S}\times\mathbf{S})$ pour obtenir la convergence vague d'une suite de mesures.

Théorème 2.6.2 *Soit $\{\mu_n\}$ une suite dans $\mathcal{M}_c^+(\mathbf{S}\times\mathbf{S})$. Si, pour tout $h \in H$, la suite $\{\mu_n h\}$ est convergente, alors la suite $\{\mu_n\}$ converge vaguement sur $\mathcal{M}_c^+(\mathbf{S}\times\mathbf{S})$.*

<u>**Démonstration**</u> : Soit $f \in\mathbf{C}_c(\mathbf{S}\times\mathbf{S})$. Comme **S** est une réunion dénombrable de compacts, il existe un compact K de **S** tel que supp$(f)\subset$K×K. De plus, comme **S** est normal, il existe une fonction continue, g, égale à 1 sur K, nulle en dehors d'un voisinage compact de K et à valeurs dans [0,1]. Par conséquent, en posant $t = g\mathbb{1}_{\mathrm{K}}$ nous obtenons une fonction continue à support compact sur **S** telle que $|f| \leq t\otimes t$. Comme, par hypothèse, la suite $\{\mu_n(t\otimes t)\}$ est convergente, alors $\{\mu_n f\}$ est une suite bornée. Donc, d'après le théorème 1.2.3, $\{\mu_n\}$ est une suite vaguement relativement compacte. Soit μ_0 une mesure adhérente de $\{\mu_n\}$. Alors, pour tous $\varepsilon > 0$ et $h \in$H, un voisinage vague de la forme

$$\{\nu \in \mathcal{M}_c^+(\mathbf{S}\times\mathbf{S}) : \ |\nu h - \mu_0 h| < \varepsilon\}$$

contient une infinité de termes de la suite convergente $\{\mu_n h\}$. Donc

$$\mu_0 h = \lim_{n\to\infty} \mu_n h,$$

et cette égalité est vérifiée pour tout $h \in$H. D'après le théorème précédent il s'ensuit alors que μ_0 est la seule mesure adhérente de $\{\mu_n\}$, et donc $\{\mu_n\}$ converge vaguement vers μ_0. ■

Définissons la notion de convergence en moyenne quadratique de mesures aléatoires.

Définition 2.6.3 *Une suite $\{\xi_n\}$ de mesures aléatoires converge en moyenne quadratique vers la mesure aléatoire ξ si la suite des mesures*

$$\mathbf{E}\left[(\xi_n - \xi) \otimes (\xi_n - \xi)\right]$$

converge vaguement vers la mesure nulle sur $\mathbf{S}\times\mathbf{S}$.

Cette définition de convergence en moyenne quadratique est justifiée par le résultat suivant.

Théorème 2.6.4 *Soit $\{\xi_n\}$ une suite de mesures aléatoires non négatives et ξ une mesure aléatoire non négative. Les conditions suivantes sont équivalentes.*

1. $\mathbf{E}[(\xi_n - \xi) \otimes (\xi_n - \xi)] \xrightarrow{v} 0$.
2. *Pour tout $f \in \mathbf{C}_c(\mathbf{S})$, $\xi_n f \longrightarrow \xi f$ en moyenne quadratique.*
3. *Pour tout $B \in \mathcal{B}$, tel que $\xi(fr(B)) = 0$ p.s., $\xi_n(B) \longrightarrow \xi(B)$, en moyenne quadratique.*

<u>**Démonstration**</u> : 1⇒2: Pour toute fonction $f \in \mathrm{C}_c(\mathbf{S})$, $f \otimes f \in \mathrm{C}_c(\mathbf{S}\times\mathbf{S})$. D'après le théorème de Fubini

$$\mathbf{E}\left[(\xi_n - \xi) \otimes (\xi_n - \xi)\right](f \otimes f) = \mathbf{E}\left[(\xi_n f - \xi f)^2\right] \tag{2.2}$$

donc l'implication 1⇒2 s'ensuit.

2⇒1: Tenant compte de (2.2) et du théorème précédent l'implication s'ensuit.

1⇒3: Soit $\mathrm{B} \in \mathcal{B}$ tel que $\xi(\mathrm{fr(B)}) = 0$ p.s.. Alors $\xi \otimes \xi(\mathrm{fr(B\times B)}) = 0$ p.s.. Les suites $\{\mathbf{E}(\xi_n \otimes \xi_n)\}$, $\{\mathbf{E}(\xi_n \otimes \xi)\}$ et $\{\mathbf{E}(\xi \otimes \xi_n)\}$ convergent vaguement vers $\mathbf{E}(\xi \otimes \xi)$, donc $\{\mathbf{E}(\xi_n^2(\mathrm{B}))\}$ et $\{\mathbf{E}(\xi_n(\mathrm{B})\xi(\mathrm{B}))\}$ convergent vers $\mathbf{E}(\xi^2(\mathrm{B}))$. Par conséquent $\{\xi_n(\mathrm{B})\}$ converge en moyenne quadratique vers $\xi(\mathrm{B})$.

3⇒1: Pour tous $\mathrm{A,B} \in \mathcal{B}$ tels que $\xi(\mathrm{fr(A)}) = \xi(\mathrm{fr(B)}) = 0$ p.s. les suites $\{\mathbf{E}(\xi_n(\mathrm{A})\xi_n(\mathrm{B}))\}$ et $\{\mathbf{E}(\xi_n(\mathrm{A})\xi(\mathrm{B}))\}$ convergent vers $\mathbf{E}(\xi(\mathrm{A})\xi(\mathrm{B}))$. Maintenant, pour terminer la démonstration on procède comme dans la démonstration du théorème 3.1 de [5]. Soit $(s,t) \in \mathbf{S}\times\mathbf{S}$ et $\varepsilon > 0$. Considérons les ensembles

$$\mathrm{D}_\delta = \{s' : d(s,s') < \delta\} \times \{t' : d(t,t') < \delta\},$$

où d est la métrique de l'espace **S**. Si $\delta \neq \delta'$ alors

$$\mathrm{fr}(\{s' : d(s,s') < \delta\}) \cap \mathrm{fr}(\{s' : d(s,s') < \delta'\}) = \emptyset$$

et

$$\mathrm{fr}(\{t' : d(t,t') < \delta\}) \cap \mathrm{fr}(\{t' : d(t,t') < \delta'\}) = \emptyset.$$

Donc, il existe au moins un valeur $\delta \in (0,\varepsilon)$ tel que l'ensemble D_δ est tel que $\xi(\mathrm{fr}(\{s' : d(s,s') < \delta\})) = 0$ p.s. et $\xi(\mathrm{fr}(\{t' : d(t,t') < \delta\})) = 0$ p.s.. En considérant sur $\mathbf{S}\times\mathbf{S}$ la métrique

$$d'((s_1,t_1),(s_2,t_2)) = \max\{d(s_1,s_2), d(t_1,t_2)\},$$

l'ensemble D_δ n'est autre que la boule de centre (s,t) et rayon δ. Donc la classe des ensembles de la forme $\mathrm{A}\times\mathrm{B}$ où $\xi(\mathrm{fr}(\mathrm{A})) = \xi(\mathrm{fr}(\mathrm{B})) = 0$ p.s. vérifie les conditions du corollaire 1, page 14 [5]. Alors, d'après les convergences démontrées il s'ensuit $\mathbf{E}(\xi_n \otimes \xi_n) \xrightarrow{v} \mathbf{E}(\xi \otimes \xi)$, $\mathbf{E}(\xi_n \otimes \xi) \xrightarrow{v} \mathbf{E}(\xi \otimes \xi)$ et $\mathbf{E}(\xi \otimes \xi_n) \xrightarrow{v} \mathbf{E}(\xi \otimes \xi)$, donc $\mathbf{E}[(\xi_n - \xi) \otimes (\xi_n - \xi)] \xrightarrow{v} 0$. ■

Il est évident que l'implication $1 \Rightarrow 2$ du théorème précédent reste vraie quand on considère des mesures aléatoires à signe. Mais l'implication réciproque devient fausse. Pour obtenir cette implication il faudra ajouter une hypothèse de compacité relative vague de la suite double $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$.

Théorème 2.6.5 *Soit $\{\xi_n\}$ une suite de mesures aléatoires et ξ une mesure aléatoire telles que:*

1. *Pour tout $f \in \mathrm{C}_c(\mathbf{S})$, $\xi_n f$ converge en moyenne quadratique vers ξf.*
2. *La suite double $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$ est vaguement relativement compacte dans $\mathcal{M}_c(\mathbf{S}\times\mathbf{S})$.*

Alors ξ_n converge en moyenne quadratique vers ξ.

Démonstration : Soit μ une mesure adhérente de $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$. En procédant comme dans la démonstration du théorème 2.6.2 on déduit que $\mu h = \lim_{n,m\to\infty} \mathbf{E}(\xi_n \otimes \xi_m)h$, pour tout $h \in \mathrm{H}$. Donc, d'après le théorème 2.6.1 il s'ensuit qu'il n'existe qu'une mesure adhérente, ce qui démontre le théorème. ■

Corollaire 2.6.6 *Soit $\{\xi_n\}$ une suite de mesures aléatoires et ξ une mesure aléatoire telles que*

1. *Pour tout $f \in \mathbf{C}_c(\mathbf{S})$, $\xi_n f$ converge en moyenne quadratique vers ξf.*
2. *Il existe une mesure aléatoire non négative γ telle que, pour tout $n \in \mathbb{N}$, $|\xi_n| \leq \gamma$.*

Alors ξ_n converge en moyenne quadratique vers ξ.

Démonstration : Pour tous $n, m \in \mathbb{N}$, $|\mathbf{E}(\xi_n \otimes \xi_m)| \leq \mathbf{E}|\xi_n \otimes \xi_m| \leq \mathbf{E}(\gamma \otimes \gamma)$. Alors, d'après le théorème 1.2.3, la suite double $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$ est vaguement relativement compacte dans $\mathcal{M}_c(\mathbf{S}\times\mathbf{S})$, donc le corollaire se déduit comme le théorème précédent. ■

Si on se restreint aux mesures aléatoires non négatives le fait de supposer la suite double $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$ convergente entraine la convergence en moyenne quadratique de la suite $\{\xi_n\}$.

Théorème 2.6.7 *Soit $\{\xi_n\}$ une suite de mesures aléatoires non négatives telles que la suite double $\{\mathbf{E}(\xi_n \otimes \xi_m)\}$ converge vaguement vers une mesure $\mu \in \mathcal{M}_c^+(\mathbf{S}\times\mathbf{S})$. Alors $\{\xi_n\}$ converge en moyenne quadratique vers une mesure aléatoire non négative ξ telle que $\mu = \mathbf{E}(\xi \otimes \xi)$.*

Démonstration : Pour tout $f \in \mathbf{C}_c(\mathbf{S})$, $\{\xi_n f\}$ est une suite de Cauchy en moyenne quadratique (d'où il s'ensuit que cette suite est convergente en moyenne quadratique), donc c'est aussi une suite de Cauchy en probabilité. Soit ρ la métrique correspondant à la topologie vague sur $\mathcal{M}_c^+$ proposée par Kallenberg [26], page 95. Comme cette métrique ne dépend que d'une quantité dénombrable de fonctions de $\mathbf{C}_c(\mathbf{S})$, il en résulte que $\{\xi_n\}$ est une suite de Cauchy en probabilité dans l'espace métrique complet $(\mathcal{M}_c^+, \rho)$. De plus, l'espace des variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbf{P})$ à valeurs dans $(\mathcal{M}_c^+, \rho)$, muni de la pseudo-distance

$$\rho'(\xi, \eta) = \inf \ \{\varepsilon > 0 : \ \mathbf{P}\{\rho(\xi, \eta) \leq \varepsilon\} \geq 1 - \varepsilon\}$$

est lui même complet. Par conséquent $\{\xi_n\}$ converge en probabilité vers une certaine mesure aléatoire non négative ξ ([46], page 94). Alors, nécessairement la limite en moyenne quadratique de $\{\xi_n f\}$ est ξf, pour tout $f \in \mathbf{C}_c(\mathbf{S})$. Alors, d'après le théorème 2.6.1 il s'ensuit $\mu = \mathbf{E}(\xi \otimes \xi)$. ■

Corollaire 2.6.8 *Soit $\{\xi_n\}$ une suite de mesures aléatoires non négatives telles que pour tout $f \in \mathbf{C}_c(\mathbf{S})$, $\{\xi_n f\}$ converge en moyenne quadratique. Alors $\{\xi_n\}$ converge en moyenne quadratique.*

Démonstration : Du fait que $\{\xi_n f\}$ converge en moyenne quadratique pour tout $f \in \mathrm{C}_c(\mathbf{S})$, il s'ensuit que la suite double $\{\mathbf{E}(\xi_n \oslash \xi_m)h\}$ converge, pour $h \in$H. Donc, d'après le théorème 2.6.2 il s'ensuit la convergence en moyenne quadratique de la suite $\{\xi_n\}$. ■

Corollaire 2.6.9 *Soit $\{\xi_n\}$ une suite de mesures aléatoires telle que*

1. *Pour tout $f \in \mathrm{C}_c(\mathbf{S})$, $\{\xi_n f\}$ converge en moyenne quadratique.*
2. *Il existe une mesure aléatoire non négative γ telle que, pour tout $n \in \mathbf{N}$, $|\xi_n| \leq \gamma$.*

Alors $\{\xi_n\}$ converge en moyenne quadratique.

Démonstration : Si $\{\xi_n f\}$ converge en moyenne quadratique, alors la mesure aléatoire non négative $\eta_n = \xi_n + \gamma$ vérifie les conditions du corollaire précédent, donc η_n converge en moyenne quadratique vers une mesure aléatoire non négative η. Donc $\{\xi_n f\}$ converge en moyenne quadratique vers $\eta f - \gamma f = \xi f$. Alors, d'après le corollaire 2.6.6 il s'ensuit que $\{\xi_n\}$ converge en moyenne quadratique vers ξ. ■

On remarquera que, pour ce type de convergence, on n'a pas réussi à démontrer de condition nécessaire et suffisante de convergence pour les mesures aléatoires à signe. De plus, ce qui posait les problèmes était toujours les caractérisations de compacité relative, comme dans l'étude de la convergence en loi.

Finalement on pourrait s'intéresser à l'étude de la convergence presque sûre. Malheureusement il semble impossible de démontrer des résultats non triviaux en utilisant des méthodes tels que celles qu'on utilise dans cet exposé.

Chapitre 3

Mesures aléatoires infiniment divisibles

Ayant établi, au cours du chapitre précédent, les résultats généraux en ce qui concerne la théorie des mesures aléatoires à signe, nous allons, dans le chapitre présent, entreprendre l'étude d'une classe particulière de mesures aléatoires: les mesures aléatoires infiniment divisibles. Le but est de trouver une représentation du type de Lévy-Khintchine pour ces mesures aléatoires. On remarquera qu'il ne semble pas possible de démontrer une représentation pour toutes les mesures aléatoires infiniment divisibles. On doit se contenter de démontrer une représentation qui correspond à la représentation de Kolmogorov pour les variables aléatoires réelles. Ce résultat semble le mieux que l'on puisse démontrer à présent, au moins en utilisant la méthode basée sur le théorème de caractérisation des lois marginales d'une mesure aléatoire, théorème 2.5.3. Comme dans le cas des mesures aléatoires non négatives on pourra démontrer un résultat plus fort concernant les mesures aléatoires à accroissements indépendants infiniment divisibles. Dans ce cas nous trouverons une représentation du type de Lévy-Khintchine des mesures à accroissements indépendants infiniment divisibles intégrables.

3.1 Le cas général

Dans cette section nous chercherons une représentation du type de Lévy-Khintchine pour la fonction caractéristique d'une mesure aléatoire infiniment divisible.

Définition 3.1.1 *Une mesure aléatoire ξ est infiniment divisible si, pour tout $n \in \mathbb{N}$, il existe des mesures aléatoires indépendantes et de même loi $\xi_1, \ldots, \xi_n$ telles que $\xi \overset{d}{=} \xi_1 + \cdots + \xi_n$.*

Il est immédiat, de cette définition, que si ξ est une mesure aléatoire infiniment divisible, les vecteurs aléatoires $(\xi(\mathrm{B}_1), \ldots, \xi(\mathrm{B}_k))$, $k \in \mathbb{N}$, $\mathrm{B}_1, \ldots, \mathrm{B}_k \in \mathcal{B}$, sont aussi infiniment divisibles.

Le problème de la représentation d'une mesure aléatoire infiniment divisible, pour le cas des mesures aléatoires non négatives, a été résolu en utilisant la transformée de Laplace. Il est possible, dans ce cas, de trouver une représentation de toutes mesures aléatoires infiniment divisibles.

Théorème 3.1.2 (Kallenberg [26], th. 6.1) *Une mesure aléatoire non négative, ξ, est infiniment divisible si et seulement si sa transformée de Laplace s'écrit*

$$\mathbf{E}(e^{-\xi f}) = \exp\left(-\alpha f - \int 1 - e^{-\mu f} \lambda(d\mu)\right),$$

où $\alpha \in \mathcal{M}_c^+$ et λ est une mesure sur $\mathcal{M}_c^+ \setminus \{0\}$ telle que

$$\int 1 - e^{-\mu(\mathrm{B})} \lambda(d\mu) < \infty, \quad B \in \mathcal{B}.$$

Dans le cas général des mesures aléatoires à signe il ne semble pas possible trouver une représentation de la fonction caractéristique de toutes les mesures aléatoires infiniment divisibles. En effet nous allons démontrer une représentation en imposant une condition sur les mesures aléatoires correspondante aux conditions de second ordre pour les variables aléatoires.

Théorème 3.1.3 *Soit ξ une mesure aléatoire infiniment divisible telle que $\mathbf{E}(\xi \otimes \xi) \in \mathcal{M}_c(\mathbf{S} \times \mathbf{S})$. Alors il existe une mesure finiment additive λ sur $\mathcal{M}_c$ vérifiant*

$$\int_{\mathbb{R}^k} ||x||^2 \lambda \pi^{-1}_{\mathrm{B}_1, \ldots, \mathrm{B}_k}(dx) < \infty, \quad k \in \mathbb{N}, \ B_1, \ldots, B_k \in \mathcal{B},$$

et une mesure $\alpha \in \mathcal{M}_c$, telles que

$$\mathbf{E}(e^{i\xi f}) = \exp\left[i\alpha f + \int_{\mathcal{M}_c} e^{i\mu f} - 1 - i\mu f \ \lambda(d\mu)\right], \quad f \in \mathbf{C}_c(\mathbf{S}).$$

Démonstration : De l'hypothèse $\mathbf{E}(\xi \otimes \xi) \in \mathcal{M}_c(\mathbf{S}\times\mathbf{S})$, nous déduisons, pour tout $\mathrm{B} \in \mathcal{B}$, $\mathbf{E}(\xi(\mathrm{B}))^2 = \mathbf{E}(\xi \otimes \xi)(\mathrm{B}\times\mathrm{B}) < \infty$, et de même pour les vecteurs aléatoires $(\xi(\mathrm{B}_1), \ldots, \xi(\mathrm{B}_k))$, $\mathrm{B}_1, \ldots, \mathrm{B}_k \in \mathcal{B}$, on déduit

$$\mathbf{E}||(\xi(\mathrm{B}_1), \ldots, \xi(\mathrm{B}_k))||^2 < \infty,$$

où $||\cdot||$ représente la norme euclidienne de $\mathbb{R}^k$. Il s'ensuit que ξ est une mesure aléatoire intégrable. Par commodité, dans ce qui suit nous supposerons $\mathbf{E}(\xi) = 0$. Alors, d'après la fórmule de représentation de Kolmogorov, il existe une mesure non négative finie sur $\mathbb{R}^k$, $\gamma_{\mathrm{B}_1,\ldots,\mathrm{B}_k}$, telle que

$$\mathbf{E}(e^{it\cdot(\xi(\mathrm{B}_1),\ldots,\xi(\mathrm{B}_k))}) = \exp\left[\int_{\mathbb{R}^k} \frac{e^{it\cdot x} - 1 - it\cdot x}{||x||^2}\, \gamma_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx)\right],\ t \in \mathbb{R},$$

où $t\cdot x, t, x \in \mathbb{R}$, représente le produit scalaire de t par x. Posons

$$\lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx) = \frac{1}{||x||^2}\gamma_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx).$$

Alors, la fonction caractéristique du vecteur aléatoire $(\xi(\mathrm{B}_1), \ldots, \xi(\mathrm{B}_k))$ est de la forme

$$\mathbf{E}(e^{it\cdot(\xi(\mathrm{B}_1),\ldots,\xi(\mathrm{B}_k))}) = \exp\left[\int_{\mathbb{R}^k} e^{it\cdot x} - 1 - itx\ \lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx)\right], t \in \mathbb{R} \tag{3.1}$$

et les mesures $\lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}$ vérifient

$$\int_{\mathbb{R}^k} ||x||^2\ \lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx) < \infty. \tag{3.2}$$

Si dans (3.1) nous prenons $t \in \mathbb{R}$ tel que toutes les composantes sont égales à 1 excepté une qui sera égale à 0, tenant compte de la forme de la fonction qu'on intègre, nous déduisons que le système des mesures $\lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}$ est compatible au sens du théorème de Kolmogorov. En effet, supposons $t_j = 1$, $j = 1, \ldots, k-1$, et $t_k = 0$, alors

$$\mathbf{E}(e^{i(t_1,\ldots,t_k)\cdot(\xi(\mathrm{B}_1),\ldots,\xi(\mathrm{B}_k))}) = \mathbf{E}(e^{i(t_1,\ldots,t_{k-1})\cdot(\xi(\mathrm{B}_1),\ldots,\xi(\mathrm{B}_{k-1}))})$$

ce qui est équivalent à

$$\int_{\mathbb{R}^k} e^{i\sum_{j=1}^{k-1} t_j x_j} - 1 - i\sum_{j=1}^{k-1} t_j x_j\ \lambda_{\mathrm{B}_1,\ldots,\mathrm{B}_k}(dx_1, \ldots, dx_k) = \tag{3.3}$$

$$= \int_{\mathbb{R}^{k-1}} e^{i\sum_{j=1}^{k-1} t_j x_j} - 1 - i\sum_{j=1}^{k-1} t_j x_j \, \lambda_{B_1,\ldots,B_{k-1}}(dx_1,\ldots,dx_{k-1}). \qquad (3.4)$$

Comme la mesure $\lambda_{B_1,\ldots,B_{k-1}}$ est de Lévy, d'après [2], 1.1, il existe une suite de mesures non négatives finies sur $\mathbb{R}^{k-1}$, $\{\nu_n\}$, telle que pour tout A$\in \mathcal{B}(\mathbb{R}^{k-1})$, $\nu_n(A) \uparrow \lambda_{B_1,\ldots,B_{k-1}}(A)$ et une suite $\{x_n\}$ dans $\mathbb{R}^{k-1}$ telle que la suite de mesures

$$\delta_{x_n} * \sum_{m=0}^{\infty} e^{-\nu_n(\mathbb{R}^{k-1})}\frac{\nu_n^m}{m!} (= \delta_{x_n} * \ Pois\ \nu_n, \text{par définition}),$$

où ν_n^m représente le produit de convolution $\nu_n * \cdots * \nu_n$ (m facteurs) et $\nu_n^0 = \delta_0$, converge faiblement vers une mesure finie ν, que nous appelerons $t - Pois\ \lambda_{B_1,\ldots,B_{k-1}}$ (t à cause des translations). Dans le cas présent nous pouvons choisir

$$x_n = -\int_{\mathbb{R}^{k-1}} x \ \nu_n(dx).$$

Alors, il est facile de vérifier que l'exponentielle de l'intégrale de l'expression (3.4) représente la fonction caractéristique de ν. Dans l'intégrale (3.3) nous intégrons en fait relativement à la mesure $\lambda_{B_1,\ldots,B_k}\pi^{-1}_{B_1,\ldots,B_{k-1}}$, la projection de $\lambda_{B_1,\ldots,B_k}$ sur les premières $k-1$ coordonées. Cette projection est une mesure de Lévy sur $\mathbb{R}^{k-1}$, donc comme dans le cas précedent il existe une suite de mesures non négatives finies sur $\mathbb{R}^{k-1}$, $\{\mu_n\}$, telle que pour tout A$\in \mathcal{B}(\mathbb{R}^{k-1})$, $\mu_n(A) \uparrow \lambda_{B_1,\ldots,B_k}\pi^{-1}_{B_1,\ldots,B_{k-1}}(A)$ et une suite $\{y_n\}$ dans $\mathbb{R}^{k-1}$ telle que la suite de mesures

$$\delta_{y_n} * \sum_{m=0}^{\infty} e^{-\mu_n(\mathbb{R}^{k-1})}\frac{\mu_n^m}{m!},$$

converge faiblement vers une mesure finie $\mu = t - Pois\ \lambda_{B_1,\ldots,B_k}\pi^{-1}_{B_1,\ldots,B_{k-1}}$. En choisissant

$$y_n = -\int_{\mathbb{R}^{k-1}} x \ \mu_n(dx).$$

on vérifie que l'exponentielle de l'intégrale de l'expression (3.3) représente la fonction caractéristique de $\mu = t - Pois\ \lambda_{B_1,\ldots,B_k}\pi^{-1}_{B_1,\ldots,B_{k-1}}$. Alors, nous avons bien la compatibilité du système de mesures $\{\lambda_{B_1,\ldots,B_k}\}$, $k \in \mathbb{N}$, $B_1,\ldots,B_k \in \mathcal{B}$.

Considérons maintenant deux boréliens relativement compacts A,B$\in \mathcal{B}$, tels que A$\cap$B $= \emptyset$. Alors, d'après l'additivité de ξ, nous avons

$$\mathbf{E}(e^{s\xi(A)+t\xi(B)+u\xi(A\cup B)}) = \mathbf{E}(e^{(s+u)\xi(A)+(t+u)\xi(B)}).$$

D'autre part

$$\mathbf{E}\left(e^{s\xi(A)+t\xi(B)+u\xi(A\cup B)}\right) =$$
$$\exp\left[\int_{\mathbb{R}^3} e^{i(sx_1+tx_2+ux_3)} - 1 - i(sx_1 + tx_2 + ux_3)\ \lambda_{A,B,A\cup B}(dx_1dx_2dx_3)\right]$$

et

$$\mathbf{E}(e^{(s+u)\xi(A)+(t+u)\xi(B)}) =$$
$$\exp\left[\int_{\mathbb{R}^2} e^{i((s+u)x_1+(t+u)x_2)} - 1 - i((s+u)x_1 + (t+u)x_2)\ \lambda_{A,B}(dx_1dx_2)\right],$$

d'où nous déduisons l'égalité des intégrales de ces deux expressions, donc, comme il s'agit de nombres complexes, l'égalité des parties réelles de ces intégrales:

$$\int_{\mathbb{R}^3} \cos(sx_1 + tx_2 + ux_3) - 1\ \lambda_{A,B,A\cup B}(dx_1dx_2dx_3) =$$
$$\int_{\mathbb{R}^2} \cos[(s+u)x_1 + (t+u)x_2] - 1\ \lambda_{A,B}(dx_1dx_2).$$

De la compatibilité du système $\{\lambda_{B_1,\dots,B_k}\}$, nous déduisons l'égalité

$$\int_{\{(x_1,x_2x_3):\,x_3=x_1+x_2\}} \cos(sx_1 + tx_2 + ux_3) - 1\ \lambda_{A,B,A\cup B}(dx_1dx_2dx_3) =$$
$$\int_{\mathbb{R}^2} \cos[(s+u)x_1 + (t+u)x_2] - 1\ \lambda_{A,B}(dx_1dx_2),$$

donc

$$\int_{\{x_3\neq x_1+x_2\}} \cos(sx_1 + tx_2 + ux_3) - 1\ \lambda_{A,B,A\cup B}(dx_1dx_2dx_3) = 0, \quad s,t,u \in \mathbb{R}.$$

Fixons $s, t, u \in\mathbb{R}$. La fonction qu'on intègre vérifie $\cos(sx_1+tx_2+ux_3)-1 \leq 0$, alors si on considère l'intégrale seulement sur l'ensemble où la fonction est strictement négative, comme la mesure $\lambda_{A,B,A\cup B}$ est non négative, on peut conclure que la mesure de cet ensemble d'intégration est nulle. Cela veut dire

$$\lambda_{A,B,A\cup B}(\{x_3 \neq x_1 + x_2\} \cap \{sx_1 + tx_2 + ux_3 \neq 2k\pi, k \in \mathbb{Z}\}) = 0, \quad s,t,u, \in \mathbb{R}.$$

Il nous faut choisir un nombre fini de triplets (s,t,u) de telle façon que l'union des ensembles $\{x_3 \neq x_1 + x_2\} \cap \{sx_1 + tx_2 + ux_3 \neq 2k\pi, k \in \mathbb{Z}\}$, pour ces triplets soit l'ensemble $\{x_3 \neq x_1 + x_2\}$, pour conclure $\lambda_{A,B,A\cup B}(\{x_3 \neq x_1 + x_2\}) = 0$. Pour cela il suffit de choisir (s_i, t_i, u_i), $i = 1, 2, 3, 4$, tels que le système

$$\begin{cases} s_1x_1 + t_1x_2 + u_1x_3 = 2k_1\pi \\ s_2x_1 + t_2x_2 + u_2x_3 = 2k_2\pi \\ s_3x_1 + t_3x_2 + u_3x_3 = 2k_3\pi \\ s_4x_1 + t_4x_2 + u_4x_3 = 2k_4\pi \end{cases} \tag{3.5}$$

soit impossible pour tout choix des nombres k_1, k_2, k_3, k_4 dans $\mathbb{N}$. Pour que le système soit possible il suffit que

$$-\begin{vmatrix} s_2 & t_2 & u_2 \\ s_3 & t_3 & u_3 \\ s_4 & t_4 & u_4 \end{vmatrix} k_1 + \begin{vmatrix} s_1 & t_1 & u_1 \\ s_3 & t_3 & u_3 \\ s_4 & t_4 & u_4 \end{vmatrix} k_2 - \begin{vmatrix} s_1 & t_1 & u_1 \\ s_2 & t_2 & u_2 \\ s_4 & t_4 & u_4 \end{vmatrix} k_3 + \begin{vmatrix} s_1 & t_1 & u_1 \\ s_2 & t_2 & u_2 \\ s_3 & t_3 & u_3 \end{vmatrix} k_4 = 0 \tag{3.6}$$

et le coefficient de k_4 dans cette équation soit non nulle. L'équation peut s'écrire sous la forme

$$ak_1 + bk_2 + ck_3 + dk_4 = 0.$$

Si l'équation (3.6) n'a pas de solutions $(k_1, k_2, k_3, k_4) \in \mathbb{N}^4$ le système (3.5) est impossible. Cela équivaut à dire que les nombres a, b, c, d sont linéairement indépendants par rapport à $\mathbb{N}$. Donc, il suffit de choisir quatre réels linéairement indépendants par rapport à $\mathbb{N}$ (l'existence ne pose pas de problèmes: en effet les algébristes savent que l'on peut choisir n'importe quel nombre de réels algébriquement indépendants, voir Lang [28], ce qui est une notion encore plus forte que celle dont nous avons besoin; le seul problème c'est que les algébristes ne connaissent pas d'exemple), et de poser

$$\begin{bmatrix} s_1 & t_1 & u_1 \\ s_2 & t_2 & u_2 \\ s_3 & t_3 & u_3 \\ s_4 & t_4 & u_4 \end{bmatrix} = \begin{bmatrix} b & 0 & 0 \\ -a & -\frac{c}{b} & 0 \\ 0 & 1 & -\frac{d}{c} \\ 0 & 0 & 1 \end{bmatrix}.$$

Nous avons donc vérifié que le système $\{\lambda_{B_1,\ldots,B_k}\}$ satisfait la condition (M1). Nous avons vérifié la condition (M1) en utilisant le complémentaire de l'ensemble $\{x_3 = x_1 + x_2\}$ parce que nous ne savons pas si les mesures $\lambda_{B_1,\ldots,B_k}$ sont finies, donc nous avons dû d'utiliser les complémentaires. Pour ce qui concerne la condition (M2) il ne semble pas possible, en général, de vérifier la validité de cette condition pour le système $\{\lambda_{B_1,\ldots,B_k}\}$. C'est pour cette raison que la représentation de la fonction caractéristique est obtenue en fonction d'une mesure finiment additive sur $\mathcal{M}_c$ et pas en fonction d'une vraie mesure sur $\mathcal{M}_c$. Cependant nous ne pouvons pas encore justifier l'existence d'une mesure finiment additive sur $\mathcal{M}_c$ par les théorèmes de la section précédente. En fait, pour utiliser ces théorèmes il faudrait que les mesures $\lambda_{B_1,\ldots,B_k}$ soient finies, ce qui n'est pas forcément vrai. Le seul fait que nous pouvons utiliser ce sont les conditions (3.2).

Rappelons maintenant l'existence de la suite croissante de compacts, $\{K_n\}$, dans S, et fixons $n \in \mathbb{N}$. Alors, pour $B_1, \ldots, B_k \subset K_n$, définissons

$$\lambda'_{n,B_1,\ldots,B_k}(A) = \int_{\mathbb{R}\times A} x^2\, \lambda_{K_n,B_1,\ldots,B_k}(dx dx_1 \cdots dx_k), \quad A \in \mathcal{B}(\mathbb{R}^k).$$

D'après (3.2), $\lambda'_{n,B_1,\ldots,B_k}(\mathbb{R}) < \infty$. Evidemment, pour n fixé, le système $\{\lambda'_{B_1,\ldots,B_k}\}$, $k \in \mathbb{N}$, $B_1, \ldots, B_k \in \mathcal{B} \cap K_n$, vérifie les mêmes propriétés que le système $\{\lambda_{n,B_1,\ldots,B_k}\}$. Alors, comme on l'a remarqué à la fin de la section précédente nous pouvons construire une mesure finiment additive λ'_n sur $\mathcal{M}_c(K_n)$. En posant

$$\lambda_n(d\mu) = \frac{1}{(\mu(K_n))^2}\lambda'_n d(\mu)$$

on définit une mesure finiment additive qui a ses projections de dimension finies égales au système donné $\{\lambda_{B_1,\ldots,B_k}\}$:

$$\begin{aligned}
\lambda_n \pi^{-1}_{B_1,\ldots,B_k}(A) &= \int_{\{(\mu(B_1),\ldots,\mu(B_k))\in A\}} \frac{1}{(\mu(K_n))^2}\, \lambda'_n d(\mu) = \\
&= \int_{\mathbb{R}\times A} \frac{1}{x^2}\, \lambda'_{n,K_n,B_1,\ldots,B_k}(dx dx_1 \cdots dx_k) = \\
&= \int_{\mathbb{R}\times\mathbb{R}\times A} y^2 \frac{1}{x^2}\, \lambda_{K_n,K_n,B_1,\ldots,B_k}(dy dx dx_1 \cdots dx_k) = \\
&= \int_{\mathbb{R}\times A} \lambda_{K_n,B_1,\ldots,B_k}(dx dx_1 \cdots dx_k) = \\
&= \lambda_{B_1,\ldots,B_k}(A),
\end{aligned}$$

du fait que $\lambda_{K_n,K_n,B_1,\ldots,B_k}$ ne charge que le sous-ensemble de $\mathbb{R}\times\mathbb{R}\times\mathbb{R}^k$ où les deux premières coordonnées sont égales, et de la projectivité. Maintenant, on a un système de mesures finiment additives $\{\lambda_n\}$ et nous pouvons construire une mesure finiment additive λ sur $\mathcal{M}_c$ en procédant comme dans la démonstration du théorème 2.5.3. Maintenant l'introduction du terme $i\alpha f$ dans l'expression de la fonction caractéristique se fait tenant compte du centrage que nous avons supposé. ■

Remarquons que dans (3.1) la variable par rapport à laquelle on intègre n'intervient que dans des produits scalaires avec la variable t. Donc, même s'il faut tenir compte d'une exponentielle la fonction dépend de la variable d'intégration x d'une façon linéaire. C'est ce fait qui nous a permis de démontrer que le système de mesures $\{\lambda_{B_1,\dots,B_k}\}$ est compatible au sens du théorème de Kolmogorov. Maintenant, quand on essaye d'étudier le cas général, soit en enlevant l'hypothèse $\mathbf{E}(\xi \otimes \xi) \in \mathcal{M}_c(\mathbf{S}\times\mathbf{S})$, nous n'arrivons pas à modifier la représentation de la fonction caractéristique de façon à rendre la fonction qu'on intègre dépendante seulement d'une fonction linéaire de la variable d'intégration. En effet, pour étudier le cas général, il faudrait utiliser la formule de Lévy-Khintchine pour la représentation de la fonction caractéristique des vecteurs aléatoires $(\xi(\mathrm{B}_1),\dots,\xi(\mathrm{B}_k))$. Dans la forme classique et dans les variations connues de cette formule (voir [3]) il intervient toujours un terme qui ne dépend pas d'une fonction linéaire de la variable d'intégration. Par cette raison il ne semble pas possible de trouver une représentation générale des mesures aléatoires à signe infiniment divisibles, au moins en utilisant les présentes méthodes.

3.2 Accroissements indépendants

Nous allons maintenant nous intéresser au cas particulier des mesures aléatoires infiniment divisibles à accroissements indépendants. Pour ce type de mesures aléatoires infiniment divisibles nous trouverons une représentation du type de Lévy-Khintchine en supposant seulement l'intégrabilité de la mesure aléatoire.

Définition 3.2.1 *Une mesure aléatoire ξ est à accroissements indépendants si, étant donné $B_1,\dots,B_n \in \mathcal{B}$ deux à deux disjoints, les variables aléatoires réelles $\xi(B_1),\dots,\xi(B_n)$ sont indépendantes.*

Ce type de mesures aléatoires est parfois appelé mesures complètemment aléatoires parce qu'il n'y a aucun lien entre les masses de boréliens disjoints.

Nous commençons par la référence au résultat concernant les mesures aléatoires non négatives. Comme d'habitude on caractérise la transformée de Laplace.

Théorème 3.2.2 (Kallenberg [26], th. 7.2) *Une mesure aléatoire non négative à acroissements indépendants ξ est infiniment divisible si et seulement si sa transformée de Laplace est de la forme*

$$\mathbf{E}(e^{-\xi f}) = \exp\left[-\alpha f - \int 1 - e^{-xf(s)}\,\gamma(dx, dt)\right], \quad f \in \mathbf{C}_c(\mathbf{S})$$

où, $\alpha \in \mathcal{M}_c^+$ et $\gamma \in \mathcal{M}_c^+((0,\infty)\times\mathbf{S})$ telle que

$$\int 1 - e^{x}\,\gamma(dx \times B) < \infty, \quad B \in \mathcal{B}.$$

Pour démontrer la formule de représentation dans ce cas nous utiliserons une méthode différente de celle du cas général. Introduisons une notion de mesure aléatoire, que nous appelerons mesure stochastique, plus large que celle que nous avons envisagée jusqu'à présent.

Définition 3.2.3 *Une mesure stochastique est une mesure à valeurs dans l'espace des variables aléatoires réelles,*

$$\xi : \mathcal{B} \longrightarrow L^0(\Omega, \mathcal{F}, \mathbf{P}),$$

telle que, si $B_1, \ldots, B_n \in \mathcal{B}$ sont deux à deux disjoints, les variables aléatoires réelles $\xi(B_1), \ldots, \xi(B_n)$ sont indépendantes et $\xi(\{s\}) = 0$ p.s., pour tout $s \in \mathbf{S}$.

Ce type de mesures a été étudié par nombreux auteurs, notamment Prékopa [40]. Il est évident que toute mesure aléatoire à acroissements indépendants non atomique est une mesure stochastique. La réciproque n'est pas vraie car les lois marginales peuvent être n'importe quoi, dans le cas des mesures stochastiques.

Théorème 3.2.4 (Prékopa [40]) *Si ξ est une mesure stochastique alors, pour tout $B \in \mathcal{B}$*

$$\mathbf{E}(e^{it\xi(\mathbf{B})}) = \exp\left[it\alpha(B) - \frac{1}{2}t^2\sigma^2(B) + \int e^{itx} - 1 - \frac{itx}{1+x^2}\nu(dx, B)\right], \quad t \in \mathbb{R},$$

où $\alpha(B) \in \mathbb{R}$, $\sigma^2(B) > 0$, $\nu(\cdot, B)$ est une mesure de Lévy sur $\mathbb{R}$. De plus, l'application $\alpha : \mathcal{B} \longrightarrow \mathbb{R}$ est une mesure de Radon, $\sigma^2 : \mathcal{B} \longrightarrow \mathbb{R}^+$ est une mesure de Radon non négative, et $\nu(I, \cdot) : \mathcal{B} \longrightarrow \mathbb{R}^+$, $I \in \mathcal{B}(\mathbb{R})$ fixé, est une mesure de Radon non négative (donc ν est une bimesure). Les mesures α, σ^2 et $\nu(I, \cdot)$ sont non atomiques.

Remarquons que ce résultat implique que les variables aléatoires $\xi(\mathrm{B})$, $\mathrm{B}\in \mathcal{B}$ sont infiniment divisibles.

Corollaire 3.2.5 *Soit ξ une mesure stochastique et $f \in \mathrm{C}_c(\mathbf{S})$. Alors*

$$\mathbf{E}(e^{i\xi f}) = \exp\left[i\alpha f - \frac{1}{2}\sigma^2 f^2 + \int e^{ixf(s)} - 1 - \frac{ixf(s)}{1+x^2}\nu(dx,ds)\right]. \tag{3.7}$$

__Démonstration__ : Soit $f = \sum_{j=1}^k a_j \mathbb{1}_{\mathrm{B}_j}$, avec $a_1,\ldots,a_k \in\mathbb{R}$, $\mathrm{B}_1,\ldots,\mathrm{B}_k \in \mathcal{B}$ deux à deux disjoints, une fonction simple à support compact. Alors d'après le théorème précédent

$$\begin{aligned}\mathbf{E}(e^{i\xi f}) &= \\ &= \exp\left\{\sum_{j=1}^k \left[ia_j\alpha(\mathrm{B}_j) - \frac{1}{2}a_j^2\sigma^2(\mathrm{B}_j) + \int e^{ia_jx} - 1 - \frac{ia_jx}{1+x^2}\,\nu(dx,\mathrm{B}_j)\right]\right\} = \\ &= \exp\left[i\alpha f - \frac{1}{2}\sigma^2 f^2 + \int e^{ixf(s)} - 1 - \frac{ixf(s)}{1+x^2}\,\nu(dx,ds)\right].\end{aligned}$$

Pour $f \in\mathrm{C}_c(\mathrm{S})$ positive il existe une suite croissante de fonctions simples à support compact convergeant vers f. Comme les mesures α, σ^2 et $\nu(\mathrm{I},\cdot)$ sont des mesures de Radon sur $\mathbf{S}$, la version vectorielle du théorème de la convergence dominée nous permet de conclure que la formule (3.7) se vérifie pour cette fonction f. Finalement, comme toute fonction $f \in\mathrm{C}_c(\mathbf{S})$ est la différence de deux fonctions de $\mathrm{C}_c(\mathbf{S})$ positives, il s'ensuit la formule (3.7) pour tout $f \in\mathrm{C}_c(\mathrm{S})$. ■

Cette représentation de la fonction caractéristique est évidemment trop large pour pouvoir représenter une mesure aléatoire à accroissements indépendants infiniment divisible presque sûrement non atomique. Pour obtenir une telle représentation il faut imposer des conditions sur σ^2 et ν pour assurer que la variation totale soit presque sûrement finie. Les parties de la mesure correspondant à σ^2 et ν seront appelées les parties gaussienne et poissonnienne de la mesure, respectivement.

Dans la suite nous supposerons toujours que la mesure aléatoire considérée est non atomique.

3.2.1 La partie gaussienne d'une mesure aléatoire à acroissements indépendants infiniment divisible

Soit ξ une mesure aléatoire à accroissements indépendants infiniment divisible. Le problème avec la formule (3.7) c'est que cette formule ne tient pas compte du fait que $\xi(\omega)$ est une mesure de Radon, pour chaque $\omega \in \Omega$, donc la variation totale de ξ sur chaque B$\in \mathcal{B}$ doit être presque sûrement finie. Nous allons démontrer que l'existence de la partie correspondant à $-\frac{1}{2}\sigma^2 f^2$ implique que la variation totale est presque sûrement infinie, sauf si $\sigma^2 \equiv 0$.

Nous pouvons définir une mesure stochastique ξ_0 telle que

$$\mathrm{E}(e^{i\xi_0 f}) = \exp(-\frac{1}{2}\sigma^2 f^2),$$

pour f mesurable et à support compact. Il est évident que cette application à valeurs dans $\mathrm{L}^0(\Omega, \mathcal{F}, \mathbf{P})$ est à accroissements indépendants, donc est bien une mesure stochastique. De plus, $\xi_0(\mathrm{B}), \mathrm{B} \in \mathcal{B}$, est une variable aléatoire réelle gaussienne centrée et de variance $\sigma^2(\mathrm{B})$. Nous allons démontrer que la mesure stochastique ξ_0 n'est pas une mesure aléatoire, car, si ξ_0 était une mesure aléatoire elle aurait une variation totale presque sûrement infinie dans chaque B$\in \mathcal{B}$, comme dans le cas des processus stochastiques à accroissements indépendants indexés par un intervalle de la droite, cf. [12], page 279.

Fixons B$\in \mathcal{B}$. Pour chaque partition finie de B, $\Pi_{\mathrm{B}} = \{\mathrm{B}_1, \ldots, \mathrm{B}_k\}$, par des éléments de l'anneau $\mathcal{B}$, et $\mu \in \mathcal{M}_c$, définissons les applications

$$\Pi_{\mathrm{B}}^m(\mu) = \sum_{i=1}^{k} |\mu(\mathrm{B}_i)|^m, \quad m \in \mathbb{N}.$$

Evidemment si $\mu \in \mathcal{M}_c$ $|\mu|(\mathrm{B}) = \sup_{\Pi_{\mathrm{B}}}\{\Pi_{\mathrm{B}}^1(\mu)\}$.

Définition 3.2.6 *1. Une suite de partitions de B, $\{\Pi_{n,\mathrm{B}}\}$ est décroissante si nous obtenons $\Pi_{n+1,\mathrm{B}}$ en prenant des partitions d'un nombre fini d'éléments de $\Pi_{n,\mathrm{B}}$.*

2. Une suite de partitions de B, $\{\Pi_{n,\mathrm{B}}\}$ est infinitésimale si

$$\max_{1 \leq k \leq k_n} diam(B_{k,n}) \longrightarrow 0, si\ n \to \infty.$$

3. Etant donné une partition Π_{B} et une mesure $\nu \in \mathcal{M}_c^+$ nous définissons la taille de la partition par rapport à ν comme:

$$\delta\Pi_{\mathrm{B}}(\nu) = \max_{1 \leq i \leq k}\{\nu(B_i)\}.$$

Théorème 3.2.7 *Soit μ une mesure de Radon non atomique. Si $\{\Pi_{n,B}\}$ est une suite décroissante de partitions de B, qui vérifie $\delta\Pi_{n,B}(|\mu|) \longrightarrow 0$, alors $\Pi^2_{n,B}(\mu) \longrightarrow 0$.*

<u>**Démonstration**</u> : La mesure μ étant de Radon est à variation totale finie sur B$\in \mathcal{B}$. Alors, il existe L> 0 tel que $\Pi^1_{n,B}(\mu) \leq$L, pour tout $n \in\mathbb{N}$. Comme $\delta\Pi_{n,B}(|\mu|) \longrightarrow 0$, il s'ensuit

$$\forall_{\varepsilon>0}\exists_{n_0\in\mathbb{N}} : n \geq n_0 \Longrightarrow |\mu|(B_{n,i}) < \varepsilon, \quad i = 1, \ldots, k_n.$$

Donc

$$\Pi^2_{n,B}(\mu) \leq \Pi^1_{n,B}(\mu)\delta\Pi_{n,B}(|\nu|) \leq L\delta\Pi_{n,B}(|\nu|) \longrightarrow 0.\blacksquare$$

Théorème 3.2.8 *ξ_0 n'est pas une mesure aléatoire.*

<u>**Démonstration**</u> : Supposons que ξ_0 soit une mesure aléatoire. Nous allons montrer que les propriétés déjà imposées sur ξ_0 impliquent que la variation totale de ξ_0 sur chaque B$\in \mathcal{B}$ est presque sûrement infinie, ce qui est en contradiction avec la définition d'une mesure aléatoire. Pour cela nous allons nous baser sur la démonstration du théorème 2.3 de Hida [20]. Considérons B fixé et choisissons une suite de partitions de B décroissante $\{\Pi_{n,B}\}$ vérifiant la propriété suivante: on passe de $\Pi_{n,B}$ à $\Pi_{n+1,B}$ en décomposant un seul élément $B_{n,r}$ de $\Pi_{n,B}$ en $B^1_{n+1} \cup B^2_{n+1}$ (où $B^1_{n+1} \cap B^2_{n+1} = \emptyset$), les autres éléments de la partition $\Pi_{n+1,B}$ étant les ensembles $B_{n,1}, \ldots, B_{n,r-1}, B_{n,r+1}, \ldots, B_{n,k_n}$.

Pour chaque $\omega \in \Omega$ posons $V^2_n(\omega) = \Pi^2_{n,B}(\xi_0(\omega))$. Démontrons d'abord que $V^2(\omega) = \lim_{n\to\infty} V^2_n(\omega)$ existe presque sûrement.

Démontrons que

$$\mathbf{E}(\xi_0(B_{n,j})\xi_0(B_{n,l})|V^2_n, V^2_{n+1}, \ldots, V^2_{n+m}) = 0, \quad n, j, l, m \in \mathbb{N}.$$

Tenant compte de la façon dont on obtient $\Pi_{n+1,B}$ à partir de $\Pi_{n,B}$, nous pouvons écrire

$$V^2_n - V^2_{n+1} = 2\xi_0(B^1_{n+1})\xi_0(B^2_{n+1}) \tag{3.8}$$

d'où, en écrivant les différences (3.8) pour V^2_{n+1}-$V^2_{n+2}, \ldots,$ V^2_{n+m-1}-V^2_{n+m}, nous obtenons

$$\begin{aligned}&\mathbf{E}(\xi_0(B_{n,j})\xi_0(B_{n,l})|V^2_n, V^2_{n+1}, \ldots, V^2_{n+m}) = \\ &\qquad \mathbf{E}(\xi_0(B_{n,j})\xi_0(B_{n,l})|V^2_n, \xi_0(B^1_{n+1})\xi_0(B^2_{n+1}), \ldots, \xi_0(B^1_{n+m})\xi_0(B^2_{n+m})).\end{aligned}$$

En plus, comme nous pouvons écrire $V_n^2 = (\xi_0(B_{n,j}))^2 + (\xi_0(B_{n,l}))^2 + \tilde{V}_n^2$, avec $(\xi_0(B_{n,j}))^2$, $(\xi_0(B_{n,l}))^2$ et $\tilde{V}_n^2$ indépendants, au lieu de conditionner par V_n^2 nous pouvons conditionner par $(\xi_0(B_{n,j}))^2$, $(\xi_0(B_{n,l}))^2$ et $\tilde{V}_n^2$. Maintenant, observons que, tenant compte de la propriété que nous avons imposée à la suite de partitions, pour tout $k \leq m$ il se produit une seule des trois possibilités suivantes:

- $B_{n+k}^1, B_{n+k}^2 \subset B_{n,j}$;
- $B_{n+k}^1, B_{n+k}^2 \subset B_{n,l}$;
- $B_{n+k}^1 \cap (B_{n,j} \cup B_{n,l}) = B_{n+k}^2 \cap (B_{n,j} \cup B_{n,l}) = \emptyset$.

Alors, du fait que ξ_0 est à accroissements indépendants, nous pouvons exclure du conditionnement les indices $n+k$ correspondant à la troisième possibilité. De même, nous pouvons exclure $\tilde{V}_n^2$.

Soit $A \subset \mathbb{R}^{r+2}$, où r est le nombre d'ensembles $B_{n+m_i}^1$ qui restent dans le conditionnement, et intégrons l'esperance conditionelle

$$\mathbf{E}(\xi_0(B_{n,j})\xi_0(B_{n,l}) | (\xi_0(B_{n,j}))^2, (\xi_0(B_{n,l}))^2, \xi_0(B_{n+m_1}^1)\xi_0(B_{n+m_1}^2), \ldots, \xi_0(B_{n+m_r}^1)\xi_0(B_{n+m_r}^2))$$

sur l'ensemble

$$\{((\xi_0(B_{n,j}))^2, (\xi_0(B_{n,l}))^2, \xi_0(B_{n+1}^1)\xi_0(B_{n+1}^2), \ldots, \xi_0(B_{n+k}^1)\xi_0(B_{n+k}^2)) \in A\}.$$

Tenant compte du fait que ξ_0 est à accroissements indépendants cette intégrale est égale à:

$$\int_{A^*} (\sum_{i=1}^{s} x_i^1 + x_i^2)^2 (\sum_{j=1}^{t} y_j^1 + y_j^2)^2$$

$$\frac{1}{\sqrt{(2\pi)^{s+t}}} \frac{1}{\sqrt{\prod_{i=1}^{s} \sigma^2(B_{n+m_i}^1)\sigma^2(B_{n+m_i}^2)}} \frac{1}{\sqrt{\prod_{j=1}^{t} \sigma^2(B_{n+m_j}^1)\sigma^2(B_{n+m_j}^2)}}$$

$$\exp\left\{ -\frac{1}{2}\sum_{i=1}^{s} \left[\frac{(x_i^1)^2}{\sigma^2(B_{n+m_i}^1)} + \frac{(x_i^2)^2}{\sigma^2(B_{n+m_i}^2)} \right] - \frac{1}{2}\sum_{j=1}^{t} \left[\frac{(y_j^1)^2}{\sigma^2(B_{n+m_j}^1)} + \frac{(y_j^2)^2}{\sigma^2(B_{n+m_j}^2)} \right] \right\}$$

$$dx_1^1 dx_1^2 \cdots dx_s^1 dx_s^2 dy_1^1 dy_1^2 \cdots dy_t^1 dy_t^2$$

où $s + t = r$ et les variables d'intégration x_i^1, x_i^2 correspondent aux variables aléatoires réelles $\xi_0(B_{n+m_i}^1)$, $\xi_0(B_{n+m_i}^2)$ telles que $B_{n+m_i}^1$, $B_{n+m_i}^2 \subset B_{n,j}$, respectivement et les y_j^1, y_j^2 correspondent aux variables aléatoires réelles $\xi_0(B^1_{n+m_j})$, $\xi_0(B^2_{n+m_j})$ telles que $B^1_{n+m_j}$, $B^2_{n+m_j} \subset B_{n,l}$, respectivement. L'ensemble d'intégration A s'exprime en termes des variables $x_1^1, x_1^2, \ldots, x_s^1, x_s^2$, $y_1^1, y_1^2, \ldots, y_t^1, y_t^2$, en imposant des conditions sur $(\sum_{i=1}^s x_i^1 + x_i^2)^2, (\sum_{j=1}^t y_j^1 + y_j^2)^2$ et sur des produits de la forme $x_i^1 x_i^2$ et $y_j^1 y_j^2$. Donc, l'ensemble A relativement aux variables d'intégration, A^*, est symétrique par rapport à l'origine. La fonction que l'on intègre est aussi symétrique par rapport à l'origine. Comme il s'agit d'une intégrale par rapport à une densité gaussienne de moyenne nulle il s'ensuit que l'intégrale est égale à zéro. Comme on a choisi l'ensemble A arbitraire il s'ensuit que

$$\mathbf{E}(\xi_0(B_{n,j})\xi_0(B_{n,l}) | V_n^2, V_{n+1}^2, \ldots, V_{n+m}^2) = 0, \quad n, j, l, m \in \mathbb{N}.$$

Alors de $V_n^2 - V_{n+1}^2 = 2\xi_0(B^1_{n+1})\xi_0(B^2_{n+1})$ et de l'égalité démontrée on déduit

$$\mathbf{E}(V_n^2 | V_{n+1}^2, \ldots, V_{n+m}^2) = V_{n+1}^2, \quad n, m \in \mathbb{N},$$

donc aussi

$$\mathbf{E}(V_n^2 | V_{n+1}^2, \ldots, V_{n+m}^2, \ldots) = V_{n+1}^2, \quad n \in \mathbb{N}.$$

Posant $X_n = V_{-n}^2$, $n \in \mathbb{Z}^-$, nous définissons une martingale. Alors, il existe presque sûrement $V^2(\omega) = \lim_{n \to -\infty} X_n(\omega) = \lim_{n \to \infty} V_n^2(\omega)$ (Hida [20], lemme 2.2).

Nous avons démontré l'existence de cette limite pour toute suite de partitions décroissante. Supposons maintenant que la suite décroissante de partitions de B, $\{\Pi_{n,B}\}$, vérifie en plus $\Pi^2_{n,B}(\sigma^2) \longrightarrow 0$, quand $n \longrightarrow \infty$. Une telle suite de partitions existe parce que σ^2 est une mesure de Radon non atomique. D'autre part $\mathbf{E}(V_n^2) = \sum_{i=1}^{k_n} \mathbf{E}(\xi_0(B_{n,i}))^2 = \sigma^2(B)$, et, tenant compte du fait que les variables aléatoires sont gaussiennes,

$$\mathbf{E}(V_n^2 - \sigma^2(B))^2 = \mathbf{E}(V_n^2)^2 - (\sigma^2(B))^2 =$$

$$= \sum_{i=1}^{k_n} \mathbf{E}\left[\xi_0^4(B_{n,i})\right] + 2\sum_{i<j} \mathbf{E}\left[\xi_0^2(B_{n,i})\right] \mathbf{E}\left[\xi_0^2(B_{n,j})\right] - (\sigma^2(B))^2 =$$

$$= \sum_{i=1}^{k_n} 3(\sigma^2(B_{n,i}))^2 + 2\sum_{i<j} (\sigma^2(B_{n,i}))^2(\sigma^2(B_{n,j}))^2 - (\sigma^2(B))^2 =$$

$$= 2\sum_{i=1}^{k_n}(\sigma^2(B_{n,i}))^2 = 2\Pi^2_{n,B}(\sigma^2).$$

Donc, $V_n^2 \xrightarrow{m.q.} \sigma^2(B)$. On sait déjà que la suite V_n^2 est presque sûrement convergente, alors $V_n^2 \longrightarrow \sigma^2(B)$ presque sûrement. Maintenant, si $\sigma^2(B) = 0$ cela signifie que la variable aléatoire $\xi_0(B)$ est presque sûrement constante et centrée. La même propriété est vraie pour tout $C \subset B$ car σ^2 est une mesure non négative. Donc $\Pi^1_{n,B}(\xi_0) = 0$ pour tout $n \in \mathbb{N}$, c'est à dire, la variation totale de ξ_0 sur B est presque sûrement nulle. Par contre, si $\sigma^2(B) > 0$, la convergence $V_n^2 \longrightarrow \sigma^2(B)$ p.s. peut s'écrire $\Pi^2_{n,B}(\xi_0) \longrightarrow \sigma^2(B) > 0$ p.s.. Donc, d'après le théorème précédent, la variation totale de ξ_0 dans B est presque sûrement infinie. ∎

Nous avons ainsi démontré que pour que la formule de représentation (3.7) puisse représenter une mesure aléatoire il est nécessaire que σ^2 soit la mesure nulle (chose déjà connue pour les processus stochastiques indexés par $\mathbb{R}$, à accroissements indépendants, cf. [12], th. 8, page 279).

3.2.2 La partie poissonnienne d'une mesure aléatoire à accroissements indépendants infiniment divisible

Pour terminer l'étude de la formule de représentation de la fonction caractéristique d'une mesure aléatoire à accroissements indépendants infiniment divisible ξ il nous reste à étudier la bimesure ν. Il faut aussi imposer des conditions pour assurer que la variation totale soit presque sûrement finie. Cette section s'inspire et généralise des résultats présentés dans le chapitre 4 de Gihman,Skorohod [12].

Définissons la bimesure ξ' à partir de ξ

$$\xi'(A,B) = \sum_{s \in B} \mathbb{I}_A(\xi\{s\}),$$

pour $B \in \mathcal{B}$, $A \in \mathcal{B}(\mathbb{R})$, tel que $0 \notin \overline{A}$. Il est évident que, pour A fixé, $\xi'(A,\cdot)$ est une mesure aléatoire positive n'ayant que des masses ponctuelles, à accroissements indépendants, et toutes ses masses ponctuelles ont la valeur 1. Nous allons approfondir l'étude de cette bimesure.

Théorème 3.2.9 *Pour tout $B \in \mathcal{B}$, et tout $A \in \mathcal{B}(\mathbb{R})$ tel que $0 \notin \overline{A}$, $\xi'(A,B)$ est une variable aléatoire qui suit une loi de Poisson.*

<u>Démonstration</u> : Soit $\Pi_{n,\mathrm{B}}=\{\mathrm{B}_{n,1},\ldots,\mathrm{B}_{n,k_n}\}$ une suite infinitésimale de partitions de B et supposons que, pour tout $n \in \mathbb{N}$, $\Pi_{n+1,\mathrm{B}}$ soit un raffinement de $\Pi_{n,\mathrm{B}}$. Une telle suite de partitions existe, cf. Kallenberg [26], page 11. Comme $\xi'(\mathrm{A},\cdot)$ a au plus une infinité dénombrable de masses ponctuelles, et comme les masses sont toutes égales à 1, il s'ensuit que $\mathbf{P}\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\} \longrightarrow 0$, quand $n \longrightarrow \infty$. Comme

$$\begin{aligned}\left\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\right\} &= \left\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) > 1\right\} \cup \\ &\cup \left\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) \leq 1, \xi'(\mathrm{A},\mathrm{B}_{n,2}) > 1\right\} \cup \\ &\cup \left\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) \leq 1, \xi'(\mathrm{A},\mathrm{B}_{n,2}) \leq 1, \xi'(\mathrm{A},\mathrm{B}_{n,3}) > 1\right\} \cup \ldots \cup \\ &\cup \left\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) \leq 1, \ldots, \xi'(\mathrm{A},\mathrm{B}_{n,k_{n-1}}) \leq 1, \xi'(\mathrm{A},\mathrm{B}_{n,k_n}) > 1\right\}\end{aligned}$$

et les ensembles du membre droit de cette expression sont deux à deux disjoints il s'ensuit, tenant compte du fait que ξ' est à accroissements indépendants

$$\begin{aligned}&\mathbf{P}\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\} = \\ &\sum_{k=1}^{k_n} \prod_{j=1}^{k} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,j}) \leq 1\} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\} \geq \\ &\left(\sum_{k=1}^{k_n} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\}\right) \left(\prod_{j=1}^{k_n} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,j}) \leq 1\}\right),\end{aligned}$$

donc

$$\begin{aligned}\textstyle\sum_{k=1}^{k_n} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\} &\leq \frac{\mathbf{P}\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k})>1\}}{\prod_{j=1}^{k_n} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,j})\leq 1\}} = \\ &= \frac{\mathbf{P}\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k})>1\}}{1-\mathbf{P}\{\sup_k \xi'(\mathrm{A},\mathrm{B}_{n,k})>1\}},\end{aligned}$$

d'où nous déduisons

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) > 1\} = 0.$$

Soit $p_{n,k} = \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) = 1\}$, et supposons qu'il existe $a > 0$ tel que

$$\lim_{n\to\infty} \max_k p_{n,k} = a.$$

Cela veut dire que pour tout $n \in \mathbb{N}$ il existe des indices k_j tels que $p_{n,k_j} > \frac{a}{2}$. Supposons qu'il y ait une infinité dénombrable d'indices k_j dans cette condition. Alors

$$\sum_j p_{n,k_j} = \sum_j \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k_j}) = 1\} = +\infty.$$

Comme les boréliens B_{n,k_j} sont, pour n fixé, deux à deux disjoints, et la mesure aléatoire ξ est à accroissements indépendants, le lemme de Borel-Cantelli implique

$$\mathbf{P}\left(\limsup \{\xi'(\mathrm{A},\mathrm{B}_{n,k_j}) = 1\}\right) = 1,$$

ce qui entraine $\xi(\mathrm{B}) = +\infty$ p.s. car $\xi(\mathrm{B}) \geq \sum_j \xi'(\mathrm{A},\mathrm{B}_{n,k_j})$, ce qui est en contradiction avec le fait que ξ est presque sûrement de Radon. Donc, pour chaque n fixé, il n'existe qu'une quantité finie d'indices k_j tels que $p_{n,k_j} > \frac{a}{2}$. Alors, pour chaque $n \in \mathbb{N}$ nous pouvons choisir un de ces indices, que nous appellerons k_n, de telle façon que la suite des éléments B_{n,k_n} des partitions soit décroissante. Nous obtenons ainsi une suite décroissante $\mathrm{B}_{n,k_n} \downarrow \emptyset$ telle que $p_{n,k_n} \not\longrightarrow 0$, ce qui implique l'existence d'atomes fixes. Donc nous avons

$$\lim_{n\to\infty} \max_k p_{n,k} = 0.$$

Du fait que la mesure aléatoire n'a pas d'atomes fixes il s'ensuit également qu'il existe $\delta > 0$ tel que si $diam(\mathrm{B}_{n,k}) < \delta$ alors $\mathbf{P}\ \{\xi'(\mathrm{A},\mathrm{B}_{n,k}) = 0\} > 0$. Comme $\xi'(\mathrm{A},\mathrm{B}) = \xi'(\mathrm{A},\mathrm{B}_{n,1}) + \cdots + \xi'(\mathrm{A},\mathrm{B}_{n,k})$ et ces variables sont non négatives, il s'ensuit

$$\begin{aligned}\mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}) = 0\} &= \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) = 0, \ldots, \xi'(\mathrm{A},\mathrm{B}_{n,k}) = 0\} = \\ &= \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,1}) = 0\} \cdots \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}_{n,k}) = 0\},\end{aligned}$$

donc $\mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}) = 0\} > 0$.

Alors, des relations

$$\mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}) = 0\} = \mathbf{P}\{\sum_{k=1}^{k_n} \xi'(\mathrm{A},\mathrm{B}_{n,k}) = 0\} = \prod_{k=1}^{k_n} (1 - p_{n,k})$$

nous déduisons

$$-\log \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}) = 0\} = \sum_{k=1}^{k_n} p_{n,k}(1 + O(\max_k p_{n,k})),$$

d'où il s'ensuit, en posant $\lambda(\mathrm{A,B}) = -\log \mathbf{P}\{\xi'(\mathrm{A,B}) = 0\}$,

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} p_{n,k} = \lambda(\mathrm{A,B}).$$

Pour terminer la démonstration, calculons maintenant $\mathbf{P}\{\xi'(\mathrm{A,B}) = \mathrm{M}\}$, $\mathrm{M}\in\mathbb{N}$.

$$\begin{aligned}\mathbf{P}\{\xi'(\mathrm{A,B}) = \mathrm{M}\} &= \\ &= \lim_{n\to\infty} \sum_{i_1<i_2<\ldots<i_\mathrm{M}} p_{n,i_1}\cdots p_{n,i_\mathrm{M}} \prod_{k=1,k\neq i_j}^{k_n} (1-p_{n,k}) = \\ &= \lim_{n\to\infty} \sum_{i_1<\ldots<i_\mathrm{M}} \frac{p_{n,i_1}\cdots p_{n,i_\mathrm{M}}}{(1-p_{n,i_1})\cdots(1-p_{n,i_\mathrm{M}})} \prod_{k=1}^{k_n}(1-p_{n,k}).\end{aligned}$$

De la définition de $\lambda(\mathrm{A,B})$ il s'ensuit

$$e^{-\lambda(\mathrm{A,B})} = \prod_{k=1}^{k_n}(1-p_{n,k}), \quad n\in\mathbb{N},$$

donc

$$\mathbf{P}\{\xi'(\mathrm{A,B}) = \mathrm{M}\} = e^{-\lambda(\mathrm{A,B})} \lim_{n\to\infty} \sum_{i_1<\ldots<i_\mathrm{M}} \frac{p_{n,i_1}\cdots p_{n,i_\mathrm{M}}}{(1-p_{n,i_1})\cdots(1-p_{n,i_\mathrm{M}})}.$$

Les quantités $p_{n,k}$ convergent vers zero quand $n \longrightarrow \infty$, avec k fixé. Alors, en ce qui concerne la limite, nous pouvons remplacer les expressions $\frac{p_{n,i}}{1-p_{n,i}}$ par $p_{n,i}$. Donc la limite qu'il faut calculer est égale à

$$\lim_{n\to\infty} \sum_{i_1<\ldots<i_\mathrm{M}} p_{n,i_1}\cdots p_{n,i_\mathrm{M}}.$$

Comme

$$\sum_{i_1<\ldots<i_\mathrm{M}} p_{n,i_1}\cdots p_{n,i_\mathrm{M}} = \frac{1}{\mathrm{M}!}(\sum_{k=1}^{k_n} p_{n,k})^\mathrm{M} + O\left[\sum_{l=1}^{\mathrm{M}-1}(\sum_{k=1}^{k_n} p_{n,k})^l(\max_k p_{n,k})^{\mathrm{M}-l}\right].$$

il s'ensuit

$$\mathbf{P}\{\xi'(\mathrm{A,B}) = \mathrm{M}\} = e^{\lambda(\mathrm{A,B})}\frac{(\lambda(\mathrm{A,B}))^\mathrm{M}}{\mathrm{M}!},$$

ce qui termine la démonstration. ■

Définissons maintenant les variables aléatoires réelles

$$\xi_{\mathrm{A}}(\mathrm{B}) = \int_{\mathrm{A}} x\xi'(dx, \mathrm{B}) = \sum_{s\in \mathrm{B}} \xi(\{s\})\mathbb{1}_A(\xi(\{s\})),$$

pour B$\in \mathcal{B}$, A$\in \mathcal{B}(\mathbb{R})$, tel que $0 \notin \overline{\mathrm{A}}$. Cela signifie que $\xi_{\mathrm{A}}(\mathrm{B})$ est la somme des masses ponctuelles dans B dont les valeurs sont dans A.

Théorème 3.2.10 *Les mesures aléatoires $\xi_{\mathrm{A}}(\cdot)$ et $\xi(\cdot) - \xi_{\mathrm{A}}(\cdot)$ sont indépendantes et à accroissements indépendants, pour $A\in \mathcal{B}(\mathbb{R})$ tel que $0 \notin \overline{A}$.*

<u>**Démonstration**</u> : Comme la mesure aléatoire ξ est supposée non atomique il s'ensuit que $\xi - \xi_{\mathrm{A}}$ est non atomique. Il est évident que $(\xi_{\mathrm{A}}, \xi - \xi_{\mathrm{A}})$ est à accroissements indépendants sur la diagonale de $\mathcal{B} \times \mathcal{B}$. Pour démontrer l'indépendance des mesures aléatoires ξ_{A} et $\xi - \xi_{\mathrm{A}}$ il suffit d'établir, pour tout B$\in \mathcal{B}$,

$$\mathbf{E}\left[e^{iz_1\xi_{\mathrm{A}}(\mathrm{B})+iz_2(\xi(\mathrm{B})-\xi_{\mathrm{A}}(\mathrm{B}))}\right] = \mathbf{E}\left[e^{iz_1\xi_{\mathrm{A}}(\mathrm{B})}\right]\mathbf{E}\left[e^{iz_2(\xi(\mathrm{B})-\xi_{\mathrm{A}}(\mathrm{B}))}\right]. \tag{3.9}$$

En effet, tenant compte de l'indépendance des accroissements, cette égalité implique

$$\mathbf{E}\left[e^{i\sum_k z_{1k}\xi_{\mathrm{A}}(\mathrm{B}_k)+i\sum_k z_{2k}(\xi(\mathrm{B}_k)-\xi_{\mathrm{A}}(\mathrm{B}_k))}\right] =$$

$$\mathbf{E}\left[e^{i\sum_k z_{1k}\xi_{\mathrm{A}}(\mathrm{B}_k)}\right]\mathbf{E}\left[e^{i\sum_k z_{2k}(\xi(\mathrm{B}_k-\xi_{\mathrm{A}}(\mathrm{B}_k))}\right]$$

pour toute partition $\{\mathrm{B}_1, \cdots, \mathrm{B}_k\}$ du borélien B, et pour tous réels z_{1k}, z_{2k}, ce qui signifie l'indépendance des mesures aléatoires ξ_{A} et $\xi - \xi_{\mathrm{A}}$.

Démontrons d'abord la relation (3.9) pour le cas où $\mathbf{E}[\xi'(\mathrm{fr(A),B})]=0$. Soit $\Pi_{n,\mathrm{B}}=\{\mathrm{B}_{n,1}, \cdots, \mathrm{B}_{n,n}\}$ une suite de partitions infinitésimale de B. Nous avons

$$\xi_{\mathrm{A}}(\mathrm{B}) = \lim_{n\to\infty} \sum_{k=1}^{n} \xi(\mathrm{B}_{n,k})\mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k})). \tag{3.10}$$

Posons

$$\xi_{n,k} = \xi(\mathrm{B}_{n,k})\mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k})), \quad \eta_{n,k} = \xi(\mathrm{B}_{n,k}) - \xi_{n,k}.$$

Alors, tenant compte de (3.10), pour démontrer (3.9) il suffit de montrer que

$$\lim_{n\to\infty}\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right| = 0.$$

Comme les couples $(\xi_{n,k}, \eta_{n,k})$ sont indépendants, pour n fixé, il s'ensuit

$$\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right| =$$

$$\left|\prod_k \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \prod_k \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right| \leq$$

$$\leq \sum_{k=1}^{n}\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right|.$$

Comme $\xi_{n,k}\eta_{n,k} = 0$, il vient

$$\begin{aligned} e^{iz_1\xi_{n,k}+iz_2\eta_{n,k}} &= 1+\textstyle\sum_{m=1}^{\infty} \frac{(iz_1\xi_{n,k}+iz_2\eta_{n,k})^m}{m!} = \\ & 1+\textstyle\sum_{m=1}^{\infty}\left[\frac{(iz_1\xi_{n,k})^m}{m!}+\frac{(iz_2\eta_{n,k})^m}{m!}\right] = \\ & e^{iz_1\xi_{n,k}}+e^{iz_2\eta_{n,k}}-1, \end{aligned}$$

donc

$$\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right| =$$

$$= \left|\mathbf{E}\left[e^{iz_1\xi_{n,k}}\right]-1\right|\left|\mathbf{E}\left[e^{iz_2\eta_{n,k}}\right]-1\right|.$$

De plus,

$$\left|\mathbf{E}\left[e^{iz_1\xi_{n,k}}\right]-1\right| \leq \mathbf{E}\left|e^{iz_1\xi_{n,k}}-1\right| \leq$$
$$2\mathbf{P}(|\xi_{n,k}|>0) = 2\mathbf{P}(\mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k}))>0),$$

et, pour $\rho > 0$,

$$\left|\mathbf{E}\left[e^{iz_1\xi_{n,k}}\right]-1\right| \leq \sup_{|x|\leq\rho}|1-e^{iz_2x}| + 2\mathbf{P}(|\eta_{n,k}|>\rho),$$

d'où il s'ensuit

$$\limsup_{n\to\infty}\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right] - \mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right| \leq$$

$$\limsup\nolimits_{n\to\infty}\left[\sup\nolimits_{|x|\leq\rho}|1-e^{iz_2x}| + 2\sup\nolimits_k \mathbf{P}(|\eta_{n,k}|>\rho)\right]\cdot$$
$$\cdot\limsup\nolimits_{n\to\infty}\left[2\textstyle\sum_k \mathbf{P}(\mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k}))>0)\right] =$$

$$2\limsup_{n\to\infty}\left[\sup_{|x|\le\rho}|1-e^{iz_2x}|+2\sup_k \mathbf{P}(|\eta_{n,k}|>\rho)\right]\cdot$$
$$\cdot\limsup_{n\to\infty}\mathbf{E}\left[\textstyle\sum_k \mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k}))\right].$$

Comme nous supposons que $0\notin\overline{\mathrm{A}}$, il existe $\varepsilon>0$ tel que $\inf_{x\in\mathrm{A}}|x|\ge\varepsilon$. Alors, pour $\rho\le\varepsilon$,

$$\mathbf{P}(|\eta_{n,k}|>\rho)\le\mathbf{P}(|\xi(\mathrm{B}_{n,k})|>\rho),$$

donc, du fait que ξ n'a pas d'atome fixe, il s'ensuit que

$$\limsup_{n\to\infty}\sup_k \mathbf{P}(|\eta_{n,k}|>\rho)=0.$$

Comme

$$\lim_{n\to\infty}\sum_k \mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k}))=\xi'(\mathrm{A,B}),$$

on trouve

$$\limsup_{n\to\infty}\mathbf{E}\left[\sum_k \mathbb{1}_{\mathrm{A}}(\xi(\mathrm{B}_{n,k}))\right]\le-\log\mathbf{P}\{\xi'(\mathrm{A,B})=0\}=\lambda(\mathrm{A,B})<\infty.$$

Donc

$$\limsup_{n\to\infty}\left|\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}+i\sum_k z_2\eta_{n,k}}\right]-\mathbf{E}\left[e^{i\sum_k z_1\xi_{n,k}}\right]\mathbf{E}\left[e^{i\sum_k z_2\eta_{n,k}}\right]\right|\le$$
$$2\lambda(\mathrm{A,B})\sup_{|x|\le\rho}|1-e^{iz_2x}|.$$

En passant à la limite quand $\rho\longrightarrow0$ on obtient (3.9), ce qui démontre le théorème pour le cas $\mathbf{E}\xi'(\mathrm{fr(A),B})=0$.

La classe des ensembles C pour lesquels le théorème est vrai est une classe monotone, que nous désignerons par $\mathcal{M}$. Les ensembles C tels que $\inf_{x\in\mathrm{C}}|x|\ge\varepsilon$ et $\mathbf{E}\xi'(\mathrm{fr(C),B})=0$ forment une algèbre $\mathcal{A}_\varepsilon$, pour $\varepsilon>0$ bien choisi. En effet, il est facile de vérifier que si $\mathrm{C_1,C_2}\in\mathcal{A}_\varepsilon$ alors $\mathrm{C_1\cup C_2}\in\mathcal{A}_\varepsilon$, du fait que $\mathrm{fr(C_1\cup C_2)}\subset\mathrm{fr(C_1)\cup fr(C_2)}$, et que, dans le cas $\mathrm{C_1}\subset\mathrm{C_2}$, du fait que $\mathrm{fr(C_2\backslash C_1)}\subset\mathrm{fr(C_1)\cup fr(C_2)}$, il s'ensuit $\mathrm{C_2\backslash C_1}\in\mathcal{A}_\varepsilon$. Les ensembles $\{-\varepsilon,\varepsilon\}$ ne peuvent vérifier $\mathbf{E}\xi'(\{-\varepsilon,\varepsilon\},\mathrm{B})>0$ que pour une quantité au plus dénombrable de valeurs de ε. Donc, nous pouvons choisir un ε aussi petit que nous le voulons de façon que $\mathbf{E}\xi'(\mathrm{fr}(\mathbb{R}\backslash(-\varepsilon,\varepsilon)),B)=0$. Nous avons donc vérifié que la classe $\mathcal{A}_\varepsilon$ est bien une algèbre. De la même façon que nous avons vérifié que nous pouvons choisir $\varepsilon > 0$ pour obtenir une algèbre, nous déduisons que si $x \in \mathbb{R}\backslash(-\varepsilon,\varepsilon)$, il existe une boule de centre en x et de rayon aussi petit que l'on veut appartenant à l'algèbre $\mathcal{A}_\epsilon$. Donc la tribu borélienne de $\mathbb{R}\backslash(-\varepsilon,\varepsilon)$ est contenue dans $\mathcal{M}$. Comme l'ensemble A est tel que $0 \notin \overline{\mathrm{A}}$, il existe un valeur $\varepsilon > 0$ telle que A$\in \mathcal{A}_\epsilon$, ce qui termine la démonstration du théorème. ■

Corollaire 3.2.11 *Soient $A_1,\ldots,A_n \in \mathcal{B}(\mathbb{R})$ tels que $0 \notin \overline{A}_1,\ldots,0 \notin \overline{A}_n$ et sont deux à deux disjoints. Alors, les mesures aléatoires $\xi_{\mathrm{A}_1},\ldots,\xi_{\mathrm{A}_n}$ et $\xi - \sum_{k=1}^n \xi_{\mathrm{A}_k}$ sont indépendantes.*

<u>**Démonstration**</u> : En effet, d'après le théorème précédent $\xi - \sum_{k=1}^n \xi_{\mathrm{A}_k} = \xi - \xi_{\bigcup_k \mathrm{A}_k}$ ne dépend pas de $\xi_{\bigcup_k \mathrm{A}_k}$. Les mesures ξ_{A_k} sont fonctions mesurables de la mesure aléatoire $\xi_{\bigcup_k \mathrm{A}_k}$. Alors $\xi_{\mathrm{A}_1},\ldots,\xi_{\mathrm{A}_n}$ sont indépendantes de $\xi - \sum_{k=1}^n \xi_{\mathrm{A}_k}$. De la même façon, les mesures aléatoires ξ_{A_k}, $k \neq j$, $\xi - \sum_{k=1}^n \xi_{\mathrm{A}_k}$, sont fonctions mesurables de $\xi - \xi_{\mathrm{A}_j}$ qui est indépendante de ξ_{A_j}, ce qui démontre le corollaire. ■

Corollaire 3.2.12 *Soient $A_1,\ldots,A_n \in \mathcal{B}(\mathbb{R})$ tels que $0 \notin \overline{A}_1,\ldots,0 \notin \overline{A}_n$ et sont deux à deux disjoints. Alors les mesures aléatoires $\xi'(A_1,\cdot),\ldots,\xi'(A_n,\cdot)$ sont indépendantes.*

<u>**Démonstration**</u> : Ceci résulte du fait que $\xi'(\mathrm{A},\cdot)$ est complètement définie par $\xi_{\mathrm{A}}(\cdot)$ et du corollaire précédent. ■

Corollaire 3.2.13 *La fonction λ définie dans la démonstration du théorème 3.2.9 est une bimesure.*

<u>**Démonstration**</u> : C'est évident d'après la définition de λ,

$$\lambda(\mathrm{A},\mathrm{B}) = -\log \mathbf{P}\{\xi'(\mathrm{A},\mathrm{B}) = 0\}$$

du fait que ξ' est une bimesure, et du corollaire précédant. ■

Théorème 3.2.14 -

$$\mathbf{E}(e^{it\xi_{\mathrm{A}}(\mathrm{B})}) = \exp\left[\int_{\mathrm{A}} e^{itx} - 1 \ \mathbf{E}\xi'(dx, B)\right].$$

<u>**Démonstration**</u> : Soit $\{A_k\}$ une partition du borélien A. Alors, si $\rho = \max_k \, diam(A_k)$, et $x_k \in A_k$,

$$|\xi_A(B) - \sum_k x_k \xi'(A_k,B)| \le \rho\xi'(A,B).$$

Du théorème précédent nous savons que les variables aléatoires $\xi'(A,B)$ suivent une loi de Poisson de paramètre $\lambda(A,B) = -\log \mathbf{P}\{\xi'(A,B) = 0\}$, donc, tenant compte de l'indépendance des variables aléatoires $\xi'(A_k,B)$,

$$\mathbf{E}(e^{it\xi_A(B)}) = \lim_{\rho\to 0} \mathbf{E}(e^{it\sum_k x_k\xi'(A_k,B)}) = \lim_{\rho\to 0} \exp\left[\sum_k (e^{itx_k} - 1)\lambda(A_k,B)\right].$$

Comme $\mathbf{E}(\xi'(A,B)) = \lambda(A,B)$, nous avons en effet

$$\begin{aligned}\mathbf{E}(e^{it\xi_A(B)}) &= \lim_{\rho\to 0} \exp\left[\textstyle\sum_k (e^{itx_k} - 1)\, \mathbf{E}\xi'(A_k,B)\right] = \\ &= \exp\left[\textstyle\int_A e^{itx} - 1\, \mathbf{E}\xi'(dx,B)\right]. \ \blacksquare\end{aligned}$$

Pour tout $\varepsilon > 0$ définissons le sous-ensemble de $\mathbb{R}$

$$\Delta_\varepsilon = \{x \in \mathbb{R} : \ |x| > \varepsilon\}.$$

Théorème 3.2.15 *Il existe une mesure (non aléatoire) $\alpha \in \mathcal{M}_c$ telle que, pour tout $B \in \mathcal{B}$,*

$$\lim_{\varepsilon\to 0} \xi_{\Delta_\varepsilon}(B) = \xi(B) - \alpha(B).$$

<u>**Démonstration**</u> : La mesure $\xi - \xi_{\Delta_\varepsilon}$ n'a pas de masses ponctuelles de valeur plus grande que ε en module. D'autre part, comme les mesures ξ_{Δ_ε} et $\xi - \xi_{\Delta_\varepsilon}$ sont concentrées sur des boréliens disjoints nous avons

$$|\xi_{\Delta_\varepsilon}|(B) + |\xi - \xi_{\Delta_\varepsilon}|(B) = |\xi|(B), \quad B \in \mathcal{B},\ \varepsilon > 0,$$

donc la limite presque sûre, $\eta(B) = \lim_{\varepsilon\to 0} |\xi_{\Delta_\varepsilon}|(B)$, existe, parce qu'il s'agit d'une fonction croissante de ε qui est bornée par $|\xi|(B)$.

De même, pour $\varepsilon_1 < \varepsilon_2$, $B \in \mathcal{B}$, on a aussi

$$|\xi_{\Delta_{\varepsilon_1}}|(B) = |\xi_{\Delta_{\varepsilon_2}}|(B) + |\xi_{\Delta_{\varepsilon_1}} - \xi_{\Delta_{\varepsilon_2}}|(B),$$

donc

$$\lim_{\varepsilon_1\to 0, \varepsilon_2\to 0} |\xi_{\Delta_{\varepsilon_1}} - \xi_{\Delta_{\varepsilon_2}}|(B) = 0.$$

Alors $\sup_{C\subset B} |\xi_{\Delta_{\varepsilon_1}}(C) - \xi_{\Delta_{\varepsilon_2}}(C)| \longrightarrow 0$, quand $\varepsilon_1 \longrightarrow 0, \varepsilon_2 \longrightarrow 0$. On a donc démontré que pour $C \in \mathcal{B}$ fixé $\{\xi_{\Delta_{\varepsilon_n}}(C)\}$ est presque sûrement une suite de Cauchy pour toute suite $\varepsilon_n \longrightarrow 0$, donc la limite $\lim_{\varepsilon\to 0} \xi_{\Delta_\varepsilon}(B) = \xi^0(B)$ existe (il est immédiat de vérifier que la limite est indépendante de la suite ε_n convergeant vers zero). Donc, nous avons aussi l'existence de la limite

$$\lim_{\varepsilon\to 0} (\xi - \xi_{\Delta_\varepsilon})(B) = (\xi - \xi^0)(B).$$

Vérifions que ξ^0 est σ-additive. L'application ξ^0 est évidemment additive, donc pour démontrer la σ-additivité il suffit de démontrer que pour toute suite $B_n \downarrow \emptyset$ il se vérifie

$$\lim_{n\to\infty} |\xi^0(B_n)| = 0.$$

Choisissons une suite $\{\varepsilon_n\}$ décroissante vers zéro. Donc, d'après les propriétés déjà démontrées il s'ensuit que les suites $\{\xi_{\Delta_{\varepsilon_m}}(B_n)\}$ sont de Cauchy uniformément par rapport à n. Alors, étant donné $\delta > 0$, il existe $m_0 \in \mathbb{N}$ tel que

$$m \geq m_0 \Rightarrow |\xi^0(B_n) - \xi_{\Delta_{\varepsilon m}}(B_n)| < \delta,\ n \in \mathbb{N}.$$

Alors

$$|\xi^0(B_n)| \leq |\xi^0(B_n) - \xi_{\Delta_{\varepsilon_{m_0}}}(B_n)| + |\xi_{\Delta_{\varepsilon_{m_0}}}(B_n)| \leq \delta + |\xi_{\Delta_{\varepsilon_{m_0}}}(B_n)|,$$

donc

$$\lim_{n\to\infty} |\xi^0(B_n)| \leq \delta + \lim_{n\to\infty} |\xi_{\Delta_{\varepsilon_{m_0}}}(B_n)| = \delta,$$

et, comme δ est arbitraire, il s'ensuit

$$\lim_{n\to\infty} |\xi^0(B_n)| = 0,$$

ce qui démontre que ξ^0 est σ-additive. De plus, pour tout $B \in \mathcal{B}$,

$$|\xi^0|(B) \leq 4 \sup_{C\subset B} |\xi^0(C)| = 4 \sup_{C\subset B} \lim_{\varepsilon\to 0} |\xi_{\Delta_\varepsilon}(C)| \leq$$

$$\leq 4 \lim_{\varepsilon\to 0} \sup_{C\subset B} |\xi_{\Delta_\varepsilon}(C)| \leq 4 \lim_{\varepsilon\to 0} |\xi_{\Delta_\varepsilon}|(B) = 4\eta(B) < \infty.$$

Remarquons que nous pouvons démontrer que η est aussi une mesure de la même façon dont nous avons démontré que ξ^0 est une mesure.

La mesure $\xi - \xi^0$ n'a pas de masse ponctuelle, donc est presque sûrement égale à la partie non aléatoire α parce qu'on a déjà démontré que la partie gaussienne est nulle. Alors nous avons

$$\lim_{\varepsilon\to 0} \xi_{\Delta_\varepsilon}(B) = \xi^0(B) = \xi(B) - \alpha(B). \ \blacksquare$$

Corollaire 3.2.16 *$\xi_{\Delta_\varepsilon} \longrightarrow \xi - \alpha$ presque sûrement, quand $\varepsilon \longrightarrow 0$, au sens de la convergence vague des mesures.*

__Démonstration__ : En effet nous avons les convergences presque sûres

$$\begin{array}{ccc} \xi_{\Delta_\varepsilon}(\mathrm{B}) & \longrightarrow & (\xi - \alpha)(\mathrm{B}) \\ |\xi_{\Delta_\varepsilon}|(\mathrm{B}) & \longrightarrow & \eta(\mathrm{B}) \end{array}$$

pour tout B$\in \mathcal{B}$, où les ensembles de probabilité nulle dépendent du borélien relativement compact B. Considérons sur **S** un anneau dénombrable $\mathcal{D} \subset \mathcal{B}$, tel que tout B$\in \mathcal{B}$ ait une couverture finie par des éléments de $\mathcal{D}$ de diamètre plus petit que $\delta > 0$ arbitrairement choisi (un tel anneau existe cf. Kallenberg [26], page 3). Soit B^δ l'ensemble définit par l'union des éléments de cette couverture et supposons que B est fermé. Comme $\mathcal{D}$ est dénombrable il existe un ensemble de probabilité 1 oú $\xi_{\Delta_\varepsilon}(\mathrm{D}) \longrightarrow (\xi - \alpha)(\mathrm{D})$ et $|\xi_{\Delta_\varepsilon}|(\mathrm{D}) \longrightarrow \eta(\mathrm{D})$, D$\in \mathcal{D}$, indépendant de D. Prenons deux suites $\{\delta_n\}$ et $\{\varepsilon_n\}$ décroissantes vers zéro et considérons la suite double $\{|\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}^{\delta_n})\}$. Dans la démonstration du théroème précédent nous avons prouvé que les suites $\{|\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}^{\delta_n})\}$ convergent uniformément par rapport à n:

$$\forall_{\varepsilon>0}\, \exists_{m_0\in\mathbb{N}}\; m \geq m_0 \;\Rightarrow\; \sup_n \left|\eta(\mathrm{B}^{\delta_n}) - |\xi_{\Delta_{\varepsilon_m}}(\mathrm{B}^{\delta_n})\right| < \varepsilon.$$

Par rapport à m nous avons

$$\sup_n \left||\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}^{\delta_n}) - |\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B})\right| = \sup_n\; |\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}^{\delta_n} \setminus \mathrm{B}) = \eta(\mathrm{B}^{\delta_n} \setminus \mathrm{B})$$

donc

$$\forall_{\varepsilon>0}\, \exists_{n_0\in\mathbb{N}}\; n \geq n_0 \;\Rightarrow\; \sup_n \left||\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}^{\delta_n}) - |\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B})\right| < \varepsilon.$$

Donc il s'ensuit que pour un ensemble de probabilité 1, qui dépend seulement de l'anneau $\mathcal{D}$, il se vérifie que $|\xi_{\Delta_{\varepsilon_m}}|(\mathrm{B}) \longrightarrow \eta(\mathrm{B})$. De la même façon nous déduisons la convergence presque sure $\xi_{\Delta_{\varepsilon_m}}(\mathrm{B}) \longrightarrow (\xi - \alpha)(\mathrm{B})$, oú l'ensemble de probabilité 1 est indépendant de B. Alors, du théorème 1.2.11 il s'ensuit la convergence vague presque sûre $\xi_{\Delta_\varepsilon} \longrightarrow \xi - \alpha$, quand $\varepsilon \longrightarrow 0$. ∎

Corollaire 3.2.17 *Pour tout B$\in \mathcal{B}$*

$$\mathbf{E}\left[e^{it\eta(\mathrm{B})}\right] = \exp\left[\int_{\{|x|>0\}} e^{it|x|} - 1\; \mathbf{E}\xi'(dx, B)\right].$$

<u>**Démonstration**</u> : Il suffit de remarquer

$$|\xi_{\Delta_\epsilon}|(\mathrm{B}) = \int_{\Delta_\epsilon} |x|\, \xi'(dx, \mathrm{B}),$$

le reste de la démonstration suit la démonstration du théorème 3.2.14. ■

Corollaire 3.2.18 $|\xi - \alpha| \leq \eta \leq |\xi|$.

<u>**Démonstration**</u> : L'inégalité $|\xi - \alpha| \leq \eta$ s'ensuit d'après le théorème 1.1.16. D'après la démonstration du théorème précédent nous savons que $\eta(\mathrm{B}) = \lim_{\epsilon\to 0} |\xi_{\Delta_\epsilon}|(\mathrm{B})$ et $|\xi_{\Delta_\epsilon}|(\mathrm{B}) \leq |\xi|(\mathrm{B})$, d'où il s'ensuit $\eta(\mathrm{B}) \leq |\xi|(\mathrm{B})$. ■

Remarquons qu'une mesure aléatoire est intégrable si et seulement si sa variation totale est intégrable.

Théorème 3.2.19 *Supposons que la mesure aléatoire ξ soit intégrable. Pour tout $B \in \mathcal{B}$*

$$\int_{\{0<|x|\leq 1\}} |x|\, \mathbf{E}\xi'(dx, B) < \infty. \tag{3.11}$$

Réciproquement, si la condition (3.11) est vérifiée alors la variable aléatoire réelle $|\xi - \alpha|(B)$ est presque sûrement finie.

<u>**Démonstration**</u> : Condition nécessaire: Soit $\varepsilon > 0$. Alors

$$\int_{\{\varepsilon<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) \leq \int_{\{|x|>\varepsilon\}} |x|\, \xi'(dx, \mathrm{B}) = |\xi_{\Delta_\epsilon}|(\mathrm{B}),$$

donc

$$\int_{\{0<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) = \lim_{\varepsilon\to 0} \int_{\{\varepsilon<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) \leq \lim_{\varepsilon\to 0} |\xi_{\Delta_\epsilon}|(\mathrm{B}) = \eta(\mathrm{B}),$$

d'où il s'ensuit

$$\mathbf{E}\left[\int_{\{0<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B})\right] \leq \mathbf{E}\,(\eta(\mathrm{B})) \leq \mathbf{E}\,[|\xi|(\mathrm{B})] < \infty,$$

d'après l'intégrabilité de la mesure aléatoire.

Condition suffisante: Nous avons, pour tout B$\in \mathcal{B}$,

$$\xi(\mathrm{B}) = \int_{\{0<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B}) + \int_{\{|x|>1\}} x\, \xi'(dx, \mathrm{B}).$$

La mesure alétoire définie par $\int_{\{|x|>1\}} x\, \xi'(dx, \mathrm{B})$ n'a qu'un nombre fini de masses ponctuelles, donc sa variation totale n'est que la somme des valeurs absolues de ces masses, donc forcément finie. De même, nous pouvons conclure que la variation totale de la mesure définie par $\int_{\{\varepsilon<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B})$ est finie, pour tout $\varepsilon > 0$. Le théorème de la convergence dominée implique

$$\int_{\{0<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B}) = \lim_{\varepsilon>0} \int_{\{\varepsilon<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B}).$$

La suite des variations totales $\int_{\{\varepsilon<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B})$ est croissante quand $\varepsilon \downarrow 0$ et bornée. En fait

$$\int_{\{\varepsilon<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) \leq \int_{\{0<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}),$$

et cette variable aléatoire est presque sûrement finie car nous supposons qu'elle est intégrable. Alors, le théorème 1.2.11 nous permet de conclure que les suites de mesures aléatoires

$$\left\{\int_{\{\varepsilon_n<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B})\right\} \quad \text{et} \quad \left\{\int_{\{\varepsilon_n<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B})\right\}$$

presque sûrement vaguement convergentes pour toute suite $\varepsilon_n \downarrow 0$, et cette limite est indépendant de la suite ε_n choisie. Alors, d'après le théorème 1.1.16 il s'ensuit que la variation totale de la mesure aléatoire

$$\int_{\{0<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B}) = \lim_{\varepsilon\downarrow 0} \int_{\{\varepsilon<|x|\leq 1\}} x\, \xi'(dx, \mathrm{B})$$

est inférieur ou égale à

$$\lim_{\varepsilon\downarrow 0} \int_{\{\varepsilon<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) = \int_{\{0<|x|\leq 1\}} |x|\, \xi'(dx, \mathrm{B}) < \infty\ p.s.$$

d'après les hypothèses. ■

3.2.3 Formule de représentation de la fonction caractéristique

Nous pouvons finalement résumer tous les résultats des deux sections précédentes et obtenir la formule de représentation de la fonction caractéristique d'une mesure aléatoire à accroissements indépendants infiniment divisible.

Théorème 3.2.20 *Soit ξ une mesure aléatoire à accroissements indépendants intégrable. La mesure aléatoire ξ est infiniment divisible si et seulement si sa fonction caractéristique est de la forme*

$$\mathbf{E}(e^{i\xi f}) = \exp\left[i\alpha f + \int_{\{|x|>0\}} e^{ixf(s)} - 1 \; \nu(dx, ds)\right], \quad f \in \mathrm{C}_c(\mathbf{S}), \qquad (3.12)$$

où $\alpha \in \mathcal{M}_c$ et ν est une bimesure non négative sur $\mathbb{R}\times\mathbf{S}$, telle que $\nu(\cdot, B)$, $B \in \mathcal{B}$, est mesure de Lévy sur $\mathbb{R}$ vérifiant

$$\int_{\{0<|x|\leq 1\}} |x| \; \nu(dx, B) < \infty, \quad B \in \mathcal{B}.$$

Pour démontrer cette formule il suffit de rassembler les résultats des deux sections précédentes qui nous donnent les caractérisations des parties gaussienne et poissonnienne pour obtenir la variation totale presque sûrement finie.

Quand on compare cette formule avec (3.7) on remarque que dans l'intégrale de (3.12) on devrait intégrer la fonction $e^{ixf(s)} - 1 - \frac{ixf(s)}{1+x^2}$. Nous allons démontrer que l'intégrale de $\frac{ixf(s)}{1+x^2}$ par rapport à ν est finie, donc, comme il s'agit d'une fonction linéaire en f, peut s'additionner à la mesure non aléatoire qui figurait dans (3.7) pour obtenir la mesure α de (3.12).

Corollaire 3.2.21 *Pour toute bimesure ν vérifiant les conditions du théorème précédent*

$$\left| \int_{\{|x|>0\}} \frac{xf(s)}{1+x^2} \; \nu(dx, ds) \right| < \infty, \quad f \in \mathrm{C}_c(\mathbf{S}).$$

__**Démonstration**__ : Soit $f \in \mathrm{C}_c(\mathbf{S})$ et posons $\mathrm{S}_f = \mathrm{supp}(f)$. Comme ν est non négative

$$\begin{aligned} \left|\int_{\{|x|>0\}} \tfrac{xf(s)}{1+x^2} \, \nu(dx, ds)\right| &\leq \left|\int_{\{|x|>0\}} \tfrac{|xf(s)|}{1+x^2} \, \nu(dx, ds)\right| \leq \\ &\leq \|f\| \int_{\{|x|>0\}} \tfrac{|x|}{1+x^2} \, \nu(dx, \mathrm{S}_f). \end{aligned}$$

Comme $f \in \mathbf{C}_c(\mathbf{S})$, $||f|| < \infty$. D'autre part

$$\int_{\{|x|>0\}} \frac{|x|}{1+x^2}\,\nu(dx, \mathrm{S}_f) = \int_{\{0<|x|\leq 1\}} \frac{|x|}{1+x^2}\,\nu(dx, \mathrm{S}_f) + \int_{\{|x|>1\}} \frac{|x|}{1+x^2}\,\nu(dx, \mathrm{S}_f).$$

Dans la seconde intégrale on intègre une fonction bornée sur le complementaire, Δ_1, d'un voisinage de zéro, et $\nu(\Delta_1, \mathrm{S}_f) < \infty$. Pour la première intégrale il suffit de remarquer que nous avons l'inégalité $\frac{|x|}{1+x^2} \leq |x|$, donc d'après la condition d'intégrabilité vérifiée par la bimesure il s'ensuit que cette intégrale est aussi finie. ■

Chapitre 4

Principes d'invariance dans $L^2[0,1]$

Nous allons maintenant nous intéresser aux applications des mesures aléatoires à signe. Dans ce chapitre nous nous servirons des mesures aléatoires pour déduire des convergences en loi dans l'espace $L^2[0,1]$. Pour cela nous ne considérerons que des mesures aléatoires à signe bornées. En effet ce type de mesures aléatoires a des propriétés convenables qui nous permettront d'interpréter les mesures aléatoires comme des variables aléatoires dans l'espace $L^2[0,1]$, selon une procédure qu'on expliquera plus tard. Cette approche de la convergence en loi sur $L^2[0,1]$ nous permettra d'établir des principes d'invariance faibles, c'est à dire des versions du théorème de Donsker (voir [5], par exemple), avec la convergence dans $L^2[0,1]$ au lieu de $D[0,1]$. L'utilisation des mesures aléatoires nous facilitera l'étude de la compacité relative des suites, réduisant cette étude à l'étude des propriétés de certaines formes quadratiques sur $\mathbb{R}^k$, quand $k \longrightarrow \infty$. Il faut remarquer que cette simplification n'est pas générale, mais spécifique du problème auquel on s'intéresse. En fait, comme on le verra dans ce chapitre, les propriétés qu'on arrive à établir dépendent beaucoup de la forme des éléments aléatoires que nous définissons.

4.1 Mesures aléatoires et espaces de Hilbert

Cette section est un résumé des résultats concernant l'introduction de topologies métrisables dans l'espace des mesures à signe bornées. Pour un exposé plus complet de ce sujet nous renvoyons à Suquet [44], que nous suivrons ici.

Définition 4.1.1 *Une fonction* $\mathcal{K}$:S×S⟶IR *est dite un noyau reproduisant si*

- $\forall_{s,t \in \mathrm{S}},\ \mathcal{K}(s,t) = \mathcal{K}(t,s)$;
- $\forall_{n \in \mathrm{N}}\ \forall_{a_1,\ldots,a_n \in \mathbb{R}}\ \forall_{s_1,\ldots,s_n \in \mathrm{S}}\ \sum_{i,j=1}^{n} a_i a_j \mathcal{K}(s_i,s_j) \geq 0.$

On sait qu'à un tel noyau nous pouvons associer un espace de Hilbert de fonctions réelles sur S, que nous appellerons $\mathrm{H}_{\mathcal{K}}$, vérifiant

- $\forall_{s \in \mathrm{S}},\ \mathcal{K}(s,\cdot) \in \mathrm{H}_{\mathcal{K}}$,
- $\forall_{s \in \mathrm{S}}\ \forall_{f \in \mathrm{H}_{\mathcal{K}}},\ f(s) =< f, \mathcal{K}(s,\cdot) >_{\mathcal{K}}$,

où $< \cdot,\cdot >_{\mathcal{K}}$ représente le produit scalaire de l'espace de Hilbert $\mathrm{H}_{\mathcal{K}}$ (Aronszajn [4]). Nous allons supposer en plus que cet espace de Hilbert $\mathrm{H}_{\mathcal{K}}$ est séparable. Cette propriété implique une forme particulière pour le noyau reproduisant $\mathcal{K}$.

Théorème 4.1.2 (Guilbart [15]) *L'espace autoreproduisant $H_{\mathcal{K}}$ associé au noyau reproduisant borné $\mathcal{K}$ est séparable si et seulement s'il existe une suite de fonctions $\{f_n\}$ telle que*

$$\mathcal{K}(s,t) = \sum_{n=0}^{\infty} f_n(s) f_n(t),$$

où la série converge simplement sur S×S. *Alors les fonctions f_n sont bornées et la convergence est uniforme en s, pour tout t* ∈S *fixé.*

Dans la suite nous supposerons de plus que le noyau reproduisant est borné et mesurable. Le fait pour le noyau être borné n'est pas indispensable mais constitue une commodité technique.

Définissons alors l'application

$$\begin{array}{rccl} \varphi : & \mathcal{M}_b & \longrightarrow & \mathrm{H}_{\mathcal{K}} \\ & \mu & \hookrightarrow & \varphi(\mu)(s) = \int_{\mathrm{S}} \mathcal{K}(s,t)\, \mu(dt). \end{array}$$

D'après Suquet [44], page 10, cette application est injective. De plus, nous pouvons écrire, tenant compte des hypothèses faites sur $\mathcal{K}$ et les fonctions f_n,

$$\varphi(\mu) = \sum_{n=0}^{\infty} \left(\int_{\mathbf{S}} f_n \, d\mu\right) f_n.$$

L'application φ nous permet d'exprimer le produit scalaire de $\mathrm{H}_{\mathcal{K}}$ en fonction d'une intégrale, si au moins un des éléments dont on calcule le produit scalaire est dans $\varphi(\mathcal{M}_b)$ ([44], page 14):

$$\forall_{f \in \mathrm{H}_{\mathcal{K}}} \forall_{\mu \in \mathcal{M}_b} \quad < f, \varphi(\mu) >_{\mathcal{K}} = \int_{\mathbf{S}} f \, d\mu. \tag{4.1}$$

Remarquons que l'application φ est mesurable, donc φ permet d'interpréter une mesure aléatoire bornée comme une variable aléatoire à valeurs dans l'espace de Hilbert $\mathrm{H}_{\mathcal{K}}$.

Evidemment, nous pourrions approfondir cette étude générale des liens qui dérivent de la fonction φ, mais nous allons nous arrêter ici car nous avons rappelé les résultats dont nous avons besoin.

4.2 Un noyau particulier et l'espace $\mathbf{L}^2[0,1]$

Etudions maintenant un cas particulier avec plus de détail. Supposons $\mathbf{S} = [0,1]$, et $\mathcal{K}(s,t) = 1 - \max(s,t)$. Comme $\mathbf{S}$ est un espace compact, $\mathcal{M}_b = \mathcal{M}_c$. L'espace autoreproduisant $\mathrm{H}_{\mathcal{K}}$ est de la forme

$$\mathrm{H}_{\mathcal{K}} = \{f(u) = \int_u^1 g(t)\, dt, g \in \mathrm{L}^2[0,1]\ \},$$

et le produit scalaire sur $\mathrm{H}_{\mathcal{K}}$ vérifie

$$< f', f'' >_{\mathcal{K}} = \int_0^1 g'(t) g''(t)\, dt, \tag{4.2}$$

où

$$f'(u) = \int_u^1 g'(t)\, dt \quad \text{et} \quad f''(u) = \int_u^1 g''(t)\, dt.$$

Considérons l'application

$$\begin{array}{rcl} \psi : \mathrm{L}^2[0,1] & \longrightarrow & \mathrm{H}_{\mathcal{K}} \\ g & \hookrightarrow & \int_u^1 g(t)\, dt. \end{array}$$

D'après les commentaires précédents on vérifie que cette application ψ définit un isomorphisme isométrique entre $L^2[0,1]$ et $H_{\mathcal{K}}$. Donc, nous pouvons étudier des problèmes de convergence de suites de variables aléatoires dans $L^2[0,1]$ ou dans $H_{\mathcal{K}}$, à notre convenance.

En rappelant les résultats de la section précédente et en calculant les expressions pour le cas particulier présent, nous obtenons pour la fonction φ une expression simple. D'abord remarquons que nous pouvons écrire le noyau de la façon suivante

$$\mathcal{K}(s,t) = \int_0^1 \mathbb{1}_{[s,1]}(u)\mathbb{1}_{[t,1]}(u)\, du = \int_s^1 \mathbb{1}_{[t,1]}(u)\, du.$$

Alors, pour toute mesure μ sur [0,1]

$$\varphi(\mu)(s) = \int_0^1 \mathcal{K}(s,t)\,\mu(dt) = \int_0^1 \int_s^1 \mathbb{1}_{[t,1]}(u)\, du\, \mu(dt) =$$
$$\int_s^1 \int_0^1 \mathbb{1}_{[0,u]}(t)\,\mu(dt)\, du = \int_s^1 \mu[0,u]\, du.$$

De plus, en supposant $f(u) = \int_u^1 g(t)\, dt$, $g \in L^2[0,1]$ (donc $f \in H_{\mathcal{K}}$), pour toute mesure μ, d'après (4.2)

$$< f, \varphi(\mu) >_{\mathcal{K}} = \int_0^1 g(u)\mu[0,u]\, du = \int_0^1 f(t)\,\mu(dt).$$

ce qui peut être déduit tenant compte de (4.1) ou en utilisant la formule d'intégration par parties pour les intégrales de Stieltjes.

Considérons maintenant une suite de variables aléatoires réelles $\{X_n\}$ et définissons la mesure aléatoire

$$\xi_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \delta_{\frac{i}{n}}.$$

Tenant compte des applications φ et ψ cette mesure aléatoire peut être interprétée comme une variable aléatoire à valeurs dans $H_{\mathcal{K}}$ ou dans $L^2[0,1]$. Comme variable aléatoire à valeurs dans $L^2[0,1]$ la mesure aléatoire ξ_n est interprétée comme

$$\psi^{-1}\varphi(\xi_n)(u) = -\xi_n[0,u] = -\frac{1}{\sqrt{n}} \sum_{i \leq nu} X_i = -\frac{1}{\sqrt{n}} S_{[nu]},$$

où S_k représente la somme $X_1 + \cdots + X_k$, et $[x]$ représente le plus grand entier plus petit ou égale à x. Donc, la mesure aléatoire ξ_n interprétée comme variable aléatoire à valeurs dans $L^2[0,1]$ est, à un signe près, la fonction intervenant dans le théorème de Donsker. Pour cette raison nous appellerons la mesure aléatoire ξ_n, eventuellement divisée par une constante, la mesure aléatoire de Donsker.

4.3 Compacité relative en loi de $\{\xi_n[0, u]\}$ dans $L^2[0,1]$

Considérons les fonctions

$$g_n(u) = \cos[(n + \frac{1}{2})\pi u], \quad n \in \mathbb{N}$$

et les nombres réels

$$\lambda_n = ((n + \frac{1}{2})\pi)^{-2}, \quad n \in \mathbb{N}.$$

Alors les fonctions $\sqrt{\lambda_n} g_n$ forment une base orthonormée de l'espace $L^2[0,1]$ telle que

$$\mathcal{K}(s,t) = \sum_{n=0}^{\infty} \lambda_n g_n(s) g_n(t)$$

et la série converge uniformément. De plus, remarquons que la série $\sum_{n=0}^{\infty} \lambda_n$ converge. Il est possible de choisir une base orthonormale de l'espace autoreproduisant $H_{\mathcal{K}}$ constituée par des fonctions de la forme $G_n = \sqrt{\lambda_n} g_n$ (évidemment cette base ne correspond pas à la base $\{\sqrt{\lambda_n} g_n\}$ de $L^2[0,1]$ par l'isomorphisme ψ, mais le choix de la base $\{G_n\}$ simplifie beaucoup les calculs que l'on va effectuer).

Théorème 4.3.1 *Supposons que les variables aléatoires réelles $\{X_n\}$ vérifient la condition*

$$\sum_{k=1}^{n} \mathbf{E}(X_k^2) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} |\mathbf{E}(X_i X_j)| = O(n) \tag{4.3}$$

Alors la suite $\{\frac{1}{\sqrt{n}} S_{[nu]}\}$ est faiblement relativement compacte dans $L^2[0,1]$.

__Démonstration__ : Pour démontrer la compacité relative faible de la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ nous allons utiliser la mesure aléatoire de Donsker comme une variable aléatoire à valeurs dans $\mathrm{H}_{\mathcal{K}}$. Alors, d'après Parthasarathy [38], page 154, il suffit de démontrer

$$\lim_{N\to\infty} \sup_{n\in\mathbb{N}} \int_{\mathrm{H}_{\mathcal{K}}} R_N(F)\,\mathbf{P}_n(dF) = 0, \tag{4.4}$$

où $R_N(F) = \sum_{i=N}^{\infty} < F, G_i >_{\mathcal{K}}^2$ et où $\mathbf{P}_n$ est la loi de probabilité de ξ_n interprétée comme variable aléatoire à valeurs dans $\mathrm{H}_{\mathcal{K}}$, tenant compte de l'isomorphisme entre les espaces de Hilbert $\mathrm{H}_{\mathcal{K}}$ et $\mathrm{L}^2[0,1]$. En calculant l'intégrale dans (4.4) nous obtenons

$$\int_{\mathrm{H}_{\mathcal{K}}} R_N(F)\,\mathbf{P}_n(dF) = \int_{\mathrm{H}_{\mathcal{K}}} \sum_{i=N}^{\infty} < F, G_i >_{\mathcal{K}}^2 \;\mathbf{P}_n(dF) =$$

$$\mathbf{E}(\sum_{i=N}^{\infty} < \varphi(\xi_n), G_i >_{\mathcal{K}}^2) = \mathbf{E}[\sum_{i=N}^{\infty} (\int G_i(u)\,\xi_n(du))^2] =$$

$$\sum_{i=N}^{\infty} \frac{\lambda_i}{n}[\sum_{k,l=1}^{n} (g_i(\frac{k}{n})g_i(\frac{l}{n})\mathbf{E}(X_k X_l))] = \sum_{i=N}^{\infty} \frac{\lambda_i}{n}\tilde{g}_{in}^t \Gamma_n \tilde{g}_{in},$$

où $\tilde{g}_{in}^t = (g_i(\frac{1}{n}), \cdots, g_i(1))$ et $\Gamma_n = (\mathbf{E}(X_i X_n))_{i,j=1,\ldots,n}$, qui coïncide avec la matrice des covariances des variables aléatoires $X_1, \ldots, X_n$ si elles sont centrées. Posons $O_n = \tilde{g}_{in}^t \Gamma_n \tilde{g}_{in}$. Alors, en remarquant que $||\tilde{g}_{in}||_\infty \le 1$, nous avons

$$O_n \le \max_{||\tilde{x}||_\infty \le 1} \tilde{x}^t \Gamma_n \tilde{x} \le \sum_{k=1}^{n} \mathbf{E}(X_k^2) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} |\mathbf{E}(X_i X_j)|.$$

Donc, d'après la condition (4.3) il s'ensuit $O_n = O(n)$. Cela veut dire qu'il existe une constante $L > 0$ telle que $O_n \le nL$, $n \in\mathbb{N}$. Alors

$$\sum_{i=N}^{\infty} \frac{\lambda_i}{n}\tilde{g}_{in}^t \Gamma_n \tilde{g}_{in} \le L \sum_{i=N}^{\infty} \lambda_i$$

et cette série converge vers zéro quand $N \longrightarrow \infty$ car la série $\sum_{n=0}^{\infty} \lambda_n$ est convergente. Donc la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ est faiblement relativement compacte dans $\mathrm{L}^2[0,1]$. ■

La condition (4.3) nous permet de trouver des conditions plus simples dans des cas particuliers.

Corollaire 4.3.2 *Supposons que les variables aléatoires X_n soient stationnaires et vérifient*

$$\sum_{k=0}^{\infty} |\mathbf{E}(X_0 X_k)| < \infty.$$

Alors la suite $\{\frac{1}{\sqrt{n}} S_{[nu]}\}$ est faiblement relativement compacte dans $L^2[0,1]$.

Démonstration : Dans ce cas

$$\sum_{k=1}^{n} \mathbf{E}(X_k^2) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} |\mathbf{E}(X_i X_j)| =$$

$$= n\mathbf{E}(X_0^2) + 2\sum_{k=1}^{n-1} (n-k)|\mathbf{E}(X_0 X_k)| \leq$$

$$\leq n\mathbf{E}(X_0^2) + 2n\sum_{k=1}^{n-1} |\mathbf{E}(X_0 X_k)| \leq$$

$$\leq 2n\sum_{k=0}^{\infty} |\mathbf{E}(X_0 X_k)|,$$

donc la condition (4.3) est vérifiée et la compacité relative faible s'ensuit d'après le théorème précédent. ■

Remarquons que cette condition est utilisée par Billingsley [5] pour déduire un principe d'invariance dans l'espace D[0,1], en supposant aussi la stationnarité et, en plus, une condition de mélangeance plus forte que celle dont nous aurons besoin plus tard.

Considérons maintenant la condition

$$\sup_{n\in\mathbb{N}} \frac{1}{n}\mathbf{E}(S_n^2) < \infty. \tag{4.5}$$

Cette condition n'est pas trop forte quand on s'intéresse aux principes d'invariance. En effet, en étudiant ce type de problèmes il est habituel de supposer la convergence de la suite $\frac{1}{n}\mathbf{E}(S_n^2)$, donc la condition (4.5) est verifiée a fortiori. Cette condition nous permet d'énoncer quelques simplifications du théorème précédent.

Corollaire 4.3.3 *Supposons que les variables aléatoires $\{X_n\}$ vérifient (4.5). Alors, si au moins une des conditions suivantes est vérifiée*

1. $\mathbf{E}(X_k X_l) \geq 0, \quad k, l \in \mathbb{N}$;
2. *les variables X_n sont non correllées et les moyennes sont non négatives,*

la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ est faiblement relativement compacte dans $L^2[0,1]$.

Démonstration : 1. Dans ce cas on vérifie facilement

$$\sum_{k=1}^{n} \mathbf{E}(X_k^2) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} |\mathbf{E}(X_iX_j)| = \mathbf{E}(S_n^2),$$

donc, tenant compte de (4.5), il s'ensuit que les variables aléatoires $\{X_n\}$ vérifient (4.3), d'où la compacité relative faible de $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ dans $L^2[0,1]$.

2. Si les X_n sont non correllées, alors

$$\sum_{k=1}^{n} \mathbf{E}(X_k^2) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} |\mathbf{E}(X_iX_j)| = \mathbf{E}(S_n^2),$$

donc, la compacité relative faible de $\{X_n\}$ dans $L^2[0,1]$ s'ensuit comme dans le cas précédent. ∎

Remarquons que, en général, la condition (4.5) semble ne pas être suffisante pour déduire la compacité relative faible. La raison pour laquelle cette condition n'est pas suffisante est liée au fait que $\mathbf{E}(S_n^2)$ ne tient compte que des covariances des variables $\{X_n\}$, tandis que, pour déduire la compacité relative faible de $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ il est nécessaire d'avoir des conditions sur les valeurs absolues des covariances.

Finalement nous présentons une condition suffisante pour la compacité relative faible en faisant des hypothèses sur les valeurs propres des matrices de covariances Γ_n.

Théorème 4.3.4 *Soit Λ_n la plus grande valeur propre de Γ_n. Supposons vérifiée la condition (4.5) et*

1. $\inf_{n\in\mathbf{N}} \frac{1}{n}\mathbf{E}(S_n^2) > 0$;
2. *il existe $\Lambda > 0$ tel que $\sup_{n\in\mathbf{N}} \Lambda_n \leq \Lambda$.*

Alors la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ est faiblement relativement compacte dans $L^2[0,1]$.

Démonstration : Nous reprenons l'idée de la démonstration du théorème 4.3.1. Nous pouvons majorer l'intégrale de (4.4) par

$$\sum_{i=N}^{\infty} \frac{\lambda_i}{n} \max_{\|\tilde{x}\|_\infty \leq 1} \tilde{x}^t\Gamma_n\tilde{x}.$$

Nous allons démontrer que dans les conditions de ce théorème la quantité $\max_{\|\tilde{x}\|_\infty \leq 1} \tilde{x}^t \Gamma_n \tilde{x}$ est comparable avec $\mathbf{E}(S_n^2) = u^t \Gamma_n u$, $u^t = (1, \ldots, 1)$. Soit $c > 0$ fixé. Pour tout $n \in \mathbb{N}$ la longueur du plus petit demi-axe de l'hyper-ellipsoïde $\tilde{x}^t \Gamma_n \tilde{x} = cu^t \Gamma_n u$ est donné par

$$\sqrt{\frac{cu^t\Gamma_n u}{\Lambda_n}} \geq \sqrt{\frac{cu^t\Gamma_n u}{\Lambda}}.$$

Si l'hyper-cube $\|\tilde{x}\|_\infty \leq 1$ est contenu dans cet hyper-ellipsoïde alors le maximum de la fonction $\tilde{x}^t \Gamma_n \tilde{x}$ pour $\|\tilde{x}\|_\infty \leq 1$ est inférieur ou égal à $cu^t \Gamma_n u$. Pour avoir l'inclusion il suffit d'avoir

$$\frac{cu^t\Gamma_n u}{\Lambda} \geq n,$$

pour tout $n \in \mathbb{N}$. Pour accomplir cette condition il suffit de choisir la constante c telle que

$$\frac{1}{c} \leq \frac{1}{\Lambda} \inf_{n \in \mathbb{N}} \frac{u^t \Gamma_n u}{n} = \frac{1}{\Lambda} \inf_{n \in \mathbb{N}} \frac{1}{n} \mathbf{E}(S_n^2).$$

Alors

$$\sum_{i=N}^{\infty} \frac{\lambda_i}{n} \max_{\|\tilde{x}\|_\infty \leq 1} \tilde{x}^t \Gamma_n \tilde{x} \leq \sum_{i=N}^{\infty} \frac{\lambda_i}{n} \mathbf{E}(S_n^2) \leq$$
$$\sup_{n \in \mathbb{N}} \frac{1}{n} \mathbf{E}(S_n^2) \sum_{i=N}^{\infty} \lambda_i,$$

ce qui converge vers zéro, d'après le choix des constantes λ_i. Donc, nous avons bien la compacité relative faible de $\{\frac{1}{\sqrt{n}} S_{[nu]}\}$ dans $L^2[0,1]$. ■

Remarquons que comme Γ_n est une matrice symétrique et définie positive, les valeurs propres de Γ_n sont toutes réelles et positives. De plus, les conditions de ce théorème semblent ne pas être, en général, comparables avec la condition (4.3).

La condition $\sup_{n \in \mathbb{N}} \Lambda_n < \infty$ implique l'existence d'une certaine régularité dans la distribution des vecteurs aléatoires $(X_1, \ldots, X_n)$, pour $n \in \mathbb{N}$. En fait, si $\sup_{n \in \mathbb{N}} \Lambda_n = \infty$, cela signifierait que quand $n \longrightarrow \infty$, il existe une direction dans $\mathbb{R}^n$ telle que la variance de la projection sur le sous-espace determiné par cette direction est croissante et la suite des variances converge vers $+\infty$.

Donc, pour n suffisamment grand, la distribution des variables aléatoires devient de plus en plus chaotique. Evidemment, si nous voulons démontrer un résultat de compacité relative on doit s'attendre à une certaine régularité.

La condition sup $\Lambda_n < \infty$ n'est pas facilement vérifiable. Tenant compte des liens entre les valeurs propres d'une matrice et les éléments définissant la matrice nous pouvons réduire cette condition à une autre plus facilement utilisable.

Corollaire 4.3.5 *Supposons que les variables aléatoires vérifient la condition (4.5) et*

1. $\inf_{n\in\mathbb{N}} \frac{1}{n}\mathbf{E}(S_n^2) > 0$,
2. $\sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} |\mathbf{E}(X_iX_j)| < \infty$.

Alors la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ *est faiblement relativement compacte dans* $L^2[0,1]$.

<u>**Démonstration**</u> : En utilisant les majorations par des disques de Gershgorin (Fröberg [11], page 52, par exemple) pour les valeurs propres nous obtenons

$$\Lambda_n \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |\mathbf{E}(X_iX_j)|,$$

donc

$$\sup_{n\in\mathbb{N}} \Lambda_n \leq \sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \sum_{j=1}^{n} |\mathbf{E}(X_iX_j)| = \sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} |\mathbf{E}(X_iX_j)| < \infty,$$

d'où il s'ensuit la compacité relative faible de la suite $\{\frac{1}{\sqrt{n}}S_{[nu]}\}$ dans $L^2[0,1]$.
■

Remarquons que pour déduire les résultats de compacité relative nous n'avons pas posé d'hypothèse sur les moyennes des variables aléatoires. En effet, il n'est pas nécessaire d'imposer que les espérances soient nulles car elles n'interviennent qu'indirectement dans les calculs.

4.4 Principes d'invariance faibles

Dans cette section nous supposerons que les variables aléatoires X_n sont centrées, $\mathbf{E}(X_n) = 0$, pour tout $n \in\mathbb{N}$, et

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{E}(S_n^2) = \sigma^2 > 0. \tag{4.6}$$

Soit W une version du mouvement brownien à valeurs dans L^2[0,1]. Nous allons étudier la convergence en loi dans L^2[0,1] de la suite $\{\frac{1}{\sigma\sqrt{n}}S_{[nu]}\}$ vers W. A cause de la normalisation éffectuée nous considérerons, au cours de cette section, comme mesure aléatoire de Donsker

$$\xi_n = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n} X_i\delta_{\frac{i}{n}}.$$

Evidemment, tous les résultats de la section précédente restent vrais, car le fait de diviser par σ ne change pas les convergences vers zéro démontrées.

Pour démontrer la convergence en loi dans L^2[0,1] on doit, comme il s'agit d'un espace de Hilbert, démontrer la compacité relative faible de la suite et prouver la convergence de la suite des fonctions caractéristiques. En ce qui concerne la caractérisation de la compacité relative faible nous pouvons nous appuyer sur les résultats de la section précédente. Calculons maintenant la fonction caractéristique de $\xi_n[0,u]$, c'est à dire, la mesure aléatoire de Donsker comme variable aléatoire à valeurs dans L^2[0,1]

$$\forall_{f\in \mathrm{L}^2[0,1]}\ \mathbf{E}(e^{i<f,\xi_n[0,u]>}) = \mathbf{E}(e^{i\int_0^1 f(u)\xi_n[0,u]\,du}) = \mathbf{E}(e^{i\int_0^1 F(u)\,\xi_n(du)}),$$

où $F(u) = \int_u^1 f(t)\,dt$, d'après les propriétés rappelées dans la première section de ce chapitre. Or

$$\int_0^1 f(u)\,\xi_n(du) = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n} F(\frac{i}{n})X_i,$$

donc, la fonction caractéristique dépend des valeurs de la fonction F dans certains points déterminés et pas du comportement global de la fonction. Ce fait rend impossible de démontrer directement cette convergence dans le cas général. Néamoins, dans le cas particulier de variables aléatoires indépendantes et de même loi il est possible faire cette démonstration directement [24]. Alors, pour démontrer la convergence en loi $\frac{1}{\sigma\sqrt{n}}S_{[nu]}\xrightarrow{d}W$ dans L^2[0,1], nous allons prouver la convergence en loi des distributions marginales

$$(\xi_n[0,u_1],\ldots,\xi_n[0,u_k])\xrightarrow{d}(W(u_1),\ldots,W(u_k)),$$

pour tout choix $u_1,\ldots,u_k \in$[0,1] et $k \in \mathbb{N}$.

Si nous supposons que les variables aléatoires réelles X_n vérifient le théorème central limite, alors, tenant compte de la condition (4.6), on déduit immédiatement la convergence en loi des marges de dimension un,

$$\xi_n[0,u] = \frac{1}{\sigma\sqrt{n}} S_{[nu]} \xrightarrow{d} W(u), \quad u \in [0,1].$$

Maintenant, pour déduire la convergence en loi des distributions marginales de dimension plus grande que un nous avons besoin de quelques conditions de plus. D'abord introduisons les coefficients de mélangeance suivants:

$$\alpha_n(k) = \begin{cases} \sup\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|,\ A \in \sigma(X_i,\ 1 \le i \le m), \\ \qquad B \in \sigma(X_i,\ m+k \le i \le n), i \le m \le n-k\} & k = 1, \ldots, n-1 \\ 0 & k \ge n, \end{cases}$$

et

$$\alpha(k) = \sup_{n \in \mathbb{N}} \alpha_n(k), \quad k \in \mathbb{N}.$$

Définition 4.4.1 *Une suite de variables aléatoires $\{X_n\}$ est α-mélangeante si $\lim_{k\to\infty} \alpha(k) = 0$.*

Théorème 4.4.2 *Supposons que la suite des variables aléatoires réelles $\{X_n\}$ en plus des propriétés déjà supposées, vérifie le théorème central limite, est α-mélangeante et*

$$\sup_{m,n \in \mathbb{N}} \{\frac{1}{n}\mathbf{E}(S_{m+n} - S_m)^2\} < \infty. \tag{4.7}$$

Alors, pour tout choix de $k \in \mathbb{N}$ et $u_1, \ldots, u_k \in [0,1]$

$$(\xi_n[0,u_1], \ldots, \xi_n[0,u_k]) \xrightarrow{d} (W(u_1), \ldots, W(u_k)).$$

Démonstration : Nous allons suivre Herrndorf [19] pour démontrer cette convergence. D'après le théorème de Prokhorov et du fait que les distributions marginales de dimension un sont faiblement relativement compactes on déduit immédiatement que l'ensemble $\{(\xi_n[0,u_1], \ldots, \xi_n[0,u_k]),\ n \in \mathbb{N}\}$ est faiblement relativement compact. En fait, étant donné $\varepsilon > 0$, il existe des compacts K_i de $\mathbb{R}$ tels que

$$\mathbf{P}(\xi_n[0,u_i] \notin K_i) < \frac{\varepsilon}{k},$$

donc, en posant $K = K_1 \times \cdots \times K_k$, nous avons

$$\mathbf{P}((\xi_n[0,u_1],\ldots,\xi_n[0,u_k]) \notin K) \leq \sum_{i=1}^{k} \mathbf{P}(\xi_n[0,u_i] \notin K_i) < \varepsilon.$$

Alors, la suite $\{(\xi_n[0,u_1],\ldots,\xi_n[0,u_k])\}$ a une sous-suite convergente en loi vers une loi de probabilité $\mathbf{Q}$ sur $\mathbb{R}^k$. Désignons par π_i la projection associée à la coordonnée u_i, et choisissons une suite $\{r_n\}$ de réels non négatifs telle que $r_n \longrightarrow 0$ et $nr_n \longrightarrow \infty$. Du fait que nous avons les convergences en loi des distributions marginales de dimension un il s'ensuit que $\mathbf{Q}\pi_i^{-1}$ est une loi de probabilité normale centrée et de variance u_i. D'autre part

$$\mathbf{E}(\xi_n[0,u_i+r_n]-\xi_n[0,u_i])^2 = \tfrac{1}{\sigma^2 n}\mathbf{E}(S_{[nu_i]+[nr_n]+z}-S_{[nu_i]})^2 =$$

$$\tfrac{[nr_n]+z}{\sigma^2 n}\tfrac{1}{[nr_n]+z}\mathbf{E}(S_{[nu_i]+[nr_n]+z}-S_{[nu_i]})^2,$$

où z prend les valeurs 0 ou 1 de façon que $[nu_i+nr_n] = [nu_i]+[nr_n]+z$. D'après (4.7),

$$\frac{1}{[nr_n]+z}\mathbf{E}(S_{[nu_i]+[nr_n]+z}-S_{[nu_i]})^2$$

est borné. Comme, d'après le choix de la suite $\{r_n\}$, $\frac{[nr_n]+z}{\sigma^2 n} \longrightarrow 0$ nous obtenons

$$\mathbf{E}(\xi_n[0,u_i+r_n]-\xi_n[0,u_i])^2 \longrightarrow 0.$$

Alors, il s'ensuit que $\mathbf{Q}(\pi_1,\pi_2-\pi_1,\ldots,\pi_k-\pi_{k-1})^{-1}$ est la limite en loi d'une sous-suite de

$$\{(\xi_n[0,u_1],\xi_n[0,u_2]-\xi_n[0,u_1+r_n],\ldots,\xi_n[0,u_k]-\xi_n[0,u_{k-1}+r_n])\}.$$

Considérons A_1, A_2 des boréliens de [0,1]. Alors

$$|\mathbf{Q}(\pi_1,\pi_2-\pi_1)^{-1}(A_1\times A_2)-\mathbf{Q}\pi_1^{-1}(A_1)\mathbf{Q}(\pi_2-\pi_1)^{-1}(A_2)| =$$

$$\lim_{n\to\infty} |\mathbf{P}(\xi_n[0,u_1]\in A_1, \xi_n[0,u_2]-\xi_n[0,u_1+r_n]\in A_2)-$$

$$-\mathbf{P}(\xi_n[0,u_1]\in A_1)\mathbf{P}(\xi_n[0,u_2]-\xi_n[0,u_1+r_n]\in A_2) \leq$$

$$\leq \lim_{n\to\infty} \alpha_n([nr_n]+z) \leq \lim_{n\to\infty} \alpha([nr_n]+z) = 0,$$

en remarquant que

$$\{\xi_n[0,u_1] \in A_1\} \in \sigma(X_i, 1 \leq i \leq [nu_1])$$

et

$$\{\xi_n[0,u_2] - \xi_n[0,u_1 + r_n] \in A_2\} \in \sigma(X_i, [nu_1 + nr_n] \leq i \leq [nu_2]),$$

donc les applications π_1 et $\pi_2 - \pi_1$ sont indépendantes par rapport à la loi de probabilité $\mathbf{Q}$. De la même façon on déduit que les applications $\pi_1, \pi_2 - \pi_1, \ldots, \pi_k - \pi_{k-1}$ sont indépendantes par rapport à la loi de probabilité $\mathbf{Q}$. Alors il s'ensuit que $\mathbf{Q}$ est la loi de probabilité du vecteur aléatoire $(W(u_1), \ldots, W(u_k))$. ∎

Maintenant nous disposons de tous les résultats dont nous avons besoin pour démontrer des principes d'invariance. En fait nous avons résolu les deux problèmes qui se pose quand on essaye de démontrer des résultats de convergence en loi dans un espace de Hilbert. Il suffit maintenant de rassembler les conditions de façon convenable.

Théorème 4.4.3 *Supposons que les variables aléatoires réelles $\{X_n\}$ sont centrées et vérifient la condition (4.6), le théorème central limite, sont α-mélangeantes et satisfont au moins une des conditions suivantes:*

1. *les conditions (4.3) et (4.7);*
2. *la condition (4.7) et sont non correllées;*
3. $\mathbf{E}(X_k X_l) \geq 0,\ k, l \in \mathbb{N}$ *et (4.7);*
4. *les variables aléatoires sont stationnaires et vérifient la condition*

$$\sum_{k=0}^{\infty} |\mathbf{E}(X_0 X_k)| < \infty.$$

Alors la suite $\{\frac{1}{\sigma\sqrt{n}} S_{[nu]}\}$ converge en loi vers W dans $L^2[0,1]$.

<u>**Démonstration**</u> : Les résultats sont immédiats au vu du théorème 4.4.2 et du théorème 4.3.1 et ses corollaires. ∎

Comme corollaire de ce théorème nous déduisons un résultat de convergence concernant les intégrales stochastiques.

Corollaire 4.4.4 *Supposons vérifiées les conditions du théorème précédent. Soit $f \in H_{\mathcal{K}}$, alors*

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i f(\frac{i}{n}) \xrightarrow{d} \int_0^1 f(u)\, W(du).$$

Démonstration : D'après les hypothèses nous déduisons la convergence

$$\frac{1}{\sigma\sqrt{n}} S_{[nu]} \xrightarrow{d} W$$

dans $L^2[0,1]$. Alors, les fonctions caractéristiques correspondantes convergent, soit

$$\mathbf{E}(e^{i<g,\xi_n[0,u]>}) \longrightarrow \mathbf{E}(e^{i<g,W>}),$$

pour $g \in L^2[0,1]$, et où $< \cdot,\cdot >$ représente le produit scalaire dans $L^2[0,1]$. D'après l'isométrie ψ entre $L^2[0,1]$ et $H_{\mathcal{K}}$ et le plongement φ de l'espace des mesures sur **S** dans $H_{\mathcal{K}}$, et ses propriétés, il s'ensuit, en posant $f(u) = \int_u^1 g(t)\, dt$,

$$< g, \xi_n[0,u] > = < f, \int_u^1 \xi_n[0,u]\, dt >_{\mathcal{K}} = < f, \varphi(\xi_n) >_{\mathcal{K}} = \int_0^1 f(u)\, \xi_n(du).$$

En rappelant l'expression de la mesure aléatoire de Donsker, nous obtenons

$$\int_0^1 f(u)\, \xi_n(du) = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i f(\frac{i}{n}).$$

D'autre part

$$< g, W > = \int_0^1 f(u) W(u)\, du = \int_0^1 f(u)\, W(du),$$

d'après les liens entre les fonctions f et g, et en utilisant les propriétés élémentaires de l'intégrale stochastique (voir Hida [20]). Alors, la convergence des fonctions caractéristiques démontre le résultat énoncé. ■

Nous pouvons encore modifier les conditions du théorème précédent en profitant du fait que les variables aléatoires sont supposées α-mélangeantes pour déduire que les variables vérifient le théorème central limite. Pour cela introduisons la condition

$$\sup_{m,n\in\mathbb{N}} \{\frac{1}{\sqrt{n}} \mathbf{E}|S_{m+n} - S_m|^{2+\varepsilon}\} < \infty, \tag{4.8}$$

pour $\varepsilon > 0$ fixé. Cette condition est plus forte que la condition (4.7). En fait, en utilisant l'inégalité de Hölder,

$$\frac{1}{n}\mathbf{E}(S_{m+n} - S_m)^2 \leq [\mathbf{E}|\frac{S_{m+n}-S_m}{\sqrt{n}}|^{2+\varepsilon}]^{\frac{2}{2+\varepsilon}} \leq$$

$$(\frac{1}{\sqrt{n}}\mathbf{E}|S_{m+n} - S_m|^{2+\varepsilon})^{\frac{2}{2+\varepsilon}}.$$

Nous énonçons seulement la modification du théorème précédent correspondant à la condition 1, les autres s'obtenant de façon évidente.

Théorème 4.4.5 *Supposons que les variables aléatoires réelles $\{X_n\}$ soient centrées, vérifient les conditions (4.3), (4.6), (4.8),*

$$\sum_{k=0}^{\infty} \sup_{|l-m|\geq k} \mathbf{E}(X_l X_m) < \infty, \tag{4.9}$$

et soient α-mélangeantes. Alors la suite $\{\frac{1}{\sigma\sqrt{n}}S_{[nu]}\}$ converge en loi vers W dans $L^2[0,1]$.

<u>**Démonstration**</u> : Les conditions (4.6), (4.8), (4.9) et le fait que les variables sont α-mélangeantes nous permet de conclure que les variables aléatoires $\{X_n\}$ vérifient le théorème central limite, d'après [48], corollaire 2.1. Comme la condition (4.8) implique (4.7) les conditions du théorème 4.4.2 sont remplies. Finalement (4.3) nous assure la compacité relative faible de la suite. Donc, la convergence s'ensuit. ■

Il est intéressant de remarquer le cas particulier où les variables aléatoires $\{X_n\}$ sont supposées non correllées. En fait, dans ce cas la condition (4.9) est trivialement vérifiée:

$$\sum_{k=0}^{\infty} \sup_{|l-m|\geq k} \mathbf{E}(X_l X_m) = \mathbf{E}(X_0^2).$$

Nous pouvons donc énoncer

Corollaire 4.4.6 *Supposons que les variables aléatoires réelles $\{X_n\}$ soient centrées, vérifient les conditions (4.6), (4.8), soient α-mélangeantes et non correllées. Alors la suite $\{\frac{1}{\sigma\sqrt{n}}S_{[nu]}\}$ converge en loi vers W dans $L^2[0,1]$.*

Remarquons qu'au cours de ce chapitre nous n'avons jamais posé de condition sur les coefficients de mélangeance, sauf la α-mélangeance, pour déduire des principes d'invariance dans $L^2[0,1]$, ce qui est très habituel dans D[0,1], cf. Herrndorf [19] ou Peligrad [39], par exemple. Evidemment nous pouvons continuer à profiter des hypothèses de mélangeance (éventuellement autres types de mélangeance) pour déduire des principes d'invariance, mais la procédure est toujours pareille à celle que nous avons utilisé ici, donc nous n'énoncerons pas ces résultats.

Chapitre 5

Estimation d'un noyau de transition

Nous allons maintenant entreprendre une étude un peu différente de ce que nous avons fait jusqu'à présent. En effet, si dans les chapitres précédents nous nous sommes intéressé aux aspects théoriques des mesures aléatoires, dans ce chapitre nous envisagerons des applications à la statistique des mesures aléatoires, en utilisant seulement les mesures aléatoires non négatives. Ce qui nous intéressera au cours de ce chapitre sera l'étude d'un estimateur d'un noyau de transition pour des mesures aléatoires non négatives. Notre approche du problème nous donnera un regard général sur quelques problèmes d'estimation qui semblent être différents: l'estimation de la densité, de la régression, de la distribution locale d'un processus chromatique [32], d'une mesure aléatoire composite [43] ou des lois de Palm d'une mesure aléatoire [34]. La parenté de tous ces problèmes, qui paraît liée aux techniques utilisées pour les résoudre, par exemple régressogramme ou noyau, provient plus profondément du fait qu'il s'agit, dans la plupart des cas, d'estimer des densités de Radon-Nikodym. Cette approche générale des problèmes référés clarifie quelques conditions de convergence des estimateurs utilisés classiquement. En outre, elle nous permet d'envisager l'étude de problèmes de convergence des processus empiriques associés à l'estimateur de façon plus commode.

Cette méthode à été utilisé par Jacob, Mendes-Lopes [23] pour l'étude de l'estimateur d'un noyau de transition particulier où on s'intéressait au cas de mesures aléatoires à densité. Ce cas se ramenait alors à l'étude de densités aléatoires. Dans cet exposé nous nous intéresserons à l'étude de mesures aléatoires non négatives générales, donc, qui n'ont pas nécessairement une densité (aléatoire) par rapport à une mesure fixée (habituellement la mesure de Lebesgue sur $\mathbb{R}^p$). Comme on le remarquera plus tard, le fait de considérer les mesures aléatoires non négatives générales change la nature du problème et des résultats.

Nous commencerons ce chapitre en référant quelques résultats généraux pour variables aléatoires réelles positives, qui nous permettrons d'établir les convergences des estimateurs que l'on envisage. Dans un premier temps nous nous intéresserons à l'estimateur du type régressogramme, pour lequel nous établirons des résultats de convergence presque complète uniforme. De plus, nous établirons des conditions pour la convergence des lois de dimension finie du processus empirique associé à l'estimateur. Ensuite nous nous intéresserons à l'estimateur du type noyau, pour lequel nous établirons des résultats du même type.

5.1 Résultats préliminaires

Nous allons reprendre quelques résultats généraux pour variables aléatoires réelles positives qui nous aideront plus tard pour l'étude de la convergence des estimateurs.

Théorème 5.1.1 *Soient X et Y des variables aléatoires réelles positives et $\varepsilon \in (0,1)$. Alors*

$$\{|X-1| > \varepsilon Y\} \subset \left\{|X-1| > \frac{\varepsilon}{2}\right\} \cup \left\{|Y-1| > \frac{\varepsilon}{2}\right\}.$$

<u>**Démonstration**</u> : Soit $\delta > 0$. Alors

$$\{|X-1| > \varepsilon Y\} = \{|X-1| > \varepsilon Y, |Y-1| < \delta\} \cup \{|X-1| > \varepsilon Y, |Y-1| \geq \delta\}.$$

Si $|Y-1| < \delta$ alors $Y > 1-\delta$, et donc

$$\{|X-1| > \varepsilon Y\} \subset \{|X-1| > \varepsilon(1-\delta)\} \cup \{|Y-1| \geq \delta\}.$$

Si nous posons $\varepsilon(1-\delta) = \delta$, il s'ensuit $\delta = \frac{\varepsilon}{\varepsilon+1} > \frac{\varepsilon}{2}$, car $\varepsilon \in (0,1)$, donc

$$\{|X-Y| > \varepsilon Y\} \subset \left\{|X-1| > \frac{\varepsilon}{2}\right\} \cup \left\{|Y-1| > \frac{\varepsilon}{2}\right\}. \blacksquare$$

Théorème 5.1.2 *Soient X et Y des variables aléatoires réelles positives et intégrables. Pour $\varepsilon > 0$ suffisamment petit il se vérifie*

$$\left\{\left|\frac{X}{Y} - \frac{\mathbf{E}(X)}{\mathbf{E}(Y)}\right| > \varepsilon\right\} \subset \left\{\left|\frac{X}{\mathbf{E}(X)} - 1\right| > \frac{\varepsilon\,\mathbf{E}(Y)}{4\,\mathbf{E}(X)}\right\} \cup \left\{\left|\frac{Y}{\mathbf{E}(Y)} - 1\right| > \frac{\varepsilon\,\mathbf{E}(Y)}{4\,\mathbf{E}(X)}\right\}.$$

<u>**Démonstration**</u> : Posons $\alpha = \frac{\mathbf{E}(Y)}{\mathbf{E}(X)}$, alors

$$\left\{\left|\frac{X}{Y} - \frac{\mathbf{E}(X)}{\mathbf{E}(Y)}\right| > \varepsilon\right\} =$$

$$\left\{\left|\frac{X}{Y}\frac{\mathbf{E}(Y)}{\mathbf{E}(X)} - 1\right| > \varepsilon\alpha\right\} = \left\{\left|\frac{X}{\mathbf{E}(X)} - \frac{Y}{\mathbf{E}(Y)}\right| > \varepsilon\alpha\frac{Y}{\mathbf{E}(Y)}\right\} \subset$$

$$\subset \left\{\left|\frac{X}{\mathbf{E}(X)} - 1\right| + \left|\frac{Y}{\mathbf{E}(Y)} - 1\right| > \varepsilon\alpha\frac{Y}{\mathbf{E}(Y)}\right\} \subset$$

$$\subset \left\{\left|\frac{X}{\mathbf{E}(X)} - 1\right| > \frac{\varepsilon\alpha}{2}\frac{Y}{\mathbf{E}(Y)}\right\} \cup \left\{\left|\frac{Y}{\mathbf{E}(Y)} - 1\right| > \frac{\varepsilon\alpha}{2}\frac{Y}{\mathbf{E}(Y)}\right\} \subset$$

$$\subset \left\{\left|\frac{X}{\mathbf{E}(X)} - 1\right| > \frac{\varepsilon\alpha}{4}\right\} \cup \left\{\left|\frac{Y}{\mathbf{E}(Y)} - 1\right| > \frac{\varepsilon\alpha}{4}\right\},$$

d'après le théorème précédent. ■

Nous démontrons ensuite une inégalité du type Bernstein-Fréchet, en posant des hypothèses sur les moments non centrés.

Théorème 5.1.3 *Soit X une variable aléatoire réelle ayant des moments de tout ordre. De plus, supposons que*

$$\exists_{M>0}\,\forall_{k>2}\;\mathbf{E}(|X|^k) \leq M^{k-2}k!\,\mathbf{E}(X^2).$$

Alors

$$\mathbf{P}\left(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbf{E}(X)| \geq u\right) \leq 2\exp\left(-\frac{nu^2}{4\mathbf{E}(X^2) + 2Mu}\right), \tag{5.1}$$

où $X_1, \ldots, X_n$ sont des copies indépendantes de X et $u > 0$.

<u>**Démonstration**</u> : Pour $|t| < M^{-1}$ nous pouvons écrire

$$\begin{aligned} e^{-t\mathbf{E}(X)}\mathbf{E}(e^{tX}) &= e^{-t\mathbf{E}(X)}\left(1 + t\mathbf{E}(X) + \textstyle\sum_{k=2}^{\infty} \frac{t^k}{k!}\mathbf{E}(X^k)\right) \leq \\ &\leq e^{-t\mathbf{E}(X)}\left(1 + t\mathbf{E}(X) + \textstyle\sum_{k=2}^{\infty} M^{k-2}|t|^k\mathbf{E}(X^2)\right) \leq \\ &\leq e^{-t\mathbf{E}(X)}e^{t\mathbf{E}(X)+\frac{t^2\mathbf{E}(X^2)}{1-|t|M}} = \\ &= e^{\frac{t^2\mathbf{E}(X^2)}{1-|t|M}}. \end{aligned}$$

Comme les copies $X_1, \ldots, X_n$ de X sont indépendantes il s'ensuit, pour tout $t > 0$,

$$\mathbf{E}\left[e^{t(\sum_i X_i - n\mathbf{E}(X))}\right] = \left[\mathbf{E}\left(e^{t(X_i-\mathbf{E}(X))}\right)\right]^n \leq e^{\frac{nt^2\mathbf{E}(X^2)}{1-|t|M}}.$$

Considérons $\varepsilon > 0$ et $t \in (0, M^{-1})$. Alors, l'inégalité de Chebyshev nous permet de déduire

$$\begin{aligned} \mathbf{P}\left(\sum_{i=1}^{n} X_i - n\mathbf{E}(X) \geq \varepsilon\sqrt{n\mathbf{E}(X^2)}\right) &\leq \frac{\mathbf{E}\left[e^{t(\sum_i X_i - n\mathbf{E}(X))}\right]}{e^{\varepsilon t\sqrt{n\mathbf{E}(X^2)}}} \leq \\ &\leq \exp\left(\frac{nt^2\mathbf{E}(X^2)}{1-|t|M} - \varepsilon t\sqrt{n\mathbf{E}(X^2)}\right). \end{aligned} \tag{5.2}$$

Posons $t = \frac{\varepsilon}{2\sqrt{n\mathbf{E}(X^2)}+\varepsilon M}$, alors $\varepsilon = \frac{2t\sqrt{n\mathbf{E}(X^2)}}{1-|t|M}$. Du fait que $\varepsilon > 0$ il s'ensuit $t < M^{-1}$, donc les inégalités déduites sont vraies pour ce choix de t. L'inégalité (5.2) peut s'écrire, pour ce choix de t,

$$\begin{aligned} \mathbf{P}\left(\sum_{i=1}^{n} X_i - n\mathbf{E}(X) \geq \varepsilon\sqrt{n\mathbf{E}(X^2)}\right) &\leq \exp\left(-\tfrac{\varepsilon}{2}t\sqrt{n\mathbf{E}(X^2)}\right) = \\ &= \exp\left(-\frac{\varepsilon^2\sqrt{n\mathbf{E}(X^2)}}{4\sqrt{n\mathbf{E}(X^2)}+2\varepsilon M}\right). \end{aligned}$$

En posant $u = \varepsilon\sqrt{n\mathbf{E}(X^2)}$, cette inégalité s'écrit

$$\mathbf{P}\left(\sum_{i=1}^{n} X_i - n\mathbf{E}(X) \geq u\right) \leq \exp\left(-\frac{u^2}{4n\mathbf{E}(X^2) + 2Mu}\right).$$

Maintenant, en écrivant cette expression avec nu au lieu de u et en divisant par n l'expression qui définit l'ensemble dont nous voulons calculer la probabilité nous obtenons l'inégalité

$$\mathbf{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbf{E}(X) \geq u\right) \leq \exp\left(-\frac{nu^2}{4\mathbf{E}(X^2)+2Mu}\right).$$

En raisonnant de la même façon pour $\mathbf{E}(X) - \frac{1}{n}\sum_i X_i$ nous obtenons une inégalité équivalente, d'où il s'ensuit (5.1), ce qui démontre le théorème. ■

5.2 L'approche générale de l'estimation d'un noyau de transition

Nous allons vérifier que nous pouvons réduire quelques problèmes d'estimation au même cadre général que nous étudierons ensuite.

Considérons une mesure aléatoire non négative ζ.

- Supposons $\mathbf{E}(\zeta(\mathbf{S})) < \infty$. Les probabilités de Palm de ζ forment un noyau de transition $\{Q_s(\Gamma) : s \in \mathbf{S}, \Gamma$ borélien vague de $\mathcal{M}_c\}$ tel que

  $$\mathbf{E}(\zeta(\mathrm{B})\mathbb{1}_{\{\zeta\in\Gamma\}}) = \int_{\mathrm{B}} Q_s(\Gamma)\,\mathbf{E}(\zeta)(ds), \quad \mathrm{B} \in \mathcal{B}.$$

  L'estimation de $Q_s(\Gamma)$, $s \in \mathrm{D} \in \mathcal{B}$ a été étudiée dans [34]. Posons $\xi = \zeta$, $\eta(\mathrm{B}) = \zeta(\mathrm{B})\mathbb{1}_{\{\zeta\in\Gamma\}}$ et $\mu = \mathbf{E}(\eta)$, $\nu = \mathbf{E}(\xi)$. Alors on vérifie facilement que $Q_s(\Gamma) = \frac{d\mu}{d\nu}(s)$ ν-presque partout.

- Supposons que $\mathbf{S}=\mathbf{S}'\times\mathbb{R}$, où $\mathbf{S}'$ est un espace métrique séparable, complet et localement compact, et que $\mathbf{E}(\zeta(\mathbf{S})) < \infty$. La distribution locale moyenne de la mesure aléatoire composite ζ est le noyau de transition $\{q^s(\mathrm{B}) : s \in \mathbf{S}', \mathrm{B} \in \mathcal{B}(\mathbb{R})$ borné$\}$ tel que

  $$\mathbf{E}(\zeta(\mathrm{A}\times\mathrm{B})) = \int_{\mathrm{A}} q^s(\mathrm{B})\,\mathbf{E}(\zeta)(ds \times \mathbb{R}), \quad \mathrm{A} \in \mathcal{B}(\mathbf{S}').$$

  L'estimation de $q^s(\mathrm{B})$, $s \in \mathrm{D} \in \mathcal{B}(\mathbf{S}')$ est faite dans [32] et [43]. Posons $\xi(\mathrm{A}) = \zeta(\mathrm{A}\times\mathbb{R})$, $\eta(\mathrm{A}) = \zeta(\mathrm{A}\times\mathrm{B})$ et $\mu = \mathbf{E}(\eta)$, $\nu = \mathbf{E}(\xi)$. Il est facile de vérifier que $q^s(\mathrm{B}) = \frac{d\mu}{d\nu}(s)$ ν-presque partout.

- Soit Y une variable aléatoire positive intégrable et X une variable aléatoire à valeurs dans $\mathbf{S}$. Nous pouvons nous intéresser à l'étude de $\mathbf{E}(Y|X = s)$, $s \in \mathrm{D} \in \mathcal{B}$. Posons $\xi(\mathrm{B}) = \mathbb{1}_{\{X \in \mathrm{B}\}}$, $\eta(\mathrm{B}) = Y\mathbb{1}_{\{X\in \mathrm{B}\}}$ et $\mu = \mathbf{E}(\eta)$, $\nu = \mathbf{E}(\xi)$. Alors
$$\mathbf{E}\left(\mathbb{1}_{\{X\in \mathrm{B}\}}\frac{d\mu}{d\nu}(X)\right) = \int_{\mathrm{B}} \frac{d\mu}{d\nu}(x)\,\nu(dx) = \int_{\mathrm{B}} \mu(dx) = \mu(\mathrm{B}) = \mathbf{E}(Y\mathbb{1}_{\{X\in\mathrm{B}\}}),$$
donc $\frac{d\mu}{d\nu}(s) = \mathbf{E}(Y|X = s)$ ν-presque partout.

Nous avons donc réduit ces trois problèmes d'estimation au problème plus général de l'estimation de $\frac{d\mu}{d\nu}$, où μ et ν sont les mesures moyennes de deux mesures aléatoires non négatives η et ξ, respectivement. Nous pouvons donc faire une étude générale de ces problèmes d'estimation en utilisant des estimateurs du type régressogramme ou, dans le cas $\mathbf{S}=\mathbb{R}^p$, des estimateurs du type noyau.

5.3 Estimateurs du type régressogramme

Dans l'espace $\mathbf{S}$ considérons une suite de partitions $\{\Pi_k\}_{k\in\mathbf{N}}$ telle que

- $\Pi_k \subset \mathcal{B}$ pour tout $k \in \mathbb{N}$;
- Π_{k+1} est un raffinement de Π_k;
- pour tout k fixé, chaque élément de $\mathcal{B}$ ne rencontre qu'un nombre fini d'éléments de Π_k;
- $\lim_{k\to\infty} \delta_k(\mathrm{B}) = 0$, pour tout $\mathrm{B}\in \mathcal{B}$, où
$$\delta_k(\mathrm{B}) = \sup\,\{diam(I), I \in \Pi_k, I \cap \mathrm{B} \neq \emptyset\}.$$

Si nous fixons un point s dans $\mathbf{S}$, pour chaque $k \in\mathbf{N}$, il existe un unique élément de la partition Π_k qui contient s. Désignons cet élément par $I_k(s)$. Evidemment $\lim_{k\to\infty} diam(I_k(s)) = 0$.

Définissons les fonctions g_k par
$$g_k(s) = \begin{cases} \frac{\mu(I_k(s))}{\nu(I_k(s))} & \text{si } \nu(I_k(s)) > 0 \\ 0 & \text{si } \nu(I_k(s)) = 0. \end{cases}$$

Alors, d'après un résultat classique de la théorie des martingales, nous savons que la suite $\{g_k\}$ converge vers $\frac{d\mu}{d\nu}$ ν-presque partout. De plus, si nous choisissons une version continue de $\frac{d\mu}{d\nu}$, f, sur un borélien borné B, la convergence de la suite $\{g_k\}$ vers f est uniforme sur B. Evidemment, du point de vue pratique, les fonctions g_k ne servent pas comme estimateurs, car en général nous ne connaissons pas les mesures μ et ν. Il nous intéresse d'étudier un estimateur construit à partir d'un échantillon du couple (η, ξ).

Soit $((\eta_1, \xi_1), \ldots, (\eta_n, \xi_n))$ un échantillon du couple (η, ξ). Posons $\overline{\eta}_n = \frac{1}{n}\sum_{i=1}^n \eta_i$, $\overline{\xi}_n = \frac{1}{n}\sum_{i=1}^n \xi_i$ et définissons les estimateurs

$$f_n(s) = \frac{\overline{\eta}_n(I_k(s))}{\overline{\xi}_n(I_k(s))},$$

où les indices k dépendent de n de façon à expliciter plus tard pour obtenir les convergences voulues.

On remarque tout de suite que les estimateurs f_n correspondent aux estimateurs des lois de Palm d'une mesure aléatoire étudiés par Niéré [34], d'une mesure aléatoire composite étudié par Saleh [43], de la distribution locale d'un processus chromatique étudié par Mendes-Lopes [32], et du régressogramme, selon le cas particulier choisi, d'après la section précédente.

Nous voulons établir la convergence de la suite $\{f_n\}$ vers le noyau de transition $f = \frac{d\mu}{d\nu}$. Pour cela nous allons étudier la convergence de

$$|f_n(s) - g_k(s)|$$

vers zéro, en choisissant de façon correcte les indices k. Nous allons essayer de démontrer la convergence presque complète de f_n vers le noyau de transition.

Remarquons que, d'après le théorème 5.1.2, pour $\varepsilon > 0$ suffisamment petit,

$$\{|f_n(s) - g_k(s)| > \varepsilon\} = \left\{\left|\frac{\overline{\eta}_n(I_k(s))}{\overline{\xi}_n(I_k(s))} - \frac{\mu(I_k(s))}{\nu(I_k(s))}\right| > \varepsilon\right\} \subset$$

$$\subset \left\{\left|\frac{\overline{\eta}_n(I_k(s))}{\mu(I_k(s))} - 1\right| > \frac{\varepsilon}{4}\frac{\nu(I_k(s))}{\mu(I_k(s))}\right\} \cup \left\{\left|\frac{\overline{\xi}_n(I_k(s))}{\nu(I_k(s))} - 1\right| > \frac{\varepsilon}{4}\frac{\nu(I_k(s))}{\mu(I_k(s))}\right\} =$$

$$= \left\{|\overline{\eta}_n(I_k(s)) - \mu(I_k(s))| > \frac{\varepsilon}{4}\nu(I_k(s))\right\} \cup \left\{|\overline{\xi}_n(I_k(s)) - \nu(I_k(s))| > \frac{\varepsilon}{4}\frac{(\nu(I_k(s)))^2}{\mu(I_k(s))}\right\},$$

donc, en applicant le théorème 5.1.3, nous obtenons l'inégalité

$$\mathbf{P}\{|f_n(s) - g_k(s)| > \varepsilon\} \leq$$

$$\leq 2\exp\left(-\frac{\frac{\varepsilon}{4}n\nu(I_k(s))}{\frac{4\mathbf{E}(\eta^2(I_k(s)))}{\frac{\varepsilon}{4}\nu(I_k(s))}+2M}\right) + 2\exp\left(-\frac{\frac{\varepsilon}{4}n\frac{(\nu(I_k(s)))^2}{\mu(I_k(s))}}{\frac{4\mathbf{E}(\xi^2(I_k(s)))}{\frac{\varepsilon}{4}\frac{(\nu(I_k(s)))^2}{\mu(I_k(s))}}+2M}\right), \tag{5.3}$$

dés que nous supposons qu'il existe une constante $M > 0$ telle que pour tout $k > 2$, se vérifie

$$\begin{aligned} \mathbf{E}(\eta^k(I)) &\leq M^{k-2}k!\,\mathbf{E}(\eta^2(I)), \quad I \in \mathcal{B} \\ \mathbf{E}(\xi^k(I)) &\leq M^{k-2}k!\,\mathbf{E}(\xi^2(I)), \quad I \in \mathcal{B}. \end{aligned} \tag{5.4}$$

Pour pouvoir démontrer la convergence presque complète de l'estimateur nous avons besoin d'améliorer l'inégalité (5.3). Pour cela il nous faudra imposer quelques conditions supplémentaires.

Démontrons d'abord quelques propriétés qui nous permettrons de choisir des conditions d'une façon naturelle.

Théorème 5.3.1 *Soit $B \in \mathcal{B}$ fixé.*

1. *Les fonctions*

$$\sum_{I\in\Pi_k} \frac{\mathbf{E}(\eta^2(I\cap B))}{\mathbf{E}(\eta(I\cap B))}\mathbb{1}_{I\cap B}$$

$$\sum_{I\in\Pi_k} \frac{\mathbf{E}(\xi^2(I\cap B))}{\mathbf{E}(\xi(I\cap B))}\mathbb{1}_{I\cap B}$$

$$\sum_{I\in\Pi_k} \frac{\mathbf{E}\left[\xi(I\cap B)\eta(I\cap B)\right]}{\mathbf{E}(\eta(I\cap B))}\mathbb{1}_{I\cap B}$$

sont des sur-martingales positives par rapport à la filtration $\mathcal{F}_k = \sigma(\Pi_k)$.

2. *La fonction*

$$\sum_{I\in\Pi_k} \frac{\mathbf{E}(\eta(I\cap B))}{\mathbf{E}(\xi(I\cap B))}\mathbb{1}_{I\cap B}$$

est une martingale positive par rapport à la filtration $\mathcal{F}_k = \sigma(\Pi_k)$.

<u>**Démonstration**</u> : 1. Posons

$$X_k = \sum_{I\in\Pi_k} \frac{\mathbf{E}(\eta^2(I\cap \mathrm{B}))}{\mathbf{E}(\eta(I\cap \mathrm{B}))} 1\!\!1_{I\cap \mathrm{B}}$$

et démontrons que, pour tout A$\in \mathcal{F}_k$, il se vérifie

$$\int_{\mathrm{A}} X_{k+1}\, d\mathbf{E}(\eta) \leq \int_{\mathrm{A}} X_k\, d\mathbf{E}(\eta).$$

Comme Π_{k+1} est un raffinement de la partition Π_k, étant donné A$\in \mathcal{F}_k$, il existe $\mathrm{A}_1,\ldots,\mathrm{A}_m \in \mathcal{F}_{k+1}$ deux à deux disjoints et tels que $\mathrm{A}=\bigcup_{j=1}^m \mathrm{A}_j$. Alors, il s'ensuit

$$\int_{\mathrm{A}} X_{k+1}\, d\mathbf{E}(\eta) = \sum_{j=1}^{m} \int_{\mathrm{A}_j} X_{k+1}\, d\mathbf{E}(\eta) =$$

$$= \sum_{j=1}^{m} \int_{\mathrm{A}_j} \sum_{I\in\Pi_{k+1}} \frac{\mathbf{E}(\eta^2(I\cap \mathrm{B}))}{\mathbf{E}(\eta(I\cap \mathrm{B}))} 1\!\!1_{I\cap \mathrm{B}}\, d\mathbf{E}(\eta) = \sum_{j=1}^{m} \frac{\mathbf{E}(\eta^2(\mathrm{A}_j\cap \mathrm{B}))}{\mathbf{E}(\eta(\mathrm{A}_j\cap \mathrm{B}))} \mathbf{E}(\eta(\mathrm{A}_j\cap\mathrm{B})) =$$

$$= \sum_{j=1}^{m} \mathbf{E}(\eta^2(\mathrm{A}_j\cap \mathrm{B})) \leq \mathbf{E}\left(\sum_{j=1}^{m} \eta(\mathrm{A}_j\cap \mathrm{B})\right)^2 =$$

$$= \mathbf{E}(\eta^2(A\cap \mathrm{B})) = \int_{\mathrm{A}} X_k\, d\mathbf{E}(\eta).$$

Pour démontrer que $\sum_{I\in\Pi_k} \frac{\mathbf{E}(\xi^2(I\cap\mathrm{B}))}{\mathbf{E}(\xi(I\cap\mathrm{B}))} 1\!\!1_{I\cap\mathrm{B}}$ et $\sum_{I\in\Pi_k} \frac{\mathbf{E}[\xi(I\cap\mathrm{B})\eta(I\cap\mathrm{B})]}{\mathbf{E}(\eta(I\cap\mathrm{B}))} 1\!\!1_{I\cap\mathrm{B}}$ sont des sur-martingales on procède de la même façon en intégrant par rapport aux mesures $\mathbf{E}(\xi)$ et $\mathbf{E}(\eta)$, respectivement.

2. On démontre que $\sum_{I\in\Pi_k} \frac{\mathbf{E}(\eta(I\cap\mathrm{B}))}{\mathbf{E}(\xi(I\cap\mathrm{B}))} 1\!\!1_{I\cap\mathrm{B}}$ est une martingale en procédant de la même façon, en intégrant par rapport à la mesure $\mathbf{E}(\xi)$, et en remarquant que la seule inégalité dans les calculs devient une égalité, car $\mathbf{E}(\eta)$ est une mesure. Ce résultat est d'ailleurs classique. ■

Nous allons alors supposer des conditions liées à la convergence de ces sur-martingales. Considérons les conditions suivantes:

$$\begin{aligned}
&\forall_{\mathrm{B}\in\mathcal{B}}\ \exists_{C_1>0,C_1'>0}\ \forall_{k\in\mathbf{N},\,I\in\Pi_k}\ C_1' \leq \frac{\mathbf{E}(\eta(I\cap\mathrm{B}))}{\mathbf{E}(\xi(I\cap\mathrm{B}))} \leq C_1;\\
&\forall_{\mathrm{B}\in\mathcal{B}}\ \exists_{C_2>0,C_2'>0}\ \forall_{k\in\mathbf{N},\,I\in\Pi_k}\ C_2' \leq \frac{\mathbf{E}(\xi^2(I\cap\mathrm{B}))}{\mathbf{E}(\xi(I\cap\mathrm{B}))} \leq C_2;\\
&\forall_{\mathrm{B}\in\mathcal{B}}\ \exists_{C_3>0,C_3'>0}\ \forall_{k\in\mathbf{N},\,I\in\Pi_k}\ C_3' \leq \frac{\mathbf{E}(\eta^2(I\cap\mathrm{B}))}{\mathbf{E}(\eta(I\cap\mathrm{B}))} \leq C_3;\\
&\forall_{\mathrm{B}\in\mathcal{B}}\ \exists_{C_4>0,C_4'>0}\ \forall_{k\in\mathbf{N},\,I\in\Pi_k}\ C_4' \leq \frac{\mathbf{E}[\xi(I\cap\mathrm{B})\eta(I\cap\mathrm{B})]}{\mathbf{E}(\eta(I\cap\mathrm{B}))} \leq C_4;
\end{aligned} \tag{5.5}$$

Remarquons que ces conditions sont vérifiées dans le cas où les sur-martingales et la martingale définies dans le théorème 5.3.1 sont convergentes (ce qui est vérifié dans le cas présent, car les sur-martingales et la martingale sont non négatives), la convergence est uniforme sur le borélien relativement compact B et la limite est strictement positive sur B. C'est ce qui se passe, par exemple, dans le cas où on suppose l'existence de la densité des mesures ν et μ par rapport à une mesure sur **S**, non nulle sur B, et que ces densités sont strictement positives sur B.

Nous pouvons maintenant démontrer une inégalité qui nous permettra d'établir la convergence presque complète ponctuelle de l'estimateur f_n.

Théorème 5.3.2 *Soient ξ et η deux mesures aléatoires non négatives intégrables telles que $\mathbf{E}(\eta) \ll \mathbf{E}(\xi)$. Supposons vérifiées les conditions (5.4) et (5.5). Alors, pour $\varepsilon > 0$ suffisamment petit,*

$$\mathbf{P}\{|f_n(s) - g_k(s)| > \varepsilon\} \leq 4\exp\left(-Cn\nu(I_k(s))\right), \tag{5.6}$$

où $C > 0$ est une constante.

<u>**Démonstration**</u> : En utilisant les conditions (5.4) nous avions déjà établi l'inégalité (5.3), que nous pouvons écrire sous la forme

$$\mathbf{P}\{|f_n(s) - g_k(s)| > \varepsilon\} \leq$$

$$2\exp\left(-\frac{\frac{\varepsilon}{4}n\nu(I_k(s))}{\frac{16\mathbf{E}(\eta^2(I_k(s)))}{\varepsilon\nu(I_k(s))}+2M}\right) + 2\exp\left(-\frac{\frac{\varepsilon}{4}n\frac{(\nu(I_k(s)))^2}{\mu(I_k(s))}}{\frac{16\mathbf{E}(\xi^2(I_k(s)))}{\varepsilon\frac{(\nu(I_k(s)))^2}{\mu(I_k(s))}}+2M}\right). \tag{5.7}$$

Des conditions (5.5) nous déduisons les majorations, pour tous $k \in \mathbb{N}$, $I \in \Pi_k$,

$$\frac{\nu^2(I \cap I_1(s))}{\mu(I \cap I_1(s))} \geq \frac{\nu(I \cap I_1(s))}{C_1};$$

$$\frac{\mathbf{E}(\eta^2(I \cap I_1(s)))}{\nu(I \cap I_1(s))} \leq C_1\frac{\mathbf{E}(\eta^2(I \cap I_1(s)))}{\mu(I \cap I_1(s))};$$

$$\frac{\mu(I \cap I_1(s))\mathbf{E}(\xi^2(I \cap I_1(s)))}{\nu^2(I \cap I_1(s))} \leq C_1\frac{\mathbf{E}(\xi^2(I \cap I_1(s)))}{\nu(I \cap I_1(s))}.$$

Donc, en substituant dans l'inégalité (5.7), nous obtenons, en rappelant que $I_k(s) \subset I_1(s), k \geq 2$

$$\mathbf{P}\{|f_n(s) - g_k(s)| > \varepsilon\} \leq$$

$$\leq 2\exp\left(-\frac{\frac{\varepsilon}{4} n\nu(I_k(s))}{\frac{16C_1}{\varepsilon}\frac{\mathbf{E}(\eta^2(I_k(s)))}{\mu(I_k(s))}+2M}\right) + 2\exp\left(-\frac{\frac{\varepsilon}{4C_1} n\nu(I_k(s))}{\frac{16C_1}{\varepsilon}\frac{\mathbf{E}(\xi^2(I_k(s)))}{\nu(I_k(s))}+2M}\right) \leq$$

$$\leq 2\exp\left(-\frac{\frac{\varepsilon}{4} n\nu(I_k(s))}{\frac{16C_1C_3}{\varepsilon}+2M}\right) + 2\exp\left(-\frac{\frac{\varepsilon}{4C_1} n\nu(I_k(s))}{\frac{16C_1C_2}{\varepsilon}+2M}\right) \leq$$

$$\leq 4\exp\left(-Cn\nu(I_k(s))\right),$$

en posant $C = \min\left(\frac{\frac{\varepsilon}{4}}{\frac{16C_1C_3}{\varepsilon}+2M}, \frac{\frac{\varepsilon}{4C_1}}{\frac{16C_1C_2}{\varepsilon}+2M}\right)$, ce qui démontre l'inégalité (5.6). ■

La convergence presque complète ponctuelle de l'estimateur f_n s'ensuit comme conséquence de ce théorème.

Corollaire 5.3.3 *Dans les conditions du théorème 5.3.2 nous avons*

$$f_n(s) \longrightarrow \frac{d\mu}{d\nu}(s) \ \ p.co..$$

<u>**Démonstration**</u> : En effet de l'inégalité (5.6) et du choix de la suite de partitions $\{\Pi_k\}$, il s'ensuit

$$\sum_n \mathbf{P}\{|f_n(s) - g_k(s)| > \varepsilon\} \leq \sum_n 4\exp\left(-Cn\nu(I_k(s))\right) \leq$$

$$\leq \sum_n 4\exp\left(-Cn\nu(I_1(s))\right) \leq \infty,$$

donc $|f_n(s) - g_k(s)| \longrightarrow 0$ p.co.. Comme nous avons aussi la convergence $g_k(s) \longrightarrow \frac{d\mu}{d\nu}(s)$ il s'ensuit la convergence énoncée. ■

Remarquons que si nous supposons que $s \in$B, où B est un borélien relativement compact fixé, la constante C intervenant dans les deux résultats précédents est indépendante du point s auxquel nous nous intéressons.

Maintenant, en choisissant convenablement la suite de partitions nous pouvons déduire la convergence presque complète uniforme sur un borélien borné B que nous considérerons fixé dans la suite. Définissons

$$R_k = \{I \in \Pi_k : \ I \cap \mathrm{B} \neq \emptyset\}.$$

D'après les hypothèses sur la suite de partitions $\{\Pi_k\}$, l'ensemble R_k n'a qu'un nombre fini d'éléments. Alors

$$\sup_{s\in\mathrm{B}} |f_n(s) - g_k(s)| = \max_{I\in R_k} \sup_{s\in I} |f_n(s) - g_k(s)|,$$

donc

$$\mathbf{P}\left\{\sup_{s\in\mathrm{B}} |f_n(s) - g_k(s)| > \varepsilon\right\} = \mathbf{P}\left\{\bigcup_{I\in R_k} |\frac{\overline{\eta}_n(I)}{\overline{\xi}_n(I)} - \frac{\mu(I)}{\nu(I)}| > \varepsilon\right\} \leq$$

$$\leq 4\#R_k \exp\left(-Cn \min_{I\in\Pi_k} \{\nu(I)\}\right),$$

où $\#R_k$ représente le nombre d'éléments de l'ensemble R_k. Fixons $\delta > 0$ tel que le dilaté $\mathrm{B}^\delta = \{s \in \mathbf{S} : d(s, \mathrm{B}) < \delta\}$ soit encore relativement compact. Pour $k \in \mathbb{N}$ assez grand nous avons

$$\#R_k \min_{I\in\Pi_k} \{\nu(I)\} \leq \mathbf{E}\left(\xi(\mathrm{B}^\delta)\right) < \infty.$$

Si nous choisissons la suite de partitions $\{\Pi_k\}$ de telle façon que

$$\min_{I\in\Pi_k} \{\nu(I)\} = \frac{\log n}{n\varepsilon_n}, \tag{5.8}$$

où $\{\varepsilon_n\}$ est une suite de réels positifs convergeant vers zéro, il s'ensuit que, pour k assez grand se vérifie

$$\#R_k \leq \nu(\mathrm{B}^\delta)\frac{n\varepsilon_n}{\log n} \leq n,$$

donc

$$\mathbf{P}\{\sup_{s\in\mathrm{B}} |f_n(s) - g_k(s)| > \varepsilon\} \leq 4n \exp\left(-Cn \min_{I\in\Pi_k} \{\nu(I)\}\right). \tag{5.9}$$

Théorème 5.3.4 *Supposons vérifiées les conditions du théorème 5.3.2 et la condition (5.8). Alors, la convergence $f_n(s) \longrightarrow \frac{d\mu}{d\nu}(s)$ est presque complète uniforme, dès que nous choisissons une version uniformément continue sur B de la densité de Radon-Nikodym $\frac{d\mu}{d\nu}(s)$.*

<u>**Démonstration**</u> : En effet, si (5.8) se vérifie, de l'inégalité (5.9) il s'ensuit

$$\mathbf{P}\{\sup_{s\in B} |f_n(s) - g_k(s)| > \varepsilon\} \leq 4n \exp\left(-C\frac{\log n}{\varepsilon_n}\right) = 4n^{1-\frac{C}{\varepsilon_n}},$$

donc

$$\sum_{n=1}^{\infty} \mathbf{P}\{\sup_{s\in B} |f_n(s) - g_k(s)| > \varepsilon\} \leq 4\sum_{n=1}^{\infty} n^{1-\frac{C}{\varepsilon_n}} < \infty.$$

Comme la convergence de g_k vers une version uniformément continue sur B de $\frac{d\mu}{d\nu}$ est uniforme sur B, il s'ensuit que la convergence de f_n vers une version uniformément continue sur B de $\frac{d\mu}{d\nu}$ est presque complète uniforme sur B. ■

5.4 Considérations du point de vue pratique

Dans la pratique se pose le problème du choix des partitions Π_k. En effet, la condition (5.8) n'est pas vérifiable car en général on ne connait pas la mesure ν. Donc les résultats théoriques démontrés ne sont pas utilisables, au moins directement. Pourtant, si nous supposons quelques propriétés sur la mesure moyenne ν nous pouvons quand même nous servir des résultats théoriques. Supposons alors que la mesure ν est absolument continue par rapport à une mesure fixée λ sur l'espace **S**. Dans ce cas nous pouvons nous servir des résultats théoriques pour bien choisir la suite de partitions. En fait, si nous choisissons la partition Π_k de telle façon que

$$\min_{I\in\Pi_k} \lambda(I) = \frac{\log n}{n\varepsilon_n}, \tag{5.10}$$

nous pouvons déduire le même résultat de convergence presque complète uniforme, car dans ce cas

$$\min_{I\in\Pi_k} \nu(I) \leq \sup_{s\in B} |\frac{d\mu}{d\nu}(s)| \min_{I\in\Pi_k} \lambda(I),$$

donc, ce qui change dans la condition (5.8) c'est le fait qu'il est nécessaire de multiplier par une constante, ce qui ne change rien en ce qui concerne la convergence de la série.

Remarquons que le fait de supposer que la mesure moyenne ν est absolument continue par rapport à une mesure fixée sur **S** ne nous ramène pas au cas étudié par Jacob, Mendes-Lopes [23], car ce fait n'implique pas que la mesure aléatoire ξ soit à densité. Un exemple évident est un processus ponctuel de Poisson sur la droite, qui n'est évidemment pas à densité mais sa mesure moyenne est à densité par rapport à la mesure de Lebesgue.

Dans le cas particulier où $S=\mathbb{R}^p$ et λ la mesure de Lebesgue sur $\mathbb{R}^p$ la condition (5.10) est équivalente à

$$\frac{n \min_{I\in\Pi_n} \lambda(I)}{\log n} \longrightarrow \infty$$

ce qui est la condition habituelle de convergence de l'estimateur étudié par Saleh pour une mesure aléatoire composite et pour le régressogramme.

5.5 Le processus empirique associé

Supposons que la mesure moyenne ν soit à densité par rapport à une mesure fixé sur S, λ, que nous supposerons non atomique. Donc, a fortiori, la mesure μ est aussi à densité par rapport à λ. De plus, nous supposerons que

$$\inf_{s\in B} \frac{d\nu}{d\lambda}(s) > 0 \quad \text{et} \quad \inf_{s\in B} \frac{d\mu}{d\lambda}(s) > 0,$$

où B est un borélien relativement compact de S fixé. D'après la section précédente, pour obtenir la convergence presque complète uniforme énoncée dans le théorème 5.3.4, nous pouvons choisir la suite de partitions en imposant des conditions en utilisant la mesure λ.

Nous allons maintenant nous intéresser à l'étude des lois de dimension finie du processus

$$\sqrt{\frac{n}{h_n}}(f_n(s) - g_k(s)),$$

où $h_n = \min_{I\in\Pi_n} \lambda(I)$. De plus, dans cette section nous supposerons que les partitions Π_n sont telles que $h_n = \lambda(I)$, pour tout $I \in \Pi_n$ (cette condition à un sens naturel dans le cas $S=\mathbb{R}^p$ et λ la mesure de Lebesgue sur $\mathbb{R}^p$). Fixons $s_1, \ldots, s_m \in B$ et notons $I_{n1}, \ldots, I_{nm}$ les éléments $I_k(s_1), \ldots, I_k(s_m) \in \Pi_k$ pour simplifier les notations (il faut ne pas oublier que l'indice k de la partition dépend de n). Remarquons que, comme les $s_1, \ldots, s_m$ sont tous différents les ensembles $I_{n1}, \ldots, I_{nm}$ sont deux à deux disjoints pour n suffisamment grand.

Théorème 5.5.1 *Supposons que les mesures aléatoires non négatives ξ et η vérifient*

- *Pour toute suite I_n d'éléments de $\bigcup_{n=1}^{\infty} \Pi_n$ décroissante vers un ensemble discret, et pour toute constante $N > 0$*

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\eta^2(I_n)>Nnh_n\}} \frac{1}{h_n}\eta^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\eta^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\eta^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

- *Pour tous $I, J \in \mathcal{B}$ tels que $I \cap J = \emptyset$*
  - $\mathbf{E}(\xi(I)\xi(J)) \leq \nu(I)\nu(J)$;
  - $\mathbf{E}(\xi(I)\eta(J)) \leq \nu(I)\mu(J)$;
  - $\mathbf{E}(\eta(I)\eta(J)) \leq \mu(I)\mu(J)$.

Alors, le vecteur aléatoire

$$\begin{aligned}\sqrt{\tfrac{n}{h_n}}(\overline{\eta}_n(I_{n1}) - \mu(I_{n1}), \ldots, \overline{\eta}_n(I_{nm}) - \mu(I_{nm}),\\ \overline{\xi}_n(I_{n1}) - \nu(I_{n1}), \ldots, \overline{\xi}_n(I_{nm}) - \nu(I_{nm}))\end{aligned} \tag{5.11}$$

converge en loi vers un vecteur aléatoire gaussien de moyenne nulle.

<u>**Démonstration**</u> : Pour démontrer la convergence en loi énoncée nous allons démontrer la convergence en loi de combinaisons linéaires de composantes du vecteur aléatoire (5.11) et ensuite en déduire la convergence du vecteur aléatoire (5.11) d'après le théorème de Cramer-Wold.

Considérons $c_1, \ldots, c_m, d_1, \ldots, d_m \in \mathbb{R}$ et définissons les variables aléatoires réelles

$$X_{ni} = \sum_{j=1}^{m} \left(c_j\xi_i(I_{nj}) + d_j\eta_i(I_{nj})\right),\ i = 1, \ldots, n,$$

et
$$X_n = \frac{1}{n}\sum_{i=1}^{m} X_{ni}.$$

Alors, la variable aléatoire réelle $Z_n = \sqrt{\frac{n}{h_n}}(X_n - \mathbf{E}(X_n))$ est une combinaison linéaire de composantes du vecteur aléatoire (5.11). C'est donc la convergence en loi de la variable aléatoire Z_n vers une variable aléatoire gaussienne de moyenne nulle que nous voulons établir.

Etudions d'abord le comportement de $\mathbf{V}(Z_n)$. Il est évident que

$$\mathbf{V}(Z_n) = \frac{n}{h_n}\mathbf{V}(X_n) = \frac{1}{h_n}\mathbf{V}(X_{n1}).$$

d'autre part
$$\mathbf{V}(X_{n1}) = \mathbf{V}\left[\sum_{j=1}^{m}(c_j\xi(I_{nj}) + d_j\eta(I_{nj}))\right] =$$

$$\sum_{j=1}^{m}\mathbf{V}(c_j\xi(I_{nj}) + d_j\eta(I_{nj})) + 2\sum_{j<l} cov(c_j\xi(I_{nj}) + d_j\eta(I_{nj}), c_l\xi(I_{nl}) + d_l\eta(I_{nl})).$$

Pour calculer ces variances nous tenons compte du fait

$$\mathbf{V}(c_j\xi(I_{nj}) + d_j\eta(I_{nj})) = \mathbf{E}(c_j\xi(I_{nj}) + d_j\eta(I_{nj}))^2 - \mathbf{E}^2(c_j\xi(I_{nj}) + d_j\eta(I_{nj})).$$

En calculant les esperances nous obtenons

$$\mathbf{E}(c_j\xi(I_{nj}) + d_j\eta(I_{nj})) = c_j\nu(I_{nj}) + d_j\mu(I_{nj}),$$

donc,

$$\frac{1}{h_n}\mathbf{E}^2(c_j\xi(I_{nj})+d_j\eta(I_{nj})) \leq (|c_j|+|d_j|)^2\left(\frac{\nu^2(I_{nj})}{h_n} + \frac{\mu^2(I_{nj})}{h_n} + 2\frac{\nu(I_{nj})\mu(I_{nj})}{h_n}\right),$$

et, d'après les hypothèses faites sur les partitions, il s'ensuit

$$\frac{\nu(I_{nj})}{h_n} \longrightarrow \frac{d\nu}{d\lambda}(s_j), \quad \frac{\mu(I_{nj})}{h_n} \longrightarrow \frac{d\mu}{d\lambda}(s_j),$$

d'où nous déduisons

$$\frac{1}{h_n}\mathbf{E}^2(c_j\xi(I_{nj}) + d_j\eta(I_{nj})) \longrightarrow 0.$$

D'autre part

$$\frac{1}{h_n}\mathbf{E}(c_j\xi(I_{nj})+d_j\eta(I_{nj}))^2 = c_j^2\frac{\mathbf{E}(\xi^2(I_{nj}))}{h_n}+d_j^2\frac{\mathbf{E}(\eta^2(I_{nj}))}{h_n}+2c_jd_j\frac{\mathbf{E}(\xi(I_{nj})\eta(I_{nj}))}{h_n}.$$

Nous pouvons écrire

$$\frac{\mathbf{E}(\xi^2(I_{nj}))}{h_n} = \frac{\mathbf{E}(\xi^2(I_{nj}))}{\nu(I_{nj})}\frac{\nu(I_{nj})}{h_n}.$$

$\frac{\mathbf{E}(\xi^2(I_{nj}))}{\nu(I_{nj})}$ converge d'après le théorème 5.3.1 et $\frac{\nu(I_{nj})}{h_n}$ converge d'après les hypothèses faites sur les partitions. De plus, toutes cettes limites sont strictement positives au vu des conditions imposées (5.5). De la même façon nous concluons que $\frac{\mathbf{E}(\eta^2(I_{nj}))}{h_n}$ et $\frac{\mathbf{E}[\xi(I_{nj})\eta(I_{nj})]}{h_n}$ convergent et les limites sont strictement positives.

Nous avons donc démontré que la suite

$$\frac{1}{h_n}\left[\sum_{j=1}^{m}\mathbf{V}\left(c_j\xi(I_{nj})+d_j\eta(I_{nj})\right)\right]$$

est convergente.

Considérons maintenant les termes de $\mathbf{V}(X_{n1})$ qui contiennent les covariances. Dans ce cas nous avons, en posant $A=(|c_j|+|d_j|)(|c_l|+|d_l|)$,

$$|\mathbf{E}\left[(c_j\xi(I_{nj})+d_j\eta(I_{nj}))(c_l\xi(I_{nl})+d_l\eta(I_{nl}))\right]| \leq$$

$$\leq A\mathbf{E}\left[(\xi(I_{nj})+\eta(I_{nj}))(\xi(I_{nl})+\eta(I_{nl}))\right] =$$

$$= A\mathbf{E}\left[\xi(I_{nj})\xi(I_{nl})+\eta(I_{nj})\xi(I_{nl})+\xi(I_{nj})\eta(I_{nl})+\eta(I_{nj})\eta(I_{nl})\right] \leq$$

$$\leq A\left[\nu(I_{nj})\nu(I_{nl})+\mu(I_{nj})\nu(I_{nl})+\nu(I_{nj})\mu(I_{nl})+\mu(I_{nj})\mu(I_{nl})\right],$$

car $I_{nj}\cap I_{nl}=\emptyset$ pour n suffisamment grand, donc, quand nous divisons par h_n nous pouvons majorer le terme par

$$A\left[\frac{\nu(I_{nj})}{h_n}\nu(I_{nl})+\frac{\mu(I_{nj})}{h_n}\nu(I_{nl})+\frac{\nu(I_{nj})}{h_n}\mu(I_{nl})+\frac{\mu(I_{nj})}{h_n}\mu(I_{nl})\right] \longrightarrow 0$$

d'après les hypothèses faites sur la suite de partitions. Donc

$$\frac{1}{h_n}\mathbf{E}\left[(c_j\xi(I_{nj})+d_j\eta(I_{nj}))(c_l\xi(I_{nl})+d_l\eta(I_{nl}))\right] \longrightarrow 0.$$

La convergence vers zéro des autres termes intervenant dans l'expression des covariances se déduit de la même façon. Donc ou

$$\lim_{n\to\infty} \mathbf{V}(Z_n) = 0,$$

et Z_n converge en loi vers une variable aléatoire presque sûrement nulle, ou

$$0 < \lim_{n\to\infty} \mathbf{V}(Z_n) < \infty.$$

Etudions alors ce dernier cas.

Reprenons la variable aléatoire Z_n. Nous pouvons écrire

$$\begin{aligned} Z_n &= \sqrt{\tfrac{n}{h_n}}(X_n - \mathbf{E}(X_n)) = \sqrt{\tfrac{n}{h_n}}\left(\tfrac{1}{n}\textstyle\sum_{i=1}^n (X_{ni} - \mathbf{E}(X_{ni}))\right) = \\ &= \textstyle\sum_{i=1}^n \sqrt{\tfrac{1}{nh_n}}(X_{ni} - \mathbf{E}(X_{ni})), \end{aligned}$$

donc, en posant $Z_{ni} = \sqrt{\frac{1}{nh_n}}(X_{ni} - \mathbf{E}(X_{ni}))$, nous avons $Z_n = \sum_{i=1}^n Z_{ni}$, et nous pouvons utiliser la condition de Lindeberg-Lévy pour déduire la convergence en loi de Z_n. Il faut donc démontrer

$$\frac{1}{\mathbf{V}(Z_n)}\sum_{i=1}^n \int_{\{|Z_{ni}|>\varepsilon\sqrt{\mathbf{V}(Z_n)}\}} Z_{ni}^2\, d\mathbf{P} \longrightarrow 0, \ \ \varepsilon > 0.$$

D'après ce que nous avons démontré pour la suite $\mathbf{V}(Z_n)$, il suffit de démontrer que la somme converge vers zéro. Or

$$\begin{aligned} \sum_{i=1}^n \int_{\{|Z_{ni}|>\varepsilon\sqrt{\mathbf{V}(Z_n)}\}} Z_{ni}^2\, d\mathbf{P} &= n\int_{\{|Z_{n1}|>\varepsilon\sqrt{\mathbf{V}(Z_n)}\}} Z_{n1}^2\, d\mathbf{P} = \\ n\int_{\{|X_{n1}-\mathbf{E}(X_{n1})|>\varepsilon\sqrt{nh_n\mathbf{V}(Z_n)}\}} &\tfrac{1}{nh_n}(X_{n1} - \mathbf{E}(X_{n1}))^2\, d\mathbf{P}, \end{aligned} \tag{5.12}$$

et nous pouvons majorer

$$\begin{aligned} |X_{n1} - \mathbf{E}(X_{n1})| &\le \sum_j \left[(|c_j| + |d_j|)\,(\xi_1(I_{nj}) + \eta_1(I_{nj}) + \nu(I_{nj}) + \mu(I_{nj}))\right] \le \\ &\le \sum_j \left[2(|c_j| + |d_j|)\,(\xi_1(I_{nj}) + \nu(I_{nj}) + \eta_1(I_{nj}) + \mu(I_{nj}))\right] \le \\ &\le T\,(\xi_1(I_{nj}) + \nu(I_{nj}) + \eta_1(I_{nj}) + \mu(I_{nj})), \end{aligned}$$

en posant $T = 2\max_{1\le j\le m}(|c_j| + |d_j|)$, et $I_n = \bigcup_{j=1}^{m} I_{nj}$. Remarquons que, d'après le choix des ensembles I_{nj}, $j = 1, \ldots, m$, il s'ensuit $I_n \downarrow \{s_1, \ldots, s_m\}$. Alors, nous pouvons majorer l'intégrale dans (5.12) par

$$\int_{\{(\xi(I_n)+\nu(I_n)+\eta(I_n)+\mu(I_n))^2 > \frac{\varepsilon^2}{T^2} n h_n \mathbf{V}(Z_n)\}} \frac{T^2}{h_n}(\xi(I_n)+\nu(I_n)+\eta(I_n)+\mu(I_n))^2 \, d\mathbf{P}.$$

Comme la suite $\mathbf{V}(Z_n)$ est minorée, il existe une constante $C > 0$ telle que $\frac{\varepsilon^2}{T^2}\mathbf{V}(Z_n) > 4C$, donc cette dernière intégrale est majorée par

$$\begin{aligned}
& T^2 \int_{\{(\xi(I_n)+\nu(I_n)+\eta(I_n)+\mu(I_n))^2 > 4Cnh_n\}} \frac{1}{h_n}(\xi(I_n)+\nu(I_n)+\eta(I_n)+\mu(I_n))^2 \, d\mathbf{P} \le \\
& \le 2T^2 \int_{\{(\xi(I_n)+\nu(I_n))^2+(\eta(I_n)+\mu(I_n))^2 > 2Cnh_n\}} \frac{1}{h_n}\left[(\xi(I_n)+\nu(I_n))^2 + (\eta(I_n)+\mu(I_n))^2\right] d\mathbf{P} \le \\
& \le 2T^2 \int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}\cup\{(\eta(I_n)+\mu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\left[(\xi(I_n)+\nu(I_n))^2 + (\eta(I_n)+\mu(I_n))^2\right] d\mathbf{P} \le \\
& \le 2T^2 \int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\left[(\xi(I_n)+\nu(I_n))^2 + (\eta(I_n)+\mu(I_n))^2\right] d\mathbf{P} + \\
& \quad +2T^2 \int_{\{(\eta(I_n)+\mu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\left[(\xi(I_n)+\nu(I_n))^2 + (\eta(I_n)+\mu(I_n))^2\right] d\mathbf{P}.
\end{aligned} \tag{5.13}$$

Vérifions que

$$\int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}(\xi(I_n)+\nu(I_n))^2 \, d\mathbf{P} \longrightarrow 0.$$

En fait, nous avons

$$\begin{aligned}
& \int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}(\xi^2(I_n)+\nu^2(I_n)+2\xi(I_n)\nu(I_n)) \, d\mathbf{P} \le \\
& \le \int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\xi^2(I_n) \, d\mathbf{P} + \int \frac{1}{h_n}(\nu^2(I_n)+2\xi(I_n)\nu(I_n)) \, d\mathbf{P}.
\end{aligned}$$

La deuxième intégrale de cette expression est égale à $\frac{3}{h_n}\nu^2(I_n)$. D'après les hypothèses faites sur les partitions et la définition de l'ensemble I_n il s'ensuit

$$\frac{\nu(I_n)}{h_n} \longrightarrow \sum_{j=1}^{m} \frac{d\mu}{d\nu}(s_j),$$

donc, comme la suite I_n est décroissante vers $\{s_1, \ldots, s_m\}$ et du fait que λ est supposée non atomique, il s'ensuit $\frac{3}{h_n}\nu^2(I_n) \longrightarrow 0$. Donc, il reste à vérifier que

$$\int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0.$$

Nous pouvons écrire l'ensemble sur lequel nous intégrons dans la forme

$$\{\frac{\xi^2(I_n)}{h_n} + 2\frac{\xi(I_n)\nu(I_n)}{h_n} + \frac{\nu^2(I_n)}{h_n} > Cn\}.$$

Nous savons que $\frac{\nu^2(I_n)}{h_n} \longrightarrow 0$, d'après le choix de la suite de partitions. Donc, il existe une constante $C_1 > 0$ telle que, pour tout $n \in \mathbb{N}$

$$\frac{\nu^2(I_n)}{h_n} \leq C_1.$$

Alors

$$\{\frac{\xi^2(I_n)}{h_n} + 2\frac{\xi(I_n)\nu(I_n)}{h_n} + \frac{\nu^2(I_n)}{h_n} > Cn\} \subset \{\frac{\xi^2(I_n)}{h_n} + 2\frac{\xi(I_n)\nu(I_n)}{h_n} > Cn - C_1\},$$

et nous pouvons remplacer $Cn - C_1$ par $C_3 n$ où C_3 est une constante positive convenablement choisie. Nous avons donc, pour n assez grand

$$\int_{\{(\xi(I_n)+\nu(I_n))^2 > Cnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \leq \int_{\{\frac{\xi^2(I_n)}{h_n} + 2\frac{\xi(I_n)\nu(I_n)}{h_n} > C_3 n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P}$$

$$\leq \int_{\{\frac{\xi^2(I_n)}{h_n} > \frac{C_3}{2}n\} \cup \{2\frac{\nu(I_n)\xi(I_n)}{h_n} > \frac{C_3}{2}n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \leq$$

$$\leq \int_{\{\frac{\xi^2(I_n)}{h_n} > \frac{C_3}{2}n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} + \int_{\{2\frac{\nu(I_n)\xi(I_n)}{h_n} > \frac{C_3}{2}n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P}.$$

La première intégrale converge vers zéro d'après les hypothèses du théorème. En ce qui concerne la seconde intégrale remarquons que l'ensemble d'intégration vérifie

$$\{2\frac{\nu(I_n)\xi(I_n)}{h_n} > \frac{C_3}{2}n\} \subset \{2C_2\frac{\xi(I_n)}{h_n} > \frac{C_3}{2}n\},$$

car $\nu(I_n) < 1$, pour n assez grand. En posant $C_4 = \frac{C_3^2}{16}$ et en rappelant que nous avons choisi $h_n = \frac{\log n}{n\varepsilon_n}$, où $\varepsilon_n \longrightarrow 0$ (5.8), nous avons

$$\{2\frac{\nu(I_n)\xi(I_n)}{h_n} > \frac{C_3}{2}n\} \subset \{\frac{\xi^2(I_n)}{h_n} > C_4 n^2 h_n\} =$$

$$= \{\frac{\xi^2(I_n)}{h_n} > C_4 n^2 \frac{\log n}{n\varepsilon_n}\} \subset$$

$$\subset \{\frac{\xi^2(I_n)}{h_n} > C_4 n\},$$

et donc, pour n assez grand

$$\int_{\{2\frac{\nu(I_n)\xi(I_n)}{h_n} > \frac{C_3}{2}n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \leq \int_{\{\frac{\xi^2(I_n)}{h_n} > C_4 n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P},$$

qui converge vers zéro d'après les hypothèses du théorème.

La convergence des autres intégrales de (5.13) se démontre de la même façon, car nous avons imposés des conditions qui nous permetent de reprendre la procédure utiliée pour ce cas.

Nous avons donc vérifié la condition de Lindeberg-Lévy, ce qui nous permet de déduire que la suite des variables aléatoires réelles Z_n converge en loi vers une variable aléatoire gaussienne centrée. Donc, le théorème de Cramer-Wold nour permet de déduire la convergence énoncée. ■

Ce théorème nous permet de déduire la convergence en loi des lois de dimension finie du processus empirique $\sqrt{\frac{n}{h_n}}(f_n(s) - g_k(s))$.

Corollaire 5.5.2 *Supposons les hypothèses du théorème 5.5.1. Alors le vecteur aléatoire*

$$\sqrt{\frac{n}{h_n}}\left(\frac{\overline{\eta}_n(I_{nj})}{\overline{\xi}_n(I_{nj})} - \frac{\mu(I_{nj})}{\nu(I_{nj})}\right)_{j=1,\ldots,m}$$

converge en loi vers un vecteur aléatoire gaussien de moyenne nulle.

Démonstration : Il suffit d'appliquer la δ-méthode à la fonction

$$\Psi(x_1, \ldots, x_m, y_1, \ldots, y_m) = (\frac{y_1}{x_1}, \ldots, \frac{y_m}{x_m}). \blacksquare$$

Remarquons que dans le théorème 5.5.1 nous avons posé une condition qui ressemble à la condition de Lindebeg-Lévy:

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

où $N > 0$ est une constante et I_n décroissante vers un ensemble discret. On remarque que cette condition n'est pas trop forte pour être utile. En effet, si, par exemple, ξ est processus ponctuel de Poisson, alors ξ vérifie cette condition. Pour le vérifier appelons $\nu_n = \mathbf{E}(\xi(I_n))$, le paramètre de la loi de Poisson suivie par la variable aléatoire $\xi(I_n)$. Comme la mesure moyenne est supposée à densité par rapport à une mesure non atomique il s'ensuit $\nu_n \downarrow 0$. Alors, en notant par $[x]$ le plus grand entier plus petit ou égale à x,

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} = \sum_{k=[\sqrt{Nnh_n}]+1}^{\infty} \frac{1}{h_n}k^2 e^{-\nu_n}\frac{\nu_n^k}{k!} =$$

$$= \frac{\nu_n^2 e^{-\nu_n}}{h_n} \sum_{k=[\sqrt{Nnh_n}]+1}^{\infty} \frac{\nu_n^{k-2}}{(k-2)!} + \frac{\nu_n e^{-\nu_n}}{h_n} \sum_{k=[\sqrt{Nnh_n}]+1}^{\infty} \frac{\nu_n^{k-1}}{(k-1)!}. \tag{5.14}$$

Les séries intervenant dans cette expression sont des restes de séries convergentes de somme e^{ν_n}. En utilisant une expression classique pour le reste d'une série convergente, nous obtenons que la seconde série est inférieure où égale à

$$e^{\nu_n}\frac{\nu_n^{[\sqrt{Nnh_n}]}}{[\sqrt{Nnh_n}]!}.$$

Comme $\nu_n \downarrow 0$, cette quantité converge vers zéro car $nh_n \longrightarrow \infty$ pour le choix effectué de la suite h_n. D'après les hypothèses faites sur les partitions nous savons que $\frac{\nu_n}{h_n} = \frac{\nu(I_n)}{h_n} \longrightarrow \frac{d\nu}{d\lambda}(s) < \infty$, donc le second terme de (5.14) converge vers zéro. En ce qui concerne le premier terme, nous déduisons de la même façon que la série multipliée par $e^{-\nu_n}$ converge vers zéro. De plus, $\frac{\nu_n^2}{h_n} = \frac{\nu^2(I_n)}{h_n} \longrightarrow 0$, ce qui démontre qu'un processus de Poisson vérifie la condition posée.

En ce qui concerne les cas particuliers que nous avons cités dans la section 5.2, dans les deux premiers il suffit de prendre les conditions sur la mesure aléatoire ξ, car, dans ces deux cas il se vérifie que $\eta \leq \xi$, donc toute la démonstration du théorème 5.5.1 peut être refaite en utilisant seulement les conditions sur ξ. De toute façon les conditions du théorème 5.5.1 nous donnent une indication sur la vitesse de convergence de la suite h_n vers zéro. Par exemple, pour le cas particulier du régressogramme, nous obtenons des conditions du type

$$\int_{\{\frac{Y^2}{h_n}\mathbb{1}_{\{X\in I_n\}}>Nn\}} \frac{Y^2}{h_n}\mathbb{1}_{\{X\in I_n\}}\, d\mathbf{P} = \int_{\{\frac{Y^2}{h_n}>Nn\}} \frac{Y^2}{h_n}\mathbb{1}_{\{X\in I_n\}}\, d\mathbf{P} \longrightarrow 0.$$

En applicant l'inégalité de Hölder

$$\int_{\{\frac{Y^2}{h_n}>Nn\}} \frac{Y^2}{h_n}\mathbb{1}_{\{X\in I_n\}}\, d\mathbf{P} \le \sqrt{\int \frac{Y^4}{h_n}\mathbb{1}_{\{X\in I_n\}}\, d\mathbf{P}\, \frac{\mathbf{P}\left\{Y^2 > Nnh_n\right\}}{h_n}},$$

donc, si nous supposons que la variable aléatoire Y a un moment d'ordre 4, il suffit que $\frac{1}{h_n}\mathbf{P}\left\{Y^2 > Nnh_n\right\} \longrightarrow 0$ pour avoir les conditions du théorème relatives aux intégrales vérifées. Cette convergence veut dire que la suite h_n ne peut pas converger vers zéro trop lentement. La vitesse de convergence minimale dépend de la loi de la variable aléatoire Y. Remarquons qu'au cours de ce chapitre, pour obtenir la convergence presque complète de l'estimateur, nous avons imposé que h_n ne converge pas vers zéro trop rapidement (nous avons besoin de $nh_n \longrightarrow \infty$) et cette condition est indépendante du cas particulier qu'on puisse envisager.

Pous les conditions concernant les covariances nous les avons dans les deux premiers cas car il s'agit d'échantillons. Pour le cas du régressogramme

$$\mathbf{E}\left(\eta(I)\eta(J)\right) = \mathbf{E}\left(Y^2\mathbb{1}_{\{X\in I\}}\mathbb{1}_{\{X\in J\}}\right) = 0$$

pour $I \cap J = \emptyset$, donc évidemment les conditions sur les covariances sont vérifées. Donc les théorèmes que nous avons démontrés s'appliquent au moins aux trois cas particuliers envisagés.

5.6 L'estimateur du type noyau

L'estimateur que nous avons étudié jusqu'à présent est un estimateur général qui peut être défini dans n'importe quel espace métrique séparable, complet et localement compact. Souvent, dans les applications, l'espace $\mathbf{S}$ n'est autre que $\mathbb{R}^p$, pour $p \in \mathbb{N}$, et dans ce cas il est naturel de considérer des estimateurs de forme différente. En effet, sur $\mathbb{R}^p$, il est naturel de considérer des estimateurs du type noyau pour la densité.

Considérons K une densité de probabilité continue et à support borné. Etant donné un échantillon $((\eta_1,\xi_1),\ldots,(\eta_n,\xi_n))$ du couple (η,ξ) nous pouvons définir l'estimateur

$$\varphi_n(x) = \frac{\overline{\eta}_n * K_n(x)}{\overline{\xi}_n * K_n(x)} = \frac{\frac{1}{h_n^p}\int K(\frac{x-u}{h_n})\,\overline{\eta}_n(du)}{\frac{1}{h_n^p}\int K(\frac{x-u}{h_n})\,\overline{\xi}_n(du)},$$

où h_n est une suite convergente vers zéro.

Remarquons que, si nous supposons l'existence de la densité $\frac{d\nu}{d\lambda}$ où λ est la mesure de Lebesgue, et si nous choisissons B$\in \mathcal{B}$ tel que $\inf_{x\in \mathrm{B}} \frac{d\nu}{d\lambda}(x) > 0$, alors

$$\nu * K_n(x) = \frac{1}{h_n^p}\int K(\frac{x-u}{h_n})\,\nu(du)$$

converge uniformément sur B vers $\frac{d\nu}{d\lambda}$. De même, nous avons la convergence uniforme sur B de $\mu * K_n(x)$ vers $\frac{d\mu}{d\lambda}$. Donc, dans ce cas nous obtenons la convergence uniforme sur B de $\frac{\mu * K_n(x)}{\nu * K_n(x)}$ vers $\frac{\frac{d\mu}{d\lambda}}{\frac{d\nu}{d\lambda}} = \frac{d\mu}{d\nu}$. Alors, pour étudier la convergence de l'estimateur φ_n vers $\frac{d\mu}{d\nu}$, il est suffisant d'étudier la convergence vers zéro de $\varphi_n(x) - \frac{\mu * K_n(x)}{\nu * K_n(x)}$, pour $x \in \mathrm{B} \in \mathcal{B}$.

Ce problème a été étudié par Jacob, Mendes-Lopes [23] pour le cas particulier de la distribution locale moyenne d'une mesure aléatoire composite, en considérant les mesures aléatoires η et ξ à densité par rapport à la mesure de Lebesgue λ. Ce fait est suffisant pour obtenir des propriétés très convenables pour l'estimateur φ_n. Pourtant, quelques de ces propriétés ne sont plus vérifées si on ne suppose pas l'existence des densités. En effet, si $\xi \ll \lambda$, alors, comme on l'a remarqué, $\xi * K_n(x) \longrightarrow \frac{d\xi}{d\lambda}(x)$ presque partout, tandis que, si on suppose qu'il n'existe pas la densité $\frac{d\xi}{d\lambda}$, il est facile de déduire que $\xi * K_n(x) \longrightarrow 0$ presque sûrement. De plus, si on suppose toutes les mesures à densité, on passe du problème de l'étude des mesures aléatoires au problème de l'étude des densités aléatoires. On est donc amené a poser des conditions sur les densités aléatoires au lieu de poser des conditions sur les mesures aléatoires elles-mêmes.

Considérons B$\in \mathcal{B}$ fixé et soient $\alpha, \beta_1, \beta_2 > 0$ des constantes telles que

$$\begin{aligned} &\inf_{x\in \mathrm{B}} \tfrac{d\nu}{d\lambda}(x) \geq \alpha \\ &\beta_1 \leq \tfrac{d\mu}{d\nu}(x) \leq \beta_2, \quad x \in \mathrm{B}. \end{aligned} \tag{5.15}$$

Comme $\frac{\mu * K_n(x)}{\nu * K_n(x)}$ converge uniformément sur B vers $\frac{d\mu}{d\nu}(x)$, il s'ensuit que, pour n assez grand, se vérifie

$$\frac{\beta_1}{2} \leq \frac{\mu * K_n(x)}{\nu * K_n(x)} \leq 2\beta_2, \quad x \in \mathrm{B},$$

donc, pour tout $\varepsilon > 0$

$$\frac{\varepsilon}{2} \frac{\nu * K_n(x)}{\mu * K_n(x)} \leq \frac{\varepsilon}{\beta_1}.$$

De même, comme $\nu * K_n(x)$ converge uniformément sur B vers $\frac{d\nu}{d\lambda}(x) \geq \alpha$, nous avons, pour n assez grand,

$$\inf_{x \in \mathrm{B}} \nu * K_n(x) \geq \frac{\alpha}{2},$$

donc

$$\frac{(\nu * K_n(x))^2}{\mu * K_n(x)} \geq \frac{\alpha}{2} \frac{1}{2\beta_2} = \frac{\alpha}{4\beta_2}.$$

Alors, d'après le théorème 5.1.2, il s'ensuit que, pour $\varepsilon > 0$ suffisamment petit, nous avons

$$\begin{aligned}
&\left\{ \sup_{x \in \mathrm{B}} \left| \frac{\overline{\eta}_n * K_n(x)}{\overline{\xi}_n * K_n(x)} - \frac{\mu * K_n(x)}{\nu * K_n(x)} \right| > \varepsilon \right\} = \\
&\qquad \bigcup_{x \in \mathrm{B}} \left\{ \left| \frac{\overline{\eta}_n * K_n(x)}{\overline{\xi}_n * K_n(x)} - \frac{\mu * K_n(x)}{\nu * K_n(x)} \right| > \varepsilon \right\} \subset \\
&\subset \left[\bigcup_{x \in \mathrm{B}} \left\{ |\overline{\eta}_n * K_n(x) - \mu * K_n(x)| > \frac{\varepsilon}{4} \nu * K_n(x) \right\} \right] \cup \\
&\qquad \cup \left[\bigcup_{x \in \mathrm{B}} \left\{ \left| \overline{\xi}_n * K_n(x) - \nu * K_n(x) \right| > \frac{\varepsilon}{4} \frac{(\nu * K_n(x))^2}{\mu * K_n(x)} \right\} \right] \subset \\
&\subset \left[\bigcup_{x \in \mathrm{B}} \left\{ |\overline{\eta}_n * K_n(x) - \mu * K_n(x)| > \frac{\varepsilon\alpha}{8} \right\} \right] \cup \\
&\qquad \cup \left[\bigcup_{x \in \mathrm{B}} \left\{ \left| \overline{\xi}_n * K_n(x) - \nu * K_n(x) \right| > \frac{\varepsilon\alpha}{16\beta_2} \right\} \right] = \\
&\left\{ \sup_{x \in \mathrm{B}} |\overline{\eta}_n * K_n(x) - \mu * K_n(x)| > \frac{\varepsilon\alpha}{8} \right\} \cup \left\{ \sup_{x \in \mathrm{B}} \left| \overline{\xi}_n * K_n(x) - \nu * K_n(x) \right| > \frac{\varepsilon\alpha}{16\beta_2} \right\},
\end{aligned}$$

donc, comme dans l'étude de l'estimateur du type régressogramme, pour étudier la convergence uniforme presque complète de φ_n vers $\frac{d\mu}{d\nu}$ on est ramené à étudier les convergences uniformes presque complètes de $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$ et $|\overline{\eta}_n * K_n(x) - \mu * K_n(x)|$ vers zéro.

Pour déduire la convergence uniforme presque complète pour l'estimateur du type régressogramme on a été forcé d'imposer des conditions sur les moments des mesures aléatoires (condition (5.4)), et que les sur-martingales définies dans le théorème 5.3.1 soient uniformément bornées dans $\bigcup_{n\in\mathbb{N}} \Pi_n$, (5.5). Dans le cas présent nous pouvons affaiblir ces conditions car il n'est pas nécessaire considérer tous les ensembles dans $\bigcup_{n\in\mathbb{N}} \Pi_n$. En effet, considérons

$$\left|\overline{\xi}_n * K_n(x) - \nu * K_n(x)\right| = \left|\frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)\frac{1}{n}\sum_{i=1}^n(\xi_i-\nu)(du)\right| =$$

$$\frac{1}{nh_n^p}\left|\sum_{i=1}^n\int K\left(\frac{x-u}{h_n}\right)(\xi_i-\nu)(du)\right|.$$

Nous avons supposé la densité de probabilité K à support borné. Soit S_K le support de K, et définissons les ensembles $S_K(x,h) = x - hS_K$, $x, h \in \mathbb{R}$. Alors, d'après l'expression de l'intégrale il est suffisant de considérer les valeurs de la variable u tels que

$$\frac{x-u}{h_n} \in S_K \Leftrightarrow u \in S_K(x,h_n),$$

donc, il est suffisant d'imposer des conditions sur les ensembles $S_K(x,h_n)$. Tous les autres ensembles n'ont pas d'importance car ils n'ajoutent rien aux intégrales. Considérons alors la condition: il existe $M > 0$ tel que, pour tout $j > 2$ se vérifie

$$\begin{aligned} &\mathbf{E}\left[\eta^j(S_K(x,h_n))\right] \le M^{j-2}j!\,\mathbf{E}\left[\eta^2(S_K(x,h_n))\right], \quad x \in \mathrm{B} \\ &\mathbf{E}\left[\xi^j(S_K(x,h_n))\right] \le M^{j-2}j!\,\mathbf{E}\left[\xi^2(S_K(x,h_n))\right], \quad x \in \mathrm{B}. \end{aligned} \tag{5.16}$$

De même, il nous faut modifier les conditions (5.5) en considérant seulement les ensembles de la forme $S_K(x,h_n)$:

$$\begin{aligned} &\exists_{C_1,C_2>0}\,\forall_{n\in\mathbb{N},x\in\mathrm{B}}\; C_1 \le \frac{\mathbf{E}\left[\xi^2(S_K(x,h_n))\right]}{h_n^p} \le C_2; \\ &\exists_{C_3,C_4>0}\,\forall_{n\in\mathbb{N},x\in\mathrm{B}}\; C_3 \le \frac{\mathbf{E}\left[\eta^2(S_K(x,h_n))\right]}{h_n^p} \le C_4. \end{aligned} \tag{5.17}$$

Remarquons que, comme $h_n \longrightarrow 0$, il s'ensuit $S_K(x,h_n) \downarrow \{x\}$.

Théorème 5.6.1 *Choisissons la suite h_n telle que*

$$h_n^p = \frac{\log n}{n\varepsilon_n},$$

où $\varepsilon_n \longrightarrow 0$. Si les mesures aléatoires ξ et η vérifient les conditions (5.16) et (5.17) alors $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$ et $|\overline{\eta}_n * K_n(x) - \mu * K_n(x)|$ convergent presque complètement vers zéro.

Démonstration : Vérifions la convergence pour $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$, l'autre se démontrant de façon équivalente.

Comme on l'a remarqué, nous avons

$$|\overline{\xi}_n * K_n(x) - \nu * K_n(x)| = \frac{1}{nh_n^p}\left|\sum_{i=1}^{n} \int K(\frac{x-u}{h_n})\,(\xi_i - \nu)(du)\right|.$$

Soit $j \geq 2$, alors

$$\mathbf{E}\left(\int K(\frac{x-u}{h_n})\,\xi_i(du)\right)^j \leq \mathbf{E}\left(\int ||K||\mathbb{1}_{S_K(x,h_n)}(u)\,\xi_i(du)\right)^j =$$

$$= \mathbf{E}\left[||K||^j \xi_i^j(S_K(x,h_n))\right] \leq$$

$$\leq ||K||^j M^{j-2} j!\,\mathbf{E}\left[\xi^2(S_K(x,h_n))\right],$$

d'après les hypothèses (5.16). Alors, en utilisant l'inégalité du théorème 5.1.3 nous obtenons, pour $\varepsilon > 0$

$$\mathbf{P}\left\{\left|\overline{\xi}_n * K_n(x) - \nu * K_n(x)\right| > \varepsilon\right\} =$$

$$\mathbf{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} \int K(\frac{x-u}{h_n})\,(\xi_i - \nu)(du)\right| > \varepsilon h_n^p\right\} \leq$$

$$\leq 2\exp\left(-\frac{n\varepsilon^2 h_n^{2p}}{4||K||^2\mathbf{E}\left[\xi^2(S_K(x,h_n))\right] + 2M||K||\varepsilon h_n^p}\right) \leq$$

$$\leq 2\exp\left(-\frac{n\varepsilon^2 h_n^p}{M_1}\right),$$

où $M_1 = 4\|K\|^2 C_2 + 2M\|K\|\varepsilon$. Donc, tenant compte du choix de la suite h_n, nous obtenons

$$\mathbf{P}\left\{\left|\overline{\xi}_n * K_n(x) - \nu * K_n(x)\right| > \varepsilon\right\} \leq 2n^{-\frac{\varepsilon^2}{M_1 \varepsilon_n}}.$$

Comme la série de terme général $n^{-\frac{\varepsilon^2}{M_1 \varepsilon_n}}$ converge, car $\varepsilon_n \longrightarrow 0$, il s'ensuit la convergence ponctuelle presque complète de $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$ vers zéro.

Théorème 5.6.2 *Supposons vérifiées les conditions du théorème précédent. Si, de plus, le noyau K est lipschitzien alors $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$ et $|\overline{\eta}_n * K_n(x) - \mu * K_n(x)|$ convergent uniformément presque complètement vers zéro sur B.*

Démonstration : Vérifions la convergence pour $|\overline{\xi}_n * K_n(x) - \nu * K_n(x)|$ en suivant la procédure proposée par Collomb [8], l'autre convergence se démontrant de façon équivalente.

Considérons les récouvrement finis de B par des boules $\mathrm{B}_{nj} = B(x_{nj}, \rho_{nj})$, $x_{nj} \in \mathrm{B}$, $j = 1, \ldots, l_n$. Alors, comme pour tout indice $j = 1, \ldots, l_n$

$$|\overline{\xi}_n * K_n(x) - \nu * K_n(x)| \leq |\overline{\xi}_n * K_n(x_{nj}) - \nu * K_n(x_{nj})| +$$

$$+ |\overline{\xi}_n * (K_n(x_{nj}) - K_n(x)) - \nu * (K_n(x_{nj}) - K_n(x))|,$$

il s'ensuit

$$\sup_{x \in \mathrm{B}} |\overline{\xi}_n * K_n(x) - \nu * K_n(x)| \leq \max_j |\overline{\xi}_n * K_n(x_{nj}) - \nu * K_n(x_{nj})| +$$

$$+ \max_j \sup_{x \in \mathrm{B}_{nj}} \left\{|\overline{\xi}_n * (K_n(x_{nj}) - K_n(x)) - \nu * (K_n(x_{nj}) - K_n(x))|\right\}.$$

Comme le noyau K est lipschitzien, il existe des constantes positives C et α telles que $|K(x) - K(y)| \leq C \parallel x - y \parallel^\alpha$, donc

$$|K_n(x_{nj}) - K_n(x)| \leq \frac{C}{h_n^{p+\alpha}} \parallel x_{nj} - x \parallel^\alpha .$$

Alors, pour tout $x \in \mathrm{B}_{nj}$

$$|\overline{\xi}_n * (K_n(x_{nj}) - K_n(x)) - \nu * (K_n(x_{nj}) - K_n(x))| \leq$$

$$\leq \int |K_n(x_{nj}) - K_n(x)|\, \overline{\xi}_n(du) + \int |K_n(x_{nj}) - K_n(x)|\, \nu(du) \leq$$

$$\leq \frac{C}{h_n^{p+\alpha}} \rho_{nj}^{\alpha} \left(\overline{\xi}_n(\mathrm{B}') + \nu(\mathrm{B}')\right)$$

où $\mathrm{B}' = \mathrm{B} - h_1 S_K \in \mathcal{B}$. D'après la loi forte des grand nombres $\left(\overline{\xi}_n(\mathrm{B}') + \nu(\mathrm{B}')\right)$ est presque sûrement convergente, donc presque sûrement borné. Donc, en choisissant les rayons telles que $\rho_n = \max_{1 \leq j \leq l_n} \rho_{nj}^{\alpha} \leq h_n^{p+\alpha+1}$, il s'ensuit

$$\max_j \sup_{x \in \mathrm{B}_{nj}} \left\{|\overline{\xi}_n * (K_n(x_{nj}) - K_n(x)) - \nu * (K_n(x_{nj}) - K_n(x))|\right\} \leq$$

$$\leq \frac{C\rho_n}{h_n^{p+\alpha}} \left(\overline{\xi}_n(\mathrm{B}') + \nu(\mathrm{B}')\right) \leq$$
$$\leq C h_n \left(\overline{\xi}_n(\mathrm{B}') + \nu(\mathrm{B}')\right).$$

Comme, $h_n \nu(\mathrm{B}') \longrightarrow 0$, pour n assez grand

$$\mathbf{P}\left\{C h_n \left(\overline{\xi}_n(\mathrm{B}') + \nu(\mathrm{B}')\right) > \varepsilon\right\} \leq$$
$$\leq \mathbf{P}\left\{C h_n \left(\overline{\xi}_n(\mathrm{B}') - \nu(\mathrm{B}')\right) > \frac{\varepsilon}{2}\right\}$$

en posant $T = \frac{\varepsilon}{2C}$, cette dernière probabilité est égale à

$$\mathbf{P}\left\{\frac{1}{n}\sum_{i=1}^{n} \xi_i(\mathrm{B}') - \nu(\mathrm{B}') > \frac{T}{h_n}\right\} \leq$$
$$\leq \exp\left[-\frac{\frac{nT^2}{h_n^2}}{4\mathbf{E}(\xi^2(\mathrm{B}')) + 2M\frac{T}{h_n}}\right] =$$
$$= \exp\left[-\frac{\frac{nT^2}{h_n}}{4h_n\mathbf{E}(\xi^2(\mathrm{B}')) + 2MT}\right]$$

en utilisant l'inégalité (5.1). Pour n assez grand le dénominateur est plus petit que $4MT$, donc nous pouvons encore majorer cette quantité par

$$\exp\left[-4\frac{T}{M}\frac{n}{h_n}\right] = \exp\left[-4\frac{T}{M}n\frac{h_n^p}{h_n^{p+1}}\right] =$$

$$= \exp\left[-4\frac{T}{M}\frac{\log n}{\varepsilon_n h_n^{p+1}}\right] = n^{-\frac{4T}{M\varepsilon_n h_n^{p+1}}} < n^{-2}$$

pour n assez grand, car $\varepsilon_n h_n^{p+1} \longrightarrow 0$.

D'après le théorème précédent nous savons que, pour tout $x \in$B fixé

$$\mathbf{P}\left(|\bar{\xi}_n * K_n(x) - \nu * K_n(x)| > \varepsilon\right) \leq 2n^{-\frac{\varepsilon^2}{M_1\varepsilon_n}}$$

où M_1 est une constante positive indépendante de x. Alors

$$\mathbf{P}\left(\max_{1\leq j\leq l_n} |\bar{\xi}_n * K_n(x_{nj}) - \nu * K_n(x_{nj})| > \varepsilon\right) \leq 2l_n n^{-\frac{\varepsilon^2}{M_1\varepsilon_n}}.$$

En choisissant le récouvrement de telle façon que $l_n \leq A(h_n^{\frac{p}{\alpha}})^{-(p+\alpha+1)}$, où A ne dépend que du borélien relativement compact B, nous obtenons une majoration de la probabilité par le terme général d'une série convergente, ce qui termine la démonstration. ■

5.7 Le processus empirique associé

Comme pour l'estimateur du type régressogramme nous allons étudier la convergence en loi des lois de dimension finie du processus empirique associé à l'estimateur φ_n:

$$\sqrt{nh_n^p}\left(\varphi_n(x) - \frac{\mu * K_n(x)}{\nu * K_n(x)}\right). \tag{5.18}$$

Remarquons que, par rapport au processus empirique associé à l'estimateur du type régressogramme, nous avons changé la constante par laquelle nous multiplions. Ce fait est lié au propriétés de convergence des variances, comme on le vera plus tard. La technique à utiliser pour démontrer la convergence en loi des lois de dimension finie sera la même que nous avons utilisé pour le processus empirique associé à l'estimateur du type régressogramme. C'est à dire, nous utiliserons le théorème de Cramer-Wold et la δ-méthode. De plus, remarquons que, comme $h_n^p \longrightarrow 0$, il s'ensuit que $\frac{\nu(S_k(x,h_n))}{h_n^p} \longrightarrow \frac{d\nu}{d\lambda}(x)$, $x \in$B. En fait, il suffit de choisir une suite de partitions $\{\Pi_n\}$ de B telle que $h_n^p = \min_{I\in\Pi_n} \lambda(I)$, d'après l'étude de l'estimateur du type régressogramme.

Les conditions à imposer pour obtenir les convergences en loi des lois de dimension finie sont équivalentes au conditions imposées dans le cas de l'estimateur du type régressogramme, théorème 5.5.1. Démontrons d'abord un résultat auxiliaire.

Théorème 5.7.1 *Suposons que les mesures aléatoires non négatives ξ et η sont presque sûrement discrètes et vérifient: pour tous A, B boréliens de $\mathbb{R}^p$ disjoints*

$$\mathbf{E}\left[\xi(A)\eta(B)\right] \leq \nu(A)\eta(B).$$

Alors, pour tout x fixé, la suite

$$\{h_n^p\mathbf{E}\left[\xi * K_n(x)\eta * K_n(x)\right]\}_{n\in\mathbb{N}}$$

converge vers une limite finie.

<u>**Démonstration**</u> : Remarquons que d'après le théorème de Fubini

$$h_n^p\mathbf{E}\left[\xi * K_n(x)\eta * K_n(x)\right] = \frac{1}{h_n^p}\mathbf{E}\left[\int K\left(\frac{x-u}{h_n}\right)\,\xi(du)\int K\left(\frac{x-u}{h_n}\right)\,\eta(du)\right]$$

$$= \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)\mathbf{E}(\xi\otimes\eta)(du,dv).$$

Soit Δ la diagonale de $\mathbb{R}^p\times\mathbb{R}^p$ et Γ la restriction de $\mathbf{E}(\xi\otimes\eta)$ à Γ. Si nous posons $H = \mathbf{E}(\xi\otimes\eta) - \Gamma$ nous définissons une mesure non négative qui n'a pas de masse sur la diagonale Δ et $\mathbf{E}(\xi\otimes\eta) = H + \Gamma$, donc

$$\frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)\mathbf{E}(\xi\otimes\eta)(du,dv) =$$

$$= \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)d\Gamma + \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)dH.$$

Comme, par hypothèse $\mathbf{E}\left[\xi(A)\eta(B)\right] \leq \nu(A)\eta(B)$ pour $A\cap B = \emptyset$, il s'ensuit

$H \leq \nu \otimes \mu$, et alors

$$\begin{aligned}
\frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)dH &\leq \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)d\nu\otimes\mu \\
&= \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)\nu(du)\int K\left(\frac{x-u}{h_n}\right)\mu(du) \\
&\leq \frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)\nu(du)\parallel K \parallel \mu\left(S_K(x,h_n)\right)
\end{aligned}$$

qui converge vers zéro.

Il reste à voir l'intégrale par rapport à la mesure Γ. Posons Γ^* la projection sur $\mathbb{R}^p$ de Γ. Alors nous pouvons écrire

$$\frac{1}{h_n^p}\int K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)d\Gamma = \frac{1}{h_n^p}\int K^2\left(\frac{x-u}{h_n}\right)d\Gamma^* \tag{5.19}$$

La fonction K^2 est un noyau continue borné sur S_K, donc pour que (5.19) converge il suffit que Γ^* soit à densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^p$, λ. Remarquons que pour tout ensemble $A \subset \Delta$ il existe un unique ensemble $A^* \subset \mathbb{R}^p$ tel que $A = \{(u,u), u \in A^*\}$, et dans ce cas $\Gamma^*(A^*) = \Gamma(A)$. Si $\lambda(A^*) = 0$, comme nous avons supposé ν absolument continue par rapport à λ il s'ensuit $\nu(A^*) = \mathbf{E}(\xi(A^*)) = 0$, donc $\xi(A^*) = 0$ presque sûrement. Comme ξ et η sont presque sûrement discrètes, il s'ensuit $\xi \otimes \eta(A) = 0$ presque sûrement, donc $\Gamma(A) = 0$ et, finalement, $\Gamma^*(A) = 0$. Donc l'intégrale (5.19) converge vers $\frac{d\Gamma^*}{d\lambda}(x)$. ■

Remarques: **1.** Ce théorème est vrai même dans le cas $\xi = \eta$. **2.** Pour que $\frac{d\Gamma^*}{d\lambda}(x)$ soit non nulle il suffit que ξ et η aient en commum au moins un atome avec probabilité non nulle, ce qui est le cas dans la plupart des applications. **3.** Si nous supposons que $\mathbf{E}[\xi(A)\eta(B)] \leq \nu(A)\mu(B)$ pour tous A, B (non forcément disjoints) alors la mesure Γ est nulle. En effet, pour chaque $n \in \mathbb{N}$, soit $\{A_{in}\}_{i\in\mathbb{N}}$ une partition de B en ensembles de diamètre inférieur à $\frac{1}{n}$ telle que $\max_{i\in\mathbb{N}} \nu(A_{in}) \longrightarrow 0$. Alors

$$\Gamma(\Delta) \leq \Gamma\left(\bigcup_{i=1}^{\infty} A_{in}\times A_{in}\right) \leq \mathbf{E}(\xi\otimes\eta)\left(\bigcup_{i=1}^{\infty} A_{in}\times A_{in}\right) \leq$$

$$\leq \sum_{i=1}^{\infty}\nu(A_{in})\mu(A_{in}) \leq \mu(B)\max_{i\in\mathbb{N}}\nu(A_{in}) \longrightarrow 0.$$

Théorème 5.7.2 *Soient $x_1, \ldots, x_m \in B$. Supposons que les mesures aléatoires non négatives ξ et η sont presque sûrement descrètes, ont au moins un atome en commun avec probabilité non nulle et vérifient*

- *si $I_n = \bigcup_{i=1}^{m} S_K(x, h_n)$,*

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\eta^2(I_n)>Nnh_n\}} \frac{1}{h_n}\eta^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\xi^2(I_n)>Nnh_n\}} \frac{1}{h_n}\eta^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

$$\int_{\{\eta^2(I_n)>Nnh_n\}} \frac{1}{h_n}\xi^2(I_n)\, d\mathbf{P} \longrightarrow 0,$$

  pour toute constante $N > 0$,

- *si, pour tous boréliens A, B disjoints*
  - $\mathbf{E}[\xi(A)\xi(B)] \leq \nu(A)\nu(B)$,
  - $\mathbf{E}[\xi(A)\eta(B)] \leq \nu(A)\mu(B)$,
  - $\mathbf{E}[\eta(A)\eta(B)] \leq \mu(A)\mu(B)$.

Alors, le vecteur aléatoire

$$\begin{aligned}\sqrt{nh_n^p}(\overline{\eta}_n * K_n(x_1) - \mu * K_n(x_1), \ldots, \overline{\eta}_n * K_n(x_m) - \mu * K_n(x_m), \\ \overline{\xi}_n * K_n(x_1) - \nu * K_n(x_1), \ldots, \overline{\xi}_n * K_n(x_m) - \nu * K_n(x_m))\end{aligned} \tag{5.20}$$

converge en loi vers un vecteur aléatoire gaussien centré.

<u>**Démonstration**</u> : Posons

$$X_{ni} = \sum_{j=1}^{m} (c_j\xi_i + d_j\eta_i) * K_n(x_j), \quad c_1, \ldots, c_m, d_1, \ldots, d_m \in \mathbb{R},$$

et

$$Z_{ni} = \sqrt{\frac{h_n^p}{n}}(X_{ni} - \mathbf{E}(X_{ni})), \; Z_n = \sum_{i=1}^{n} Z_{ni}.$$

Alors, Z_n n'est autre qu'une combinaison linéaire de composantes du vecteur aléatoire (5.20). Etudions la variance de Z_n.

$$\mathbf{V}(Z_n) = n\mathbf{V}(Z_{n1}) = h_n^p\mathbf{V}(X_{n1}) =$$

$$h_n^p \sum_{j=1}^{m} \mathbf{V}\left[(c_j\xi + d_j\eta) * K_n(x_j)\right] +$$

$$+2h_n^p \sum_{j>l} cov\left((c_j\xi + d_j\eta) * K_n(x_j), (c_l\xi + d_l\eta) * K_n(x_l)\right).$$

d'autre part

$$h_n^p\mathbf{V}\left[(c_j\xi + d_j\eta) * K_n(x_j)\right] = h_n^p\mathbf{V}\left[\frac{1}{h_n^p}\int K(\frac{x_j - u}{h_n})\,(c_j\xi + d_j\eta)(du)\right] =$$

$$h_n^p\mathbf{E}\left[\frac{1}{h_n^p}\int K(\frac{x_j - u}{h_n})\,(c_j\xi + d_j\eta)(du)\right]^2 - h_n^p\left[\frac{1}{h_n^p}\int K(\frac{x_j - u}{h_n})\,(c_j\nu + d_j\mu)(du)\right]^2.$$

Or

$$h_n^p\left[\frac{1}{h_n^p}\int K(\frac{x_j - u}{h_n})\,(c_j\nu + d_j\mu)(du)\right]^2 \leq$$

$$h_n^p\left[||K||\frac{|c_j|\nu(S_K(x,h_n)) + |d_j|\mu(S_K(x,h_n))}{h_n^p}\right]^2,$$

ce qui converge vers zéro, d'après la remarque qui précéde ce théorème, et tenant compte du fait $h_n^p \longrightarrow 0$. De plus, d'après le théorème précédent, 5.7.1, il s'ensuit que

$$h_n^p\mathbf{E}\left[\frac{1}{h_n^p}\int K(\frac{x_j - u}{h_n})\,(c_j\xi + d_j\eta)(du)\right]^2$$

converge et la limite est strictement positive.

En ce qui concerne les termes avec les covariances nous avons, d'une part, en posant $A = (|c_j| + |d_j|)(|c_l| + |d_l|)$

$$h_n^p\frac{1}{h_n^{2p}}\int K(\frac{x_j - u}{h_n})\,(c_j\nu + d_j\mu)(du)\int K(\frac{x_l - u}{h_n})\,(c_l\nu + d_l\mu)(du) \leq$$

$$\leq h_n^p\frac{1}{h_n^{2p}}||K||^2 A[\nu(S_K(x_j,h_n))\nu(S_K(x_l,h_n)) + \nu(S_K(x_j,h_n))\mu(S_K(x_l,h_n)) +$$

$$+\mu(S_K(x_j,h_n))\nu(S_K(x_l,h_n)) + \mu(S_K(x_j,h_n))\mu(S_K(x_l,h_n))],$$

ce qui converge vers zéro d'après la remarque qui précède ce théorème. D'autre part

$$h_n^p \left| \mathbf{E}\left[\frac{1}{h_n^{2p}} \int K(\frac{x_j-u}{h_n})\,(c_j\xi + d_j\eta)(du) \int K(\frac{x_l-u}{h_n})\,(c_l\xi + d_l\eta)(du) \right] \right| \leq$$

$$\leq h_n^p \mathbf{E}\left[||K||^2 A \frac{(\xi(S_K(x_j,h_n)) + \eta(S_K(x_j,h_n)))\,(\xi(S_K(x_l,h_n)) + \eta(S_K(x_l,h_n)))}{h_n^{2p}} \right] \leq$$

$$\leq h_n^p \frac{||K||^2 A}{h_n^{2p}} [\nu(S_K(x_j,h_n))\nu(S_K(x_l,h_n)) + \nu(S_K(x_j,h_n))\mu(S_K(x_l,h_n)) +$$

$$+\mu(S_K(x_j,h_n))\nu(S_K(x_l,h_n)) + \mu(S_K(x_j,h_n))\mu(S_K(x_l,h_n))],$$

ce qui converge vers zéro.

Nous avons donc

$$0 \leq \lim_{n\to\infty} \mathbf{V}(Z_n) < \infty.$$

Si $\lim_{n\to\infty} \mathbf{V}(Z_n) = 0$ alors la variable aléatoire Z_n converge en loi vers une variable aléatoire presque sûrement égale à zéro.

Etudions donc le cas $0 < \lim_{n\to\infty} \mathbf{V}(Z_n) < \infty$. Alors, pour vérifier la condition de Lindeberg-Lévy, pour la suite Z_n, il suffit de démontrer

$$\sum_{i=1}^{n} \int_{\{|Z_{ni}|>\varepsilon\sqrt{\mathbf{V}(Z_n)}\}} Z_{ni}^2 \, d\mathbf{P} \longrightarrow 0, \quad \varepsilon > 0.$$

Nous pouvons majorer les variables Z_{ni} de la façon suivante

$$|Z_{ni}| = |X_{ni} - \mathbf{E}(X_{ni})| \leq \sum_{j=1}^{m} |(c_j\xi_i + d_j\eta_i) * K_n(x_j)| + \sum_{j=1}^{m} |(c_j\nu + d_j\mu) * K_n(x_j)| =$$

$$= \frac{1}{h_n^p} \sum_{j=1}^{m} \left[\left| \int K(\frac{x_j-u}{h_n})\,(c_j\xi_i + d_j\eta_i)(du) \right| + \left| \int K(\frac{x_j-u}{h_n})\,(c_j\nu + d_j\mu)(du) \right| \right] \leq$$

$$\leq \frac{||K||(|c_j| + |d_j|)}{h_n^p} \sum_{j=1}^{m} [\xi_i(S_K(x_j,h_n)) + \eta_i(S_K(x_j,h_n)) + \nu(S_K(x_j,h_n)) + \mu(S_K(x_j,h_n))] =$$

$$= \frac{T}{h_n^p} \sum_{j=1}^{m} (\xi_i + \eta_i + \nu + \mu)(S_K(x_j,h_n)) = \frac{T}{h_n^p} (\xi_i + \eta_i + \nu + \mu)(I_n),$$

en posant $T = ||K|| \max_{1\leq j\leq m}(|c_j| + |d_j|)$. Alors

$$\sum_{i=1}^{n} \int_{\left\{|Z_{ni}|>\epsilon\sqrt{\mathbf{V}(Z_n)}\right\}} Z_{ni}^2 \, d\mathbf{P} = n \int_{\left\{|Z_{n1}|>\epsilon\sqrt{\mathbf{V}(Z_n)}\right\}} Z_{n1}^2 \, d\mathbf{P} =$$

$$= \int_{\left\{|X_{n1}-\mathbf{E}(X_{n1})|>\epsilon\sqrt{\frac{n}{h_n^p}\mathbf{V}(Z_n)}\right\}} h_n^p (X_{n1} - \mathbf{E}(X_{n1}))^2 \, d\mathbf{P} \leq$$

$$\leq T^2 \int_{\left\{(\xi(I_n)+\eta(I_n)+\nu(I_n)+\mu(I_n))^2>\frac{\epsilon^2}{T^2}nh_n^p\mathbf{V}(Z_n)\right\}} \frac{1}{h_n^p}(\xi(I_n)+\eta(I_n)+\nu(I_n)+\mu(I_n))^2 \, d\mathbf{P}.$$

Nous pouvons démontrer que cette intégrale converge vers zéro comme nous l'avons fait pour l'intégrale (5.13), tenant compte du fait que les mesures vérifient les conditions (5.17) au lieu de (5.5). Donc nous avons vérifié la condition de Lindeberg-Lévy pour la suite Z_{ni}, ce qui termine la démonstration du théorème. ■

Maintenant, comme dans le cas de l'estimateur du type régressogramme, nous déduisons la convergence en loi des lois de dimension finie du processus empirique (5.18).

Corollaire 5.7.3 *Supposons vérifiées les conditions du théorème 5.7.2. Alors, le vecteur aléatoire*

$$\sqrt{nh_n^p}\left(\frac{\overline{\eta}_n * K_n(x_j)}{\overline{\xi}_n * K_n(x_j)} - \frac{\mu * K_n(x_j)}{\nu * K_n(x_j)}\right)_{j=1,\dots,m}$$

converge en loi vers un vecteur aléatoire gaussien de moyenne nulle.

Démonstration : D'après le théorème précédent il suffit d'appliquer la δ-méthode à la fonction

$$\psi(x_1,\dots,x_m,y_1,\dots,y_m) = \left(\frac{y_1}{x_1},\dots,\frac{y_m}{x_m}\right). \blacksquare$$

Bibliographie

[1] Alexandroff, A.D., Additive set functions in abstract spaces, Mat. Sb. 8(1940), 307-348; 9(1941), 563-628; 13(1943), 169-238.

[2] Araujo, A., Giné, E., Type, cotype and Lévy measures in Banach spaces, Ann. Probab., 6(1978), n°4, 637-643.

[3] Araujo, A., Giné, E., *The Central Limit Theorem for real and Banach space valued random variables*, John Wiley & Sons, New York, 1980.

[4] Aronszajn, N., The theory of reproducing kernels, Trans. Amer. Math. Soc., 68(1950), 337-404.

[5] Billingsley, P., *Convergence of Probability Measures*, John Wiley & Sons, 1968.

[6] Bonkian, S.M., *Contribution a l'étude des mesures aléatoires du second ordre*, Thèse du 3ème cycle, Université des Sciences et Techniques de Lille I, 1983.

[7] Bourbaki, N., *Intégration*, 1re éd., Chap. 1-4, Hermann, Paris 1952.

[8] Collomb, G., Propriétés de convergence presque complète du prédicteur à noyau, Z. Wahrsch. verw. Gebiete, 66(1984), 441-460.

[9] Dinculeanu, N., *Vector Measures*, Pergamon Press, 1967.

[10] Dunford, N., Schwartz, J., *Linear Operators, Part I: General Theory*, Interscience Publishers Inc., New York, 1957.

[11] Fröberg, C.-E., *Introduction to Numerical Analysis*, Addison-Wesley Publishing Co., 1965.

[12] Gihman, I.I., Skorohod, A.V., *The Theory of Stochastic Processes II*, Springer-Verlag, 1975.

[13] Geffroy, J., Quidel, P. Convergences stochastiques des répartitions pontuelles aléatoires, Revista da Faculdade de Ciências da Univ. de Coimbra, vol. XLIX.

[14] Geffroy, J., Etude de la convergence de régressogramme, Pub. Inst. Stat. Univ. Paris, vol. XXV, fasc. 1-2, 41-55.

[15] Guilbart, C., Caractérisations des produits scalaires sur l'espace des mesures bornées à signa dont la topologie trace sur l'espace des mesures de probabilité est la topologie faible et théorème de "Glivenko-Cantelli" associé, Université de Lille I, Publications Internes de l'U.E.R. de Mathématiques Pures et Appliquées, n$^{\underline{o}}$108(1977).

[16] Halmos, P.R., *Measure Theory*, D. Van Nostrand Co. Inc., Pinceton, New Jersey, 1950.

[17] Harris, T.E., Counting measures, monotone random set functions, Z. Wahrsch. verw. Gebiete, 10(1968), 102-119.

[18] Harris, T.E., Random measures and motions of point processes, Z. Wahrsch. verw. Gebiete, 18(1971), 85-115.

[19] Herrndorf, N., A functional central limit theorem for weakly dependent sequences of random variables, Ann. Probab., 12(1984), n$^{\underline{o}}$1, 141-153.

[20] Hida, T., *Brownian Motion*, Springer-Verlag, 1980.

[21] Jacob, P., Convergence uniforme à distance finie de mesures signées, Ann. Inst. Henri Poincaré, 15(1979), n$^{\underline{o}}$4, 355-373.

[22] Jacob, P., Niéré, L., Conbtribution a l'estimation des lois de Palm d'une mesure aléatoire, Pub. Inst. Stat. Univ. Paris (à paraître)

[23] Jacob, P., Mendes-Lopes, N., Un théorème central limite functionnel pour un estimateur d'un noyau de transition, Portugaliae Mathematica (à paraître)

[24] Jacob, P., Comunication privé.

[25] Jagers, P., Aspects of random measures and point processes, dans Advances in Probability and Related Topics 3, 179-239, Marcel Decker, New York, 1974.

[26] Kallenberg, O., *Random Measures*, Academic Press, 1976.

[27] Kelly, J.K., *General Topology*, D. Van Nostrand Co. Inc., New York, 1955.

[28] Lang, S., *Algebra*, Addisson-Wesley Publishing Co., 1965.

[29] Lima, E.L., *Espaços métricos*, Projecto Euclides, IMPA, Rio de Janeiro, 1983.

[30] Marle, C.-M., *Mesures et Probabilités*, Enseignement des Sciences, Hermann, Paris, 1974.

[31] Matthes, K., Kerstan, J., Mecke, J., *Infinitely divisible point processes*, Wiley, 1978.

[32] Mendes-Lopes, N., Convergence et optimisation d'un estimateur de la répartition locale des couleurs d'un processus ponctuel chromatique, Pub. Inst. Stat. Univ. Paris, XXIX, 2(1984), 49-68.

[33] Nawrotzki, K., Ein Grenzwertsatz für homogene zufällige Punktfolgen (Verallgemeinerung eines Satzes von A. Rényi), Math. Nachr. 24(1962), 201-217.

[34] Niéré, L., Estimation des distributions de Palm. Mesure intensité conditionnelle, Thèse de 3ème cycle, Lille, 1987.

[35] Oliveira, P.E., Convergence de suite de mesures et convergence des masses, Pub. IRMA, Lille, 13(1988), II.

[36] Oliveira, P.E., Invariance principles in $L^2[0,1]$, Comment. Math. Univ. Carolinae, 31(1990), nº2.

[37] Oliveira, P.E., Relative compactness in weak invariance principles, Transactions of the 11^{th} Prague Conference on information theory, statistical decision functions and random processes, 1990.

[38] Parthasarathy, K.R., *Probability measures on metric spaces*, Academic Press, London, 1967.

[39] Peligrad, M., An invariance principle for dependent random variables, Z. wahrsch. verw. Gebiete, 57(1981), n°4, 495-507.

[40] Prékopa, A., On stochastic set functions I, II, III, Acta Math. Hungaricae, 7(1956), 215-263; 8(1957), 337-374; 8(1957), 375-400.

[41] Prokhorov, Yu.V., Convergence of random processes and limit theorems in probability theory, Theory Prob. Appl., 1(1956), n°2, 157-214.

[42] Prokhorov, Yu.V., Random measures on a compactum, Soviet Math. Dokl., 2(1961),539-541.

[43] Saleh, S., Estimation de la distribution locale moyenne d'une mesure composite aléatoire, C. R. Academie Sci. Paris, 301(1985), Série 1, 10, 545-548.

[44] Suquet, C., *Espaces autoreproduisants et mesures aléatoires*, Thèse de 3ème cycle, Lille, 1986.

[45] Suquet, C., Une topologie pre-Hilbertiènne sur l'espace des mesures à signe bornées, Pub. IRMA, Lille, 13(1988), IV.

[46] Tortrat, A., *Calcul des probabilités et introduction aux processus aléatoires*, Masson, Paris, 1971.

[47] Varadarajan, V.S., Measures on Topological Spaces, Transl. of Ame. Math. Soc., Series 2, 48(1965), 161-228.

[48] Withers, C.S., Central limit theorems for dependent variables I, Z. Wahrsch. verw. Gebiete, 57(1981), n°4, 509-534; Corrigendum to Cental limit theorems for dependent variables I, Z. Wahrsch. verw. Gebiete, 63(1983), 555.