Huayi CHEN

COURS M1 PROBABILITÉS

Huayi CHEN
Université Grenoble Alpes, Institut Fourier (UMR5582), 38402 Saint Martin d'Hères, France.
E-mail : huayi.chen@ujf-grenoble.fr

CHAPITRE 3

CONVERGENCE DE VARIABLES ALÉATOIRES

3.1. Diverses notions de convergence

Dans ce paragraphe, on fixe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

3.1.1. Convergence presque sûre. — Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur $(\Omega, \mathcal{F})$. On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement par rapport à $\mathbb{P}$ s'il existe un ensemble $\mathbb{P}$-negligeable $\Omega_0 \in \mathcal{F}$ tel que la suite numérique $(X_n(\omega))_{n\in\mathbb{N}}$ converge quel que soit $\omega \in \Omega\backslash\Omega_0$. Si X est une variable aléatoire $\mathcal{F}$-mesurable sur Ω telle que la reation

$$\lim_{n\to+\infty} X_n(\omega) = X(\omega)$$

est vérifiée pour tout $\omega \in \Omega$ sauf sur un ensemble negligeable, on dit que X est une limite $\mathbb{P}$-presque sûre de la suite $(X_n)_{n\in\mathbb{N}}$, noté comme

$$\lim_{n\to+\infty} X_n = X \quad \mathbb{P}\text{-p.s.} \quad \text{ou} \quad X_n \xrightarrow{\text{p.s.}} X.$$

La proposition 1.1.8 montre que, si la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement, alors elle admet toujours une limite $\mathbb{P}$-presque sûre qui est $\mathcal{F}$-mesurable. En outre, la limite presque sûre d'une suite est unique à une fonction negligeable près.

Proposition 3.1.1. — *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires $\mathcal{F}$-mesurables. Si*

$$\sum_{n\in\mathbb{N}} \mathbb{E}[|X_n - X_{n+1}| \wedge 1] < +\infty,$$

alors la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement.

Démonstration. — La condition du théorème montre qu'il existe un ensemble negligeable $A \in \mathcal{F}$ tel que

$$\sum_{n\in\mathbb{N}} |X_n(\omega) - X_{n+1}(\omega)| \wedge 1 < +\infty$$

pour tout $\omega \in \Omega \setminus A$. Cette condition montre que la série $\sum_{n\in\mathbb{N}} |X_n(\omega) - X_{n+1}(\omega)|$ est convergente, et en particulier, la suite $(X_n(\omega))_{n\in\mathbb{N}}$ est convergente. □

3.1.2. Convergence en probabilité. — Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur $(\Omega, \mathcal{F})$ et X une variable aléatoire sur $(\Omega, \mathcal{F})$. On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X si, pour tout $\varepsilon > 0$, on a

$$\lim_{n\to+\infty} \mathbb{P}(|X_n - X| \geqslant \varepsilon) = 0,$$

noté comme

$$X_n \xrightarrow{\mathbb{P}} X.$$

Proposition 3.1.2. — *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires qui converge presque sûrement vers une variable aléatoire X. Alors la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X.*

Démonstration. — Quitte à modifier les valeurs des X_n sur un ensemble negligeable, on peut supposer que la suite de fonctions $(X_n)_{n\in\mathbb{N}}$ converge simplement vers X. En particulier, pour tout $\varepsilon > 0$, la suite $(\mathbb{1}_{\{|X_n - X| \geqslant \varepsilon\}})_{n\in\mathbb{N}}$ converge vers la fonction nulle. Le théorème de convergence dominée montre alors que la suite $(\mathbb{P}(|X_n - X| \geqslant \varepsilon))_{n\in\mathbb{N}}$ converge vers 0. La démonstration est donc achevée. □

La réciproque de la proposition précédent n'est cependant pas vrai. On considère l'espace $\Omega = [0, 1]$ muni de la tribue borélienne et la mesure de probabilité uniforme. Pour tout entier $n \in \mathbb{N}$, $n \geqslant 1$ il existe un unique $k(n) \in \mathbb{N}$ telle que $n \in [2^{k(n)}, 2^{k(n)+1}[$. Soit X_n la fonction sur $[0, 1]$ telle que

$$X_n(t) = \begin{cases} 1, & \text{si } n/2^{k(n)} \leqslant t + 1 < (n+1)/2^{k(n)}, \\ 0, & \text{sinon.} \end{cases}$$

Pour tout $\varepsilon > 0$, on a $\mathbb{P}(|X_n| \leqslant \varepsilon) \leqslant 1/n$, donc la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers 0. Cependant, pour tout $t \in [0, 1[$, la suite $(X_n(t))_{n\in\mathbb{N}}$ prend la valeur 1 pour une infinité de fois. Donc la suite $(X_n)_{n\in\mathbb{N}}$ ne converge pas presque sûrement.

Bien que la convergence en probabilité n'est pas équivalente à la convergence presque sûre, la proposition suivante montre que toute suite de Cauchy

pour la convergence en probabilité admet une sous-suite qui converge presque sûrement.

Proposition 3.1.3. — *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur $(\Omega, \mathcal{F})$ telle que,*

$$\forall\, \varepsilon > 0, \quad \lim_{N\to+\infty} \sup_{n,m\geqslant N} \mathbb{P}(|X_n - X_m| \geqslant \varepsilon) = 0,$$

alors il existe une sous-suite de $(X_n)_{n\in\mathbb{N}}$ qui converge presque sûrement. En outre, la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers une variable aléatoire $\mathcal{F}$-mesurable.

Démonstration. — Par récurrence on peut tirer une sous-suite $(X_{n_k})_{k\in\mathbb{N}}$ telle que

$$\forall\, k \in \mathbb{N}, \quad \mathbb{P}(|X_{n_k} - X_{n_{k+1}}| \geqslant 2^{-k}) \leqslant 2^{-k}.$$

Donc

$$\mathbb{E}[|X_{n_k} - X_{n_{k+1}}| \wedge 1] \leqslant 2^{1-k}.$$

D'après la proposition 3.1.1, on obtient que la suite $(X_{n_k})_{k\in\mathbb{N}}$ converge presque sûrement. Soit X une limite presque sûre de la suite $(X_{n_k})_{k\in\mathbb{N}}$. D'après la proposition 3.1.2, on obtient que la suite $(X_{n_k})_{k\in\mathbb{N}}$ converge en probabilité vers X. La condition de la présente proposition montre alors que la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X. □

3.1.3. Convergence L^p. — On a vu dans le chapitre précédent que la convergence presque sûre dominée implique la convergence pour la semi-norme $\|.\|_{\mathscr{L}^p}$. Cette implication n'est pas vraie sans condition de dominance. On peut considérer l'espace $\Omega = [0,1]$ muni de sa tribu borélienne et la mesure de probabilité uniforme. Pour tout $n \in \mathbb{N}$, $n \geqslant 1$ soit $X_n = n\mathbb{1}_{[0,n^{-1}]}$. Alors la suite $(X_n)_{n\geqslant 1}$ converge presque sûrement vers la fonction nulle. Cependant, $\|X_n\|_{\mathscr{L}^p} = n^{(p-1)/p}$ ne converge pas vers 0 lorsque n tend vers l'infini.

Proposition 3.1.4. — *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires qui est une suite de Cauchy dans $\mathscr{L}^p(\Omega, \mathcal{F}, \mathbb{P})$, alors elle converge en probabilité vers une variable aléatoire X. En outre, elle converge aussi vers X par rapport à la semi-norme $\|.\|_{\mathscr{L}^p}$.*

Démonstration. — En effet, pour tout $\varepsilon > 0$ on a

$$\forall\, (n,m) \in \mathbb{N}^2, \quad \mathbb{E}[|X_n - X_m|^p] \geqslant \varepsilon^p \mathbb{P}(|X_n - X_m| \geqslant \varepsilon),$$

qui implique que

$$\lim_{N\to+\infty} \sup_{n,m\geqslant N} \mathbb{P}(|X_n - X_m| \geqslant \varepsilon) = 0.$$

D'après la proposition 3.1.3, on obtient que la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers une variable aléatoire X. En outre, il existe une sous-suite $(X_{n_k})_{k\in\mathbb{N}}$ qui converge presque sûrement vers X. D'après le théorème de Fatou, pour tout $N \in \mathbb{N}$ on a

$$\mathbb{E}[|X_N - X|^p] \leqslant \lim_{k\to+\infty} \mathbb{E}[|X_N - X_{n_k}|^p] \leqslant \sup_{n,m\geqslant N} \mathbb{E}[|X_n - X_m|^p],$$

d'où

$$\lim_{n\to+\infty} \mathbb{E}[|X_n - X|^p] = 0.$$

□

Corollaire 3.1.5. — *L'espace vectoriel $L^p(\Omega, \mathcal{F}, \mathbb{P})$ est compact par rapport à la norme $\|.\|_{L^p}$.*

3.2. Loi de grand nombre

3.2.1. Rappels sur espérace conditionnelle. — Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $\mathcal{A}$ une sous-tribu de $\mathcal{F}$. Si X est une application $\mathcal{F}$-mesurable de Ω vers $[0, +\infty]$, alors

$$A \longmapsto \mathbb{E}[\mathbb{1}_A X]$$

définit une mesure sur la tribu $\mathcal{A}$, qui est absolument continue par rapport à la restriction de $\mathbb{P}$ à $\mathcal{A}$. Le théorème de Radon-Nikodym (théorème 1.2.20) montre qu'il existe une application $\mathcal{A}$-mesurable $\mathbb{E}[X|\mathcal{A}]$ de Ω vers $[0, +\infty]$, unique à une modification près, telle que, pour toute application $\mathcal{A}$-mesurable Y de Ω vers $[0, +\infty]$, on ait

$$\mathbb{E}[Y\mathbb{E}[X|\mathcal{A}]] = \mathbb{E}[XY].$$

Si de plus X est une variable aléatoire intégrable, alors $\mathbb{E}[X|\mathcal{A}]$ est aussi intégrable. Plus généralement, pour toute variable aléatoire positive et $\mathcal{A}$-mesurable Y, la variable aléatoire produit XY est intégrable si et seulement si $\mathbb{E}[X|\mathcal{A}]Y$ l'est. L'unicité presque sûre de l'espérance conditionnelle montre que, pour toutes applications $\mathcal{F}$-mesurables X_1 et X_2 de Ω vers $[0, +\infty]$ et tout couple (a, b) de nombres positifs, on a

$$\mathbb{E}[aX_1 + bX_2|\mathcal{A}] = a\mathbb{E}[X_1|\mathcal{A}] + b\mathbb{E}[X_2|\mathcal{A}] \quad \mathbb{P}\text{-p.s..} \tag{3.1}$$

En particulier, si X et Y sont deux applications $\mathcal{F}$-mesurables de Ω vers $[0, +\infty]$ telles que $X \leqslant Y$ p.s., alors on a $\mathbb{E}[X|\mathcal{A}] \leqslant \mathbb{E}[Y|\mathcal{A}]$ p.s.

La relation (3.1) permet d'étendre la construction d'espérance conditionnelle à $\mathscr{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Soit X une variable aléatoire $\mathcal{F}$-mesurable qui est intégrable. On peut écrire X comme la différence de deux variables aléatoires positives X_1 et X_2 dans $\mathscr{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. On définit alors $\mathbb{E}[X|\mathcal{A}]$ comme la différence $\mathbb{E}[X_1|\mathcal{A}] - \mathbb{E}[X_2|\mathcal{A}]$. D'après la relation (3.1), la classe de $\mathbb{E}[X|\mathcal{A}]$ dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ ne dépend pas du choix de la décomposition $X = X_1 - X_2$. Cette construction induit une application linéaire de $L^1(\Omega, \mathcal{F}, \mathbb{P})$ vers $L^1(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition 3.2.1. — *L'application linéaire de $L^1(\Omega, \mathcal{F}, \mathbb{P})$ vers $L^1(\Omega, \mathcal{A}, \mathbb{P})$ induite par $\mathbb{E}[\,\cdot\,|\mathcal{A}]$ est de norme* 1.

Démonstration. — Pour toute $X \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$, on a $X = \mathbb{E}[X|\mathcal{A}]$ p.s., donc l'opérateur linéaire $L^1(\Omega, \mathcal{F}, \mathbb{P}) \to L^1(\Omega, \mathcal{A}, \mathbb{P})$ induit par l'espérance conditionnelle est un opérateur de projection (dans un sous-espace vectoriel non-nul), sa norme est donc $\geqslant 1$. Il reste à établir $\mathbb{E}[|X|] \geqslant \mathbb{E}\Big[\big|\mathbb{E}[X|\mathcal{A}]\big|\Big]$ pour tout $X \in \mathscr{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Pour simplifier les notations, on désigne par Y la variable aléatoire $\mathbb{E}[X|\mathcal{A}]$. On a

$$|X| \geqslant X(\mathbb{1}_{\{Y \geqslant 0\}} - \mathbb{1}_{\{Y<0\}}).$$

Donc

$$\mathbb{E}[|X|] \geqslant \mathbb{E}[X(\mathbb{1}_{\{Y \geqslant 0\}} - \mathbb{1}_{\{Y<0\}})] = \mathbb{E}[Y(\mathbb{1}_{\{Y \geqslant 0\}} - \mathbb{1}_{\{Y<0\}})] = \mathbb{E}[|Y|]$$

car $\mathbb{1}_{\{Y \geqslant 0\}} - \mathbb{1}_{\{Y<0\}}$ est $\mathcal{A}$-mesurable. La proposition est donc démontrée. □

Proposition 3.2.2. — *Soit X une variable aléatoire $\mathcal{F}$-mesurable à valeur dans $[0, +\infty]$ ou intégrable.*

(1) Si X est indépendante à $\mathcal{A}$, alors $\mathbb{E}[X|\mathcal{A}] = \mathbb{E}[X]$ p.s..

(2) Si $X \geqslant 0$, alors pour toute application $\mathcal{A}$-mesurable de Ω vers $[0, +\infty]$, on a $\mathbb{E}[XY|\mathcal{A}] = \mathbb{E}[X|\mathcal{A}]Y$.

(3) Si X est intégrable, alors pour toute variable aléatoire $\mathcal{A}$-mesurable Y telle que $XY \in \mathscr{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, on a $\mathbb{E}[XY|\mathcal{A}] = \mathbb{E}[X|\mathcal{A}]Y$.

Démonstration. — (1) Sans perte de généralité, on peut supposer que $X \geqslant 0$. Pour tout $A \in \mathcal{A}$, on a $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X]\mathbb{P}(A)$, d'où $\mathbb{E}[X|\mathcal{A}] = \mathbb{E}[X]$ p.s..

(2) Pour tout $A \in \mathcal{A}$ on a

$$\mathbb{E}[XY\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{A}]Y\mathbb{1}_A],$$

d'où le résultat.

(3) La linéairité de l'espérance condintionnelle permet de ramener le problème au cas où X et Y sont positives, c'est-à-dire le cas du point (2).

□

Les propriétés d'espérance se généralisent naturellement au cas d'espérance conditionnelle.

Théorème 3.2.3. — *Soit $(X_n)_{n\in\mathbb{N}}$ est une famille croissante d'applications $\mathcal{F}$-mesurables de Ω vers $[0,+\infty]$, $X = \lim\limits_{n\to+\infty} X_n$, alors*

$$\mathbb{E}[X \mid \mathcal{A}] = \lim_{n\to+\infty} \mathbb{E}[X_n \mid \mathcal{A}] \quad p.s. \tag{3.2}$$

Démonstration. — Pour tout $n \in \mathbb{N}$, on a $X_n \leqslant X_{n+1}$, et donc $\mathbb{E}[X_n|\mathcal{A}] \leqslant \mathbb{E}[X_{n+1}|\mathcal{A}]$ p.s.. Il existe alors un ensemble $\mathbb{P}$-negligeable $A \in \mathcal{A}$ tel que la suite $(\mathbb{E}[X_n|\mathcal{A}](\omega))_{n\in\mathbb{N}}$ soit croissante (et convergeante dans $[0,+\infty]$) pour tout $\omega \in \Omega \setminus A$. La suite $(\mathbb{E}[X_n|\mathcal{A}])_{n\in\mathbb{N}}$ converge donc presque sûrement vers une variable aléatoire $\mathcal{A}$-mesurable à valeurs dans $[0,+\infty]$ que l'on note comme Z. Pour toute application $\mathcal{A}$-mesurable $Y : \Omega \to [0,+\infty]$, d'après le théorème de Beppo Levi, on a

$$\mathbb{E}[YZ] = \lim_{n\to+\infty} \mathbb{E}[Y\mathbb{E}[X_n|\mathcal{A}]] = \lim_{n\to+\infty} \mathbb{E}[YX_n] = \mathbb{E}[YX].$$

Comme Y est arbitraire, on obtient $Z = \mathbb{E}[X|\mathcal{A}]$ p.s.. □

Théorème 3.2.4. — *Soit $(X_n)_{n\geqslant 1}$ une suite d'applications $\mathcal{F}$-mesurables de Ω vers $[0,+\infty]$. On a*

$$\liminf_{n\to+\infty} \mathbb{E}[X_n \mid \mathcal{A}] \geqslant \mathbb{E}[\liminf_{n\to+\infty} X_n \mid \mathcal{A}] \qquad p.s.$$

Démonstration. — Pour tout entier $n \in \mathbb{N}$, soit $Y_n = \inf_{m\geqslant n} X_m$. On a $Y_n \leqslant X_n$. En outre, la suite $(Y_n)_{n\in\mathbb{N}}$ est croissante et converge vers $\liminf\limits_{n\to+\infty} X_n$. D'après le théorème 3.2.3, on a

$$\mathbb{E}[\liminf_{n\to+\infty} X_n] = \lim_{n\to+\infty} \mathbb{E}[Y_n|\mathcal{A}] \leqslant \liminf_{n\to+\infty} \mathbb{E}[X_n|\mathcal{A}] \quad \text{p.s..}$$

□

Théorème 3.2.5. — *Soient $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires $\mathcal{F}$-mesurables et $Y \in \mathscr{L}^1(\Omega,\mathcal{F},\mathbb{P})$ telles que $|X_n| \leqslant Y$ quel que soit n. Si X_n converge presque sûrement vers une variable aléatoire X, alors la suite d'espérances conditionnelle $(\mathbb{E}[X_n \mid \mathcal{A}])_{n\in\mathbb{N}}$ converge vers $\mathbb{E}[X \mid \mathcal{A}]$ dans $\mathscr{L}^1(\Omega,\mathcal{A},\mathbb{P})$ et presque sûrment.*

Démonstration. — D'après le théorème de convergence dominée, on obtient que la suite $(X_n)_{n\in\mathbb{N}}$ converge dans $\mathscr{L}^1(\Omega,\mathcal{F},\mathbb{P})$ vers X, d'où la suite $(\mathbb{E}[X_n|\mathcal{A}])_{n\in\mathbb{N}}$ converge dans $\mathscr{L}^1(\Omega,\mathcal{A},\mathbb{P})$ vers $\mathbb{E}[X|\mathcal{A}]$, compte tenu de la

proposition 3.4.5. Enfin, le théorème 3.2.4 appliqué aux suites $(Y - X_n)_{n\in\mathbb{N}}$ et $(Y + X_n)_{n\in\mathbb{N}}$ montre que

$$\liminf_{n\to+\infty} \mathbb{E}[X_n|\mathcal{A}] \geqslant \mathbb{E}[X|\mathcal{A}] \geqslant \limsup_{n\to+\infty} \mathbb{E}[X_n|\mathcal{A}] \quad \text{p.s.},$$

d'où la convergence presque sûre. □

3.2.2. Filtration et martingales. —

Définition 3.2.6. — On appelle *filtration* toute suite $\mathbb{F} = (\mathcal{F}_n)_{n\geqslant 0}$ de sous-tribus de $\mathcal{F}$ telle que $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ quel que soit n. Étant donnée une filtration $\mathbb{F}$, on dit qu'une suite $M = (M_n)_{n\in\mathbb{N}}$ de variables aléatoires est une $\mathbb{F}$-*martingale* si, pour chaque n, M_n est une variable aléatoire $\mathcal{F}_n$-mesurable et intégrable, et si la relation $\mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n] = M_n$ p.s. est vérifiée pour tout $n \in \mathbb{N}$. On dit qu'une suite $(A_n)_{n\in\mathbb{N}}$ de variables aléatoires est une $\mathbb{F}$-*sous-martingale* si chaque A_n est dans $\mathscr{L}^1(\Omega, \mathcal{F}_n, \mathbb{P})$ et si $\mathbb{E}[A_{n+1} \,|\, \mathcal{F}_n] \geqslant A_n$ p.s. pour tout $n \in \mathbb{N}$.

Soit $(Y_n)_{n\geqslant 1}$ une famille de variables aléatoires indépendantes d'espérance nulle. Pour tout $n \geqslant 1$, soit $\mathcal{F}_n$ la sous-tribu de $\mathcal{F}$ engendrée par $Y_1, \cdots, Y_n$, et soit $M_n = Y_1 + \cdots + Y_n$. On pose $\mathcal{F}_0 = \{\varnothing, \Omega\}$ et $M_0 = 0$, alors M est une $\mathbb{F}$-martingale.

Lemme 3.2.7. — *Soient X une variable aléatoire $\mathcal{F}$-mesurable intégrable et $\mathcal{A}$ une sous-tribu de $\mathcal{F}$. Si f est une fonction convexe telle que $f(X)$ soit intégrable, alors*

$$\mathbb{E}[f(X)|\mathcal{A}] \geqslant f\big(\mathbb{E}[X|\mathcal{A}]\big).$$

Démonstration. — Soit $(x_n)_{n\in\mathbb{N}}$ une suite dense dans $\mathbb{R}$. Pour tout n, il existe une fonction affine $f_n(x) = a_n x + b_n$ telle que $f_n \leqslant f$ et que $f_n(x_n) = f(x_n)$. On en déduit $f = \sup_n f_n$ car les deux fonctions sont convexes (donc continues) et d'accord sur un sous-ensemble dense de $\mathbb{R}$. Par conséquent,

$$\mathbb{E}[f(X) \,|\, \mathcal{A}] \geqslant \sup_n \mathbb{E}[f_n(X) \,|\, \mathcal{A}] = \sup_n f_n(\mathbb{E}[X \,|\, \mathcal{A}]) = f(\mathbb{E}[X \,|\, \mathcal{A}]) \quad \text{p.s..}$$

□

On en déduit la proposition suivante :

Proposition 3.2.8. — *Soit $\mathbb{F}$ une filtration de la tribu $\mathcal{F}$. Si $(M_n)_{n\geqslant 1}$ est une $\mathbb{F}$-matingale et si f est une fonction convexe telle que $f(M_n)$ soit intégrable pour tout n, alors $(f(M_n))_{n\geqslant 1}$ est une $\mathbb{F}$-sous-martingale. En particulier, $(|M_n|)_{n\geqslant 1}$ est une $\mathbb{F}$-sous-martingale.*

Démonstration. — D'après le lemme 3.2.7, on a

$$\mathbb{E}[f(M_{n+1}) \,|\, \mathcal{F}_n] \geqslant f(\mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n]) = f(M_n) \quad \text{p.s..}$$

□

Proposition 3.2.9. — *Soit $\mathbb{F}$ une filtration et $A = (A_n)_{n\in\mathbb{N}}$ une $\mathbb{F}$-sous-martigale. Si $m < n$ sont deux entiers dans $\mathbb{N}$, on a, pour tout $U \in \mathcal{F}_m$,*

$$\mathbb{E}[\mathbb{1}_U A_n] \geqslant \mathbb{E}[\mathbb{1}_U A_m].$$

Démonstration. — On a

$$\mathbb{E}[\mathbb{1}_U A_n] = \mathbb{E}[\mathbb{1}_U \mathbb{E}[A_n \,|\, \mathcal{F}_m]] \geqslant \mathbb{E}[\mathbb{1}_U A_m].$$

□

La proposition suivant est important dans la théorie des martingale.

***Théorème 3.2.10* (Doob).** — *Soient $\mathbb{F}$ une filtration et $M = (M_n)_{n\in\mathbb{N}}$ une $\mathbb{F}$-martingale telle que $M_n \in \mathscr{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ pour tout n. Alors, pour tout entier $n \in \mathbb{N}$, on a*

$$\mathbb{E}[\max_{k\in\{0,\dots,n\}} |M_k|^2] \leqslant 4\mathbb{E}[|M_n|^2] \tag{3.3}$$

En outre,

$$\mathbb{E}[\sup_{n\in\mathbb{N}} |M_n|^2] \leqslant 4 \sup_{n\in\mathbb{N}} \mathbb{E}[|M_n|^2]. \tag{3.4}$$

Démonstration. — Pour tout $n \in \mathbb{N}$, soit $Y_n = \max_{k\in\{0,\dots,n\}} |M_k|$. Sans perte de généralité, on suppose que Y_n n'est pas p.s. nulle. Pour tout a, soit

$$\tau_a = \inf\{k \geqslant 0 \,:\, |M_k| \geqslant a\}.$$

On a $\{Y_n \geqslant a\} = \{\tau_a \leqslant n\} \in \mathcal{F}_n$. En outre

$$\begin{aligned} a\mathbb{P}(Y_n \geqslant a) = a\mathbb{P}(\tau_a \leqslant n) &\leqslant \mathbb{E}[|M_{\tau_a}|\mathbb{1}_{\{\tau_a\leqslant n\}}] = \sum_{k=1}^{n} \mathbb{E}[\mathbb{1}_{\{\tau_a=k\}}|M_k|] \\ &\leqslant \sum_{k=1}^{n} \mathbb{E}[\mathbb{1}_{\{\tau_a=k\}}|M_n|] = \mathbb{E}[\mathbb{1}_{\{\tau_a\leqslant n\}}|M_n|] = \mathbb{E}[\mathbb{1}_{\{Y_n\geqslant a\}}|M_n|], \end{aligned}$$

où dans la dernière inégalité, on a utilisé la proposition 3.2.9 et le fait que

$$\{\tau_a = k\} = \{\tau_a \leqslant k\} \setminus \{\tau_a \leqslant k-1\} \in \mathcal{F}_k.$$

En prenant l'intégrale par rapport à $a \in [0, +\infty[$, on obtient

$$\int_0^{+\infty} a\mathbb{P}(Y_n \geqslant a)\,\mathrm{d}a \leqslant \mathbb{E}\big[Y_n \cdot |M_n|\big] \leqslant \mathbb{E}[Y_n^2]^{\frac{1}{2}}\mathbb{E}[|M_n|^2]^{\frac{1}{2}},$$

la dernière inégalité provient de l'inégalité de Cauchy-Schwarz. Mais l'intégrale à gauche n'est rien d'autre que $\frac{1}{2}\mathbb{E}[Y_n^2]$. D'où l'inégalité (3.3). Enfin, par le théorème de Beppo Levi, le terme à gauche de (3.3) converge vers celui de (3.4) lorsque $n \to +\infty$, donc ce dernier est majoré par la borne supérieure du terme à droite de (3.3). □

Corollaire 3.2.11. — *Soit M une $\mathbb{F}$-martingale telle que $\sup_n \mathbb{E}[M_n^2] < +\infty$, alors la suite $(M_n)_{n\geqslant 1}$ converge dans $\mathscr{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ et p.s. vers une variable aléatoire M_∞.*

Démonstration. — Soit $K = \sup_n \mathbb{E}[M_n^2]$. Comme la fonction $x \mapsto x^2$ est convexe, d'après l'inégalité de Jensen, on a

$$\mathbb{E}[M_{n+1}^2 \mid \mathcal{F}_n] \geqslant \mathbb{E}[M_{n+1} \mid \mathcal{F}_n]^2 = M_n^2.$$

En prenant l'espérance, on obtient que la suite $(\mathbb{E}[M_n^2])_{n\geqslant 1}$ est croissante. On a donc [1]

$$\mathbb{E}[(M_{n+k} - M_n)^2] = \mathbb{E}[M_{n+k}^2] - \mathbb{E}[M_n^2] \leqslant K - \mathbb{E}[M_n^2] \to 0 \quad (n \to +\infty).$$

Cela montre que $(M_n)_{n\geqslant 1}$ est une suite de Cauchy dans $\mathscr{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, et donc converge vers un certain $M_\infty \in \mathscr{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ car l'espace $\mathscr{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ est complet.

Il reste à établir que $(M_n)_{n\geqslant 0}$ converge p.s., ou de façon équivalente, est p.s. une suite de Cauchy. Pour cela, on considère des majorations de $\mathbb{P}(\sup_{i,j\geqslant n} |M_i - M_j| > \varepsilon)$ pour tous ε et n. On a

$$\mathbb{P}(\sup_{i,j\geqslant n} |M_i - M_j| > \varepsilon) \leqslant \mathbb{P}(\sup_{i\geqslant n} |M_i - M_n| > \varepsilon/2) \leqslant \frac{4}{\varepsilon^2}\mathbb{E}[\sup_{i\geqslant n} |M_i - M_n|^2].$$

Mais $(M_i - M_n)_{i\geqslant n}$ est une martingale à partir d'indice n. Donc l'inégalité de Doob (3.4) implique

$$\mathbb{E}[\sup_{i\geqslant n} |M_i - M_n|^2] \leqslant 4 \sup_{i\geqslant n} \mathbb{E}[|M_i - M_n|^2]$$

converge vers 0 lorsque n tend vers l'infini. Donc $(M_n)_{n\geqslant 1}$ est p.s. une suite de Cauchy. □

Corollaire 3.2.12. — *Soient $(X_n)_{n\geqslant 1}$ des variables aléatoires indépendantes carré intégrables, et $S_n = \sum_{k=1}^n X_k$. Si*

$$b := \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i]$$

1. $\mathbb{E}[M_{n+k}M_n] = \mathbb{E}[M_n\mathbb{E}[M_{n+k} \mid \mathcal{F}_n]] = \mathbb{E}[M_n^2]$.

existe dans $\mathbb{R}$ *et si la série* $\sum_{n\geqslant 1}\mathrm{var}(X_n)/n^2$ *converge, alors* S_n/n *converge presque sûrement vers* b.

Démonstration. — Sans perte de généralité on peut supposer que $\mathbb{E}[X_n]=0$ pour tout n. Soient

$$M_n=\frac{X_1}{1}+\frac{X_2}{2}+\cdots+\frac{X_n}{n},\quad (n\in\mathbb{N},\ n\geqslant 1),$$

et $M_0=0$. C'est une martingale par rapport à la filtration $\mathbb{F}=(\mathcal{F}_n)_{n\geqslant 0}$ telle que $\mathcal{F}_0=\{\varnothing,\Omega\}$ et $\mathcal{F}_n=\sigma(X_1,\cdots,X_n)$. Comme

$$\mathbb{E}[M_n^2]=\sum_{k=1}^{n}\frac{1}{k^2}\mathrm{var}(X_k)$$

est uniformément borné, le corollaire 3.2.11 montre que la suite $(M_n)_{n\in\mathbb{N}}$ converge p.s. vers une certaine variable aléatoire $M_\infty\in\mathscr{L}^2(\Omega,\mathcal{F},\mathbb{P})$. Par conséquent, [2]

$$\frac{S_n}{n}=\frac{1}{n}\sum_{k=1}^{n}k(M_n-M_{n-1})=M_n-\frac{M_1+\cdots+M_{n-1}}{n}\to M_\infty-M_\infty=0\quad p.s.$$

□

3.2.3. Loi de grand nombre. —

***Théorème 3.2.13* (Kolmogorov).** — *Soit* $(X_n)_{n\geqslant 1}$ *une suite de variables aléatoires mutuellement indépendantes et qui suivent la même loi (c'est-à-dire que les mesures boréliennes* μ_{X_n} *sont la même). Pour tout* $n\geqslant 1$*, soit* $S_n=X_1+\cdots+X_n$*. On suppose que* X_1 *admette une espérance* $b\in\mathbb{R}$*, alors la suite de variable aléatoires* $(S_n/n)_{n\geqslant 1}$ *converge presque sûrement vers* b.

2. Si $(x_n)_{n\geqslant 1}$ est une suite convergente dans $\mathbb{R}$ dont la limite est x, alors on a

$$\lim_{n\to+\infty}\frac{x_1+\cdots+x_n}{n}=x.$$

En effet, la suite $|x_n-x|$ est bornée par une constante α. Fixons $m\in\mathbb{N}$. Pour $n>m$, on a

$$\begin{aligned}\Big|\frac{x_1+\cdots+x_n}{n}-x\Big|&\leqslant\frac{|x_1-x|+\cdots+|x_m-x|}{n}+\frac{|x_{m+1}-x|+\cdots+|x_n-x|}{n}\\&\leqslant\alpha\frac{m}{n}+\sup_{k\geqslant m}|x_k-x|.\end{aligned}$$

En prenant limsup, on obtient

$$\limsup_{n\to+\infty}\Big|\frac{x_1+\cdots+x_n}{n}-x\Big|\leqslant\sup_{k\geqslant m}|x_k-x|.$$

Démonstration. — Sans perte de généralité, on peut supposer $\mathbb{E}[X_n] = 0$. Posons $Y_n = X_n \mathbb{1}_{\{|X_n| \leqslant n\}}$. D'après le théorème de convergence dominée, on obtient

$$\mathbb{E}[Y_n] = \mathbb{E}[X_n \mathbb{1}_{\{|X_n| \leqslant n\}}] = \mathbb{E}[X_1 \mathbb{1}_{\{|X_1| \leqslant n\}}] \to 0.$$

D'où $\frac{1}{n}(\mathbb{E}[Y_1] + \cdots + \mathbb{E}[Y_n]) = 0$. En outre, on a

$$\begin{aligned}\sum_{n \geqslant 1} \frac{1}{n^2} \mathbb{E}[Y_n^2] &= \sum_{n \geqslant 1} \frac{1}{n^2} \mathbb{E}[X_1^2 \mathbb{1}_{\{|X_1| \leqslant n\}}] = \sum_{n \geqslant 1} \frac{1}{n^2} \sum_{k=1}^{n} \mathbb{E}[X_1^2 \mathbb{1}_{\{k-1 < |X_1| \leqslant k\}}] \\ &= \sum_{k \geqslant 1} E[X_1^2 \mathbb{1}_{\{k-1 < |X_1| \leqslant k\}}] \sum_{n \geqslant k} \frac{1}{n^2} \\ &\leqslant \sum_{k \geqslant 1} \mathbb{E}[X_1^2 \mathbb{1}_{\{k-1 < |X_1| \leqslant k\}}](1/k^2 + 1/k),\end{aligned}$$

où dans la dernière inégalité, on utilise l'estimation

$$\sum_{n \geqslant k} \frac{1}{n^2} \leqslant \frac{1}{k^2} + \frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)} + \cdots = \frac{1}{k^2} + \frac{1}{k}.$$

Par conséquent, on a

$$\sum_{n \geqslant 1} \frac{1}{n^2} \mathbb{E}[Y_n^2] \leqslant \sum_{k \geqslant 1} \mathbb{E}[|X_1| \mathbb{1}_{\{k-1 < |X_1| \leqslant k\}}](1/k + 1) \leqslant 2\mathbb{E}[|X_1|] < +\infty.$$

D'après le corollaire 3.2.12, on a

$$\lim_{n \to +\infty} \frac{Y_1 + \cdots + Y_n}{n} = 0 \qquad \text{p.s.}$$

Enfin,

$$\begin{aligned}\sum_{n \geqslant 1} \mathbb{P}(X_n \neq Y_n) &= \sum_{n \geqslant 1} \mathbb{P}(|X_n| > n) \\ = \sum_{n \geqslant 1} \mathbb{P}(|X_1| > n) &= \mathbb{E}[\sum_{n < |X_1|} 1] \leqslant 1 + \mathbb{E}[|X_1|].\end{aligned}$$

Cela montre que presque sûrement les deux suites $(X_n)_{n \geqslant 1}$ et $(Y_n)_{n \geqslant 1}$ ne diffère qu'un nombre fini de termes, d'où le théorème. □

3.3. Mesures sur un espace localement compact

3.3.1. Mesures de Radon. — Soit E un espace topologique. On désigne par $C_c(E)$ l'ensemble des fonctions continues sur E qui sont de support compact. On dit que E est *localement compact* si tout point de E admet un voisinage qui est compact.

Proposition 3.3.1. — *Soit E un espace topologique.*

(1) L'intersection d'un sous-ensemble compact et un sous-ensemble fermé de E est compact.

(2) On suppose que E est séparé. Si C est un sous-ensemble compact de E et si $y \in E \setminus C$, alors il existe deux ouverts U et V de E tels que $C \subset U$, $y \in V$ et $U \cap V = \varnothing$.

(3) Si E est séparé, alors tout sous-ensemble compact de E est fermé.

Démonstration. — (1) Soient F un fermé de E et $C \subset E$ un sous-ensemble compact. Soit $(U_i)_{i\in I}$ une famille d'ouverts de E qui recouvre $C \cap F$. Alors $(U_i)_{i\in I} \cup \{F^c\}$ est un recouvrement de C, qui admet donc un sous-recouvrement fini. On en déduit l'eixstence d'un sous-ensemble fini J de I tel que $C \subset \bigcup_{j\in J} U_j \cup F^c$. On en déduit donc $C \cap F \subset \bigcup_{j\in J} U_j$.

(2) Pour tout $x \in C$, il existe deux ouverts U_x et V_x tels que $x \in U_x$, $y \in V_x$ et $U_x \cap V_x = \varnothing$. L'ensemble C est recouvert par $(U_x)_{x\in C}$, il existe alors un sous-ensemble fini $I \subset C$ tel que $C \subset U := \bigcup_{x\in I} U_x$. Soit $V = \bigcap_{x\in I} V_x$. Les sous-ensembles U et V sont ouverts, et on a $U \cap V = \varnothing$.

(3) Soit C un sous-ensemble compact de E. D'après (2), tout point $y \in E \setminus C$ admet un voisinage qui ne rencontre pas C. L'ensemble C est donc fermé. □

Lemme 3.3.2. — *Soit E un espace topologique séparé et compact.*

(1) Si U un ouvert de E, alors pour tout $x \in U$, il existe un voisinage ouvert V de x tel que $\overline{V} \subset U$.

(2) L'espace topologique E est normal, c'est-à-dire que, pour tout couple (F_1, F_2) de fermés de E tels que $F_1 \cap F_2 = \varnothing$, il existe des ouverts V_1 et V_2 de E tels que $V_1 \supset F_1$, $V_2 \supset F_2$ et $V_1 \cap V_2 = \varnothing$.

Démonstration. — (1) Comme E est compact, tout sous-ensemble fermé de E l'est aussi. En particulier, $E \setminus U$ est compact. D'après la proposition 3.3.1.(2), il existe deux ouverts V et W de E tels que $x \in V$, $E \setminus U \subset W$ et $V \cap W = \varnothing$. En particulier, on a $\overline{V} \subset E \setminus W \subset U$.

(2) D'après la proposition 3.3.1.(2), pour tout $x \in F_2$, il existe des ouverts U_x et V_x tels que $x \in U_x$, $F_1 \subset V_x$ et $U_x \cap V_x = \varnothing$. Comme F_2 est compact, il existe un sous-ensemble fini I de F_2 tel que

$$F_2 \subset V_2 := \bigcup_{x\in I} U_x.$$

Soit $V_1 = \bigcap_{x\in I} V_x$. Les sous-ensembles V_1 et V_2 sont tous ouvert, et on a $F_1 \subset V_1$, $F_2 \subset V_2$ et $V_1 \cap V_2 = \varnothing$. □

***Lemme 3.3.3* (Uryson).** — *Soit E un espace topologique normal. Si F_0 et F_1 sont deux parties fermés de E telles que $F_0 \cap F_1 = \varnothing$, alors il existe une application continue $\varphi : E \to [0,1]$ telle que $\varphi(x) = 0$ si $x \in F_0$ et $\varphi(x) = 1$ si $x \in F_1$.*

Démonstration. — On écrit l'ensemble $\mathbb{Q} \cap [0,1]$ sous forme d'une suite $(a_n)_{n\in\mathbb{N}}$ de sorte que $a_0 = 0$ et $a_1 = 1$. On construit par récurrence une famille $(E_n)_{n\geqslant 0}$ de fermés de E contenus dans $E \setminus F_0$ qui vérifie la condition suivante : si n et m sont deux indices distincts dans $\mathbb{N}_{\geqslant 1}$ tels que $a_n < a_m$, alors on a $E_n^\circ \supset E_m$. On pose $E_1 = F_1$ et on choisit un ouvert $U_0 \supset F_1$ tel que $\overline{U}_0 \cap F_0 = \varnothing$. On note $E_0 = \overline{U}_0$.

Supposons que les fermés $E_0, \ldots, E_n$ sont construits, qui satisfont à la propriété annoncé. Il existe deux indices district i et j dans $\{0, \ldots, n\}$ tels que $a_i < a_{n+1} < a_j$ et que $]a_i, a_j[\cap \{a_0, \ldots, a_n\} = \varnothing$. Comme $E_i^\circ \supset E_j$, il existe un ouvert U de E tel que $E_j \subset U \subset \overline{U} \subset E_i^\circ$. On pose $E_{n+1} = \overline{U}$. La suite de fermés $(E_n)_{n\geqslant 0}$ ainsi construite par récurrence vérifie la propriété annocée.

Pour tout élément $x \in E$, soit

$$\varphi(x) = \sup_{n\in\mathbb{N}} \{a_n \,:\, x \in E_n\},$$

où par convention $\sup \varnothing = 0$. Par définition,

$$\{x \in E \,:\, \varphi(x) > t\} = \bigcup_{n\in\mathbb{N},\, a_n > t} E_n = \bigcup_{n\in\mathbb{N},\, a_n > t} E_n^\circ$$

car pour tout $n \in \mathbb{N}$ tel que $a_n > t$, il existe un $m \in \mathbb{N}$ tel que $a_n > a_m > t$ et on a $E_n \subset E_m^\circ$. C'est donc un ouvert de E. Montrons que

$$\{x \in E \,:\, \varphi(x) \geqslant t\} = \bigcap_{n\in\mathbb{N},\, a_n < t} E_n.$$

Si $x \in \bigcap_{a_n<t} E_n$, alors $\{a_n \,:\, x \in E_n\} \supset \mathbb{Q} \cap [0,t[$, donc on a $\varphi(x) \geqslant t$. Réciproquement, si $x \in E$ est tel que $\varphi(x) \geqslant t$, alors pour tout $\varepsilon > 0$ il existe $m(\varepsilon) \in \mathbb{N}$ tel que $a_{m(\varepsilon)} > t - \varepsilon$ et que $x \in E_{m(\varepsilon)}$. Si $n \in \mathbb{N}$ est tel que $a_n < t$, il existe alors $\varepsilon > 0$ tel que $a_n < t - \varepsilon < a_{m(\varepsilon)}$ et donc $x \in E_{m(\varepsilon)} \subset E_n$. Cela montre que φ est une fonction continue. Enfin, par construction on a $\varphi|_{F_0} = 0$ et $\varphi|_{F_1} = 1$. □

Remarque 3.3.4. — Soit E un espace topologique. On peut toujours plonger E dans un espace topologique compact de manière suivante. Soit $\widetilde{E}$ la réunion disjointe de E avec l'ensemble d'un point que l'on note comme ∞. Un sous-ensemble U de $\widetilde{E}$ est ouvert si et seulement s'il est un ouvert de E ou s'il s'écrit sous la forme $(E \setminus F) \cup \{\infty\}$, où F est un sous-ensemble compact de E. On

vérifie facilement que $\widetilde{E}$ est compact. Si E est séparé et localement compact, alors l'espace topologique $\widetilde{E}$ est séparé. On en déduit que le premier point du lemme 3.3.2 est encore valable pour un espace topologique localement compact.

Soient E un espace topologique séparé, F un sous-ensemble compact de E et U un sous-ensemble ouvert de E qui contient F. Alors $\widetilde{E} \setminus U$ est sous-ensemble fermé de $\widetilde{E}$. De plus, comme F est compact, il est fermé dans $\widetilde{E}$. Si on applique le lemme d'Uryson à F et $\widetilde{E} \setminus U$, on obtient l'existence d'une fonction continue φ sur E à valeurs dans $[0,1]$ qui est identiquement nulle sur $E \setminus U$ et est telle que $\varphi(x) = 1$ pour tout $x \in F$.

Proposition 3.3.5. — *Soit E un espace séparé et localement compact, qui admet une base de topologie dénombrable. Alors la tribu engendrée par $C_c(E)$ s'identifie à la tribu borélienne de E.*

Démonstration. — Soient U un ouvert de E et x est un point de U. Comme E est localement compact, il existe un voisiange ouvert V_x de x tel que $\overline{V}_x$ soit compact [3]. D'après le lemme 3.3.2 (voir aussi la remarque 3.3.4), il existe un voisinage ouvert W_x de x tel que

$$x \in W_x \subset \overline{W_x} \subset U \cap V_x.$$

En particulier, $\overline{W_x}$ est compact. On choisit une fonction continue φ_x sur E qui est identiquement nulle sur $E \setminus (U \cap V_x)$ et telle que $\varphi_x(y) = 1$ si $y \in \overline{W}_x$. La fonction φ_x appartient à $C_c(E)$ et $W'_x := \varphi_x^{-1}(\mathbb{R} \setminus \{0\}) \subset U$. L'ensemble U est recouvert par les ouverts $(W'_x)_{x \in U}$. Il existe donc un sous-ensemble dénombrable I de U tel que $U = \bigcup_{x \in I} W'_x$. En particulier, on a $U \in \sigma(C_c(E))$. Le résultat est donc démontré. □

***Lemme 3.3.6* (Dini)**. — *Soit E un espace topologique compact. Si $(f_n)_{n \in \mathbb{N}}$ est une suite décroissante de fonctions continues sur E qui converge simplement vers 0, alors la convergence est aussi uniforme.*

Démonstration. — Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, soit $F_n := \{x \in E \,:\, f_n(x) \geqslant \varepsilon\}$. Comme la suite $(f_n)_{n \in \mathbb{N}}$ converge simplement vers la fonction nulle, on a $\bigcap_{n \in \mathbb{N}} F_n = \varnothing$, ou de façon équivalent, $E = \bigcup_{n \in \mathbb{N}} F_n^c$. Comme E est compact, il existe alors $n \in \mathbb{N}$ tel que

$$E = \bigcup_{k \in \{0,\ldots,n\}} F_k^c = F_n^c,$$

3. Le point x admet un voisinage compact F, qui est un sous-ensemble fermé de E puisque E est séparé. On peut prendre V_x comme l'intérieur de F.

d'où $F_n = \varnothing$. La suite $(f_n)_{n\in\mathbb{N}}$ converge donc uniformément. □

***Théorème 3.3.7* (Représentation de Riesz).** — *Soit E un espace topologique séparé et localement compact, qui adment une base de topologie dénombrable. Alors toute application linéaire positive*[4] *$I : C_c(E) \to \mathbb{R}$ définit une unique mesure borélienne μ sur E telle que $\mu(F)$ soit fini pour tout sous-ensemble compact F de E.*

Démonstration. — Il suffit de montrer que toute application linéaire positive $I : C_c(E) \to \mathbb{R}$ définit un opérateur d'intégration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite décroissante de fonctions positives qui converge simplement vers 0. Soit F le support de la fonction f_0, qui est un sous-ensemble compact de E. Pour tout $x \in F$, soit C_x un voisinage compact de x dans E. Comme $(C_x^\circ)_{x\in F}$ est un recouvrement de F, il existe un sous-ensemble fini $I \subset F$ tel que $F \subset \bigcup_{x\in I} C_x^\circ$. Soit $F' := \bigcup_{x\in I} C_x$. C'est un sous-ensemble compact de E tel que $F \subset F'^\circ$. D'après la remarque (3.3.4), il existe une fonction continue φ sur E, qui s'annule sur $E \setminus F'^\circ$ et telle que $\varphi(x) = 1$ dès que $x \in F$. C'est donc une fonction dans $C_c(E)$. En outre, comme les supports des fonctions f_n sont dans F, on obtient que la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément (cf. le lemme 3.3.6). Il existe donc une suite $(\varepsilon_n)_{n\in\mathbb{N}}$ qui tend vers 0 lorsque $n \to +\infty$ telle que $f_n \leqslant \varepsilon_n\varphi$ pour tout $n \in \mathbb{N}$, d'où $I(f) \leqslant \varepsilon_n I(\varphi)$ tend vers 0 lorsque $n \to +\infty$.

Comme E est localement compact et à base dénombrable, il existe donc une suite $(F_n)_{n\in\mathbb{N}}$ de sous-ensembles compacts de E tel que $\bigcup_{n\in\mathbb{N}} F_n = E$. On en déduit que la fonction constante 1 appartient à $C_c(E)^\uparrow$. D'après la proposition 1.2.24, on obtient que l'application linéaire I définit en fait une unique mesure σ-finie sur $\sigma(C_c(E)) = \mathcal{B}(E)$, compte tenu de la proposition 3.3.5. □

Définition 3.3.8. — Soit E un espace topologique séparé et localement compact, qui admet une base de topologie dénombrable. On appelle *mesure de Radon* sur E toute mesure borélienne sur E qui prend valeurs finies sur tout compact de E. Le théorème 3.3.7 montre que l'ensemble des mesures de Radon sur E est en correspondance biunivoque avec l'ensemble des formes linéaires positives sur $C_c(E)$.

3.4. Convergence de mesures

Dans ce paragraphe, on fixe un espace topologique E, qui est supposé être séparé, localement compact, et admettre une base de topologie dénombrable.

4. C'est-à-dire que $I(f) \geqslant 0$ pour toute fonction positive dans $C_c(E)$.

Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de Radon sur E, qui correspond à une suite $(I_n)_{n\in\mathbb{N}}$ de formes linéaires positives sur $C_c(E)$. On dit que la suite de mesures $(\mu_n)_{n\in\mathbb{N}}$ *converge vaguement* si la suite de formes linéaires positives $(I_n)_{n\in\mathbb{N}}$ converge simplement. Sous l'hypothèse de convergence vague pour la suite de mesures $(\mu_n)_{n\in\mathbb{N}}$, la limite de la suite de formes linéaires $(I_n)_{n\in\mathbb{N}}$ est une forme linéaire positive sur $C_c(E)$, qui détermine une mesure de Radon sur E, que l'on appelle la *limite vague* de la suite de mesure de Radon $(\mu_n)_{n\in\mathbb{N}}$.

Remarque 3.4.1. — Soit E un espace topologique séparé et localement compact, qui admet une base de topologie dénombrable. Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilité de Radon sur E, qui converge vaguement vers une mesure de Radon μ. Il est pas vrai en générale que μ est une mesure de probabilité. En effet, on peut considérer le cas où $E = \mathbb{R}$ et $\mu_n = \delta_n$ pour tout $n \in \mathbb{N}$. Pour toute fonction continue φ sur $\mathbb{R}$ à support compact, on a

$$\lim_{n\to+\infty} \int_{\mathbb{R}} \varphi(x)\, \delta_n(\mathrm{d}x) = 0.$$

Donc la suite $(\delta_n)_{n\in\mathbb{N}}$ converge vaguement vers la mesure nulle.

Définition 3.4.2. — Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesure de Radon sur E. On suppose que chaque mesure μ_n est finie. On dit que la suite $(\mu_n)_{n\in\mathbb{N}}$ *converge étroitement* vers une mesure de Radon μ si la relation

$$\lim_{n\to+\infty} \int_E h(x)\, \mu_n(\mathrm{d}x) = \int_E h(x)\, \mu(\mathrm{d}x)$$

est vérifiée pour toute fonction continue et bornée h sur E.

Proposition 3.4.3. — *Soient $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de Radon finies sur E, et μ une mesure de Radon finie sur E. Alors la suite $(\mu_n)_{n\in\mathbb{N}}$ converge étroitement vers μ si et seulement si elle converge vaguement vers μ et vérifie*

$$\lim_{n\to+\infty} \mu_n(E) = \mu(E).$$

Démonstration. — Comme E est un espace localement compact qui admet une base de topologie dénombrable, il existe une suite d'ouverts $(U_k)_{k\in\mathbb{N}}$ tels que $\overline{U}_k$ soit compact et continu dans U_{k+1} pour tout $k \in \mathbb{N}$ et que

$$\bigcup_{k\in\mathbb{N}} U_k = E.$$

Pour tout $k \in \mathbb{N}$, soit $\varphi_k : E \to [0,1]$ une fonction continue telle que $\varphi_k(x) = 0$ si $x \in U_{k+1}^c$ et $\varphi_k(x) = 1$ si $x \in \overline{U}_k$. En outre, comme la suite de fonction

$(\varphi_k)_{k\geqslant 0}$ est croissante est converge vers la fonction constante 1. On obtient

$$\lim_{k\to+\infty}\int_E \varphi_k \,\mathrm{d}\mu = \mu(E). \tag{3.5}$$

Soient f une fonction continue et bornée sur E, et $M = \|f\|_{\sup}$, on a, pour tout $k \in \mathbb{N}$

$$\begin{aligned}\left|\int_E f\,\mathrm{d}\mu_n - \int_E f\,\mathrm{d}\mu\right| \leqslant \left|\int_E f\varphi_k\,\mathrm{d}\mu_n - \int_E f\varphi_k\,\mathrm{d}\mu\right| \\ + M\int_E (1-\varphi_k)\,\mathrm{d}\mu_n + M\int_E (1-\varphi_k)\,\mathrm{d}\mu.\end{aligned}$$

Comme la suite de mesures $(\mu_n)_{n\in\mathbb{N}}$ converge vaguement vers μ, on obtient

$$\limsup_{n\to+\infty}\left|\int_E f\,\mathrm{d}\mu_n - \int_E f\,\mathrm{d}\mu\right| \leqslant 2M\Big(\mu(E) - \int_E \varphi_k\,\mathrm{d}\mu\Big),$$

où on a utilisé l'hypothèse que $\mu_n(E)$ tend vers $\mu(E)$ lorsque $n \to +\infty$. On en déduit, en utilisant (3.5),

$$\lim_{n\to+\infty}\int_E f\,\mathrm{d}\mu_n = \int_E f\,\mathrm{d}\mu.$$

□

On désigne par $\mathcal{R}_b(E)$ l'ensemble des mesures de Radon bornée sur E. La convergence étroite définit une topologie sur l'ensemble $\mathcal{R}_b(E)$, appelée la topologie *étroite*. C'est la topologie la moins fine qui rend continue toutes les applications

$$M_h : (\mu \in \mathcal{R}_b(E)) \longmapsto \int_E h(x)\,\mu(\mathrm{d}x),$$

où h parcourt l'ensemble des fonctions continues et bornées sur E. L'ensemble $\mathcal{R}_b(E)$ muni de la topologie étroite est un espace séparé. En effet, si μ_1 et μ_2 sont deux mesures de Radon distinctes, alors le théorème de représentation de Riesz montre qu'il existe une fonction $h \in C_c(E)$ tel que $M_h(\mu_1) \neq M_h(\mu_2)$. Les deux points μ_1 et μ_2 sont donc séparés par la fonction continue M_h.

Si f est une fonction borélienne bornée sur E, on désigne par M_f l'application de $\mathcal{R}_b(E)$ vers $[0, +\infty]$ qui envoie toute mesure de Radon bornée μ en

$$\int_E f(x)\,\mu(\mathrm{d}x).$$

On dit qu'une fonction $f : E \to \mathbb{R}$ est *semi-continue supérieurement* si l'ensemble

$$\{x \in E \mid f(x) \geqslant t\}$$

est fermé pour tout $t \in \mathbb{R}$.

Proposition 3.4.4. — *Soit $(f_i)_{i\in I}$ une famille de fonctions bornées et semi-continues supérieurement sur E. On suppose que, pour tout couple $(i,i')\in I^2$, il existe $j\in I$ tel que $f_j\leqslant f_i\wedge f_{i'}$. Alors il existe une suite $(i_m)_{m\in\mathbb{N}}$ dans I telle que la suite de fonction $(f_{i_m})_{m\in\mathbb{N}}$ soit décroissante et converge vers $\inf_{i\in I} f_i$.*

Démonstration. — Sans perte de généralité, on peut supposer que les fonctions f_i sont à valeurs dans $[0,1[$. Si f est une fonction semi-continue supérieurement sur E, qui est à valeurs dans $[0,1[$, on définit, pour tout entier $n\in\mathbb{N}$, une nouvelle fonction

$$T_nf := \sum_{k=0}^{2^n-1}\frac{k+1}{2^n}\mathbb{1}_{\{k/2^n\leqslant f<(k+1)/2^n\}} = \frac{1}{2^n}\sum_{k=0}^{2^n-1}\mathbb{1}_{A_{n,k}(f)},$$

où $A_{n,k}(f) := \{f\geqslant k/2^n\}$. C'est une fonction semi-continue supérieurement sur E telle que $T_nf\geqslant f$. De plus, la suite de fonctions $(T_nf)_{n\in\mathbb{N}}$ est décroissante et on a

$$\lim_{n\to+\infty} T_nf = f.$$

En outre, pour tout $n\in\mathbb{N}$ on a

$$T_n(\inf_{i\in I} f_i) = \inf_{i\in I} T_n(f_i).$$

Soit $g=\inf_{i\in I} f_i$. Soient $n\in\mathbb{N}$ et $k\in\{0,\dots,2^n-1\}$. On considère la famille $(A_{n,k}(f_i))_{i\in I}$ de sous-ensembles fermés de E. On a

$$\bigcap_{i\in I} A_{n,k}(f_i) = A_{n,k}(g).$$

Comme E admet une base de topologie dénombrable, il existe un sous-ensemble dénombrable $I_{n,k}$ de I tel que

$$\bigcap_{i\in I_{n,k}} A_{n,k}(f_i) = A_{n,k}(g).$$

On en déduit l'existence d'une application $\varphi_n:\mathbb{N}\to I$ telle que

$$\forall\, k\in\{0,\dots,2^n-1\},\quad \bigcap_{m\in\mathbb{N}} A_{n,k}(f_{\varphi_n(m)}) = A_{n,k}(g).$$

En outre, l'hypothèse de la proposition montre qu'il existe une application $\psi_n:\mathbb{N}\to I$ telle que $\psi_n(0)=\varphi_n(0)$ et que

$$\forall\, m\geqslant 1,\quad f_{\psi_n(m)}\leqslant f_{\psi_n(m-1)}\wedge f_{\varphi_n(0)}\wedge\cdots\wedge f_{\varphi_n(m)}.$$

On en déduit alors $\inf_{m\in\mathbb{N}} T_nf_{\psi_n(m)} = T_ng$ et donc

$$\inf_{(n,m)\in\mathbb{N}} f_{\psi_n(m)} = g.$$

En utilisant l'hypothèse de la proposition pour une fois encore, on obtient l'existence d'une suite $(i_m)_{m\in\mathbb{N}}$ telle que la suite $(f_{i_m})_{n\in\mathbb{N}}$ soit décroissante et vérifie $\inf_{m\in\mathbb{N}} = g$. La proposition est donc démontrée. □

Proposition 3.4.5. — *Soit f une fonction borélienne bornée sur E. Si f est semi-continue supérieurement*[5] *(resp. inférieurement), alors il en est de même de M_f.*

Démonstration. — Comme $M_{-f} = -M_f$, il suffit de traiter le cas où f est semi-continue supérieurement. Soit $\mathcal{C}_f$ le sous-ensemble de $\mathcal{C}_b(E)$ des fonctions qui sont bornées inférieurement par f. Montrons que l'on a

$$f = \inf_{g\in C_f} g.$$

Soient x_0 un point de E et a un nombre réel tel que $f(x_0) < a$. Montrons qu'il existe une fonction continue $g : E \to \mathbb{R}$ telle que $g \geqslant f$ et $g(x_0) \leqslant a$. Comme $f(x_0) < a$, il existe un voisinage ouvert U de x_0 tel que $f(x) < a$ pour tout $x \in U$. D'après le lemme d'Uryson (voir aussi la remarque), il existe une fonction continue $\varphi : E \to [0,1]$ tel que $\varphi(x_0) = 0$ et que $\varphi(x) = 1$ pour tout $x \in E\setminus U$. Soit $g = (b-a)\varphi + a$, où $b = \sup_{x\in E} f(x)$. On a $g \geqslant f$ et $g(x_0) = a$.

Montrons que $M_f = \inf_{g\in C_f} M_g$. Soit μ une mesure de Radon finie sur E. D'après la proposition précédente, il existe une suite décroissante $(g_n)_{n\in\mathbb{N}}$ dans C_f telle que $\inf\limits_{n\in\mathbb{N}} g_n = f$. On obtient alors

$$I_\mu(f) = \lim_{n\to+\infty} I_\mu(g_n) \geqslant \inf_{g\in C_f} I_\mu(g) \geqslant I_\mu(f),$$

d'où $M_f(\mu) = \inf_{g\in C_f} M_g(\mu)$. Comme M_g est une fonction continue sur $\mathcal{R}_b(E)$ pour toute $g \in C_f$, on obtient que M_f est semi-continue supérieurement. □

Théorème 3.4.6. — *Soient f une fonction borélienne et bornée sur E et μ est une mesure de Radon finie sur E. Si l'ensemble des points de discontinuité de f est μ-negligeable, alors la fonction M_f est continue en μ.*

Démonstration. — Soient $\overline{f}$ l'enveloppe semi-continue supérieurement de f, et $\underline{f}$ l'enveloppe semi-continue inférieurement de f. Par définition, on a

$$\forall x \in E, \quad \overline{f}(x) = \limsup_{y\to x} f(y), \quad \underline{f}(x) = \liminf_{y\to x} f(y).$$

5. On dit qu'une fonction $f : E \to \mathbb{R}$ est semi-continue supérieurement (resp. inférieurement) si $\{x \in E \,|\, f(x) \geqslant t\}$ (resp. $\{x \in E \,|\, f(x) \leqslant t\}$) est un fermé de E quel que soit t.

L'ensemble $A := \{x \in E \,|\, \overline{f}(x) = \underline{f}(x)\}$ s'identifie à l'ensemble des points de continuité de la fonction f. En outre, les relations $\overline{f} \geqslant f \geqslant \underline{f}$ montrent que $M_{\overline{f}} \geqslant M_f \geqslant M_{\underline{f}}$. Comme le complémentaire de A est μ-negligeable, on a $M_{\overline{f}}(\mu) = M_{\underline{f}}(\mu)$. Cela montre que la fonction M_f est continue en μ. □

Corollaire 3.4.7. — *Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de Radon bornées sur E, qui converge étroitement vers une mesure μ. Si f est une fonction borélienne et bornée telle que l'ensemble des points de discontinuité de f est μ-negligeable, alors on a*

$$\lim_{n\to+\infty} \int_E f(x)\,\mu_n(\mathrm{d}x) = \int_E f(x)\,\mu(\mathrm{d}x).$$

Démonstration. — C'est une conséquence immédiate de la continuité de M_f en μ. □

Définition 3.4.8. — Soit μ une mesure de Radon sur E. On appelle *support* de μ le complémentaire du sous-ensemble ouvert de E des points $x \in E$ tels qu'il existe une fonction continue et positive f, vérifiant $f(x) > 0$ et $I_\mu(f) = 0$. On désigne par $\mathrm{supp}(\mu)$ le support de μ.

Proposition 3.4.9. — *Soit μ une mesure de Radon sur E. Si $f : E \to \mathbb{R}$ est une fonction borélienne qui s'annule sur le support de μ, alors on a $I_\mu(f) = 0$.*

Démonstration. — Soit x un point en dehors du support de μ. D'après le théorme d'Uryson et la remarque 3.3.4, il existe une fonction continue $\varphi : E \to [0,1]$ qui s'annule sur le support de E et telle que $\varphi(x) = 1$. On en déduit donc qu'il existe un voisinage ouvert de x qui est μ-negligeable. Comme E admet une base de topologie dénombrable, on obtient que le complémentaire du support de μ est une réunion dénombrable d'ensembles μ-negligeables, donc est lui-même μ-negligeable. Le lemme est donc démontré. □

3.5. Fonction caractéristique

Dans ce sous-paragraphe, on étudie les mesures de Radon finies sur $\mathbb{R}^d$, où $d \geqslant 1$ est un entier. On désigne par $\langle\,,\,\rangle$ le produit scalaire usuel sur $\mathbb{R}^d$.

Définition 3.5.1. — Soit μ une mesure borélienne finie sur $\mathbb{R}^d$. On appelle *fonction caractéristique* de μ l'application $\chi_\mu : \mathbb{R}^d \to \mathbb{C}$ définie comme

$$\chi_\mu(\boldsymbol{t}) = \int_E \mathrm{e}^{i\langle \boldsymbol{t},\boldsymbol{x}\rangle}\,\mu(\mathrm{d}\boldsymbol{x}) := \int_E \cos(\langle \boldsymbol{t},\boldsymbol{x}\rangle)\,\mu(\mathrm{d}\boldsymbol{x}) + i\int_E \sin(\langle \boldsymbol{t},\boldsymbol{x}\rangle)\,\mu(\mathrm{d}\boldsymbol{x}).$$

Si μ est une mesure de probabilité et si $\boldsymbol{X}$ est une variable aléatoire à valeur dans $\mathbb{R}^d$ qui suit la loi μ, alors on a

$$\forall\, \boldsymbol{t} \in \mathbb{R}^d, \quad \chi_\mu(\boldsymbol{t}) = \mathbb{E}[\mathrm{e}^{i\langle \boldsymbol{t}, \boldsymbol{X}\rangle}].$$

Exemple 3.5.2. — Pour tout $\sigma > 0$ soit $N(0, \sigma^2)$ la loi gausienne de paramètre $(0, \sigma^2)$ sur $\mathbb{R}$. On a

$$\chi_{N(0,\sigma^2)}(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \mathrm{e}^{-x^2/2\sigma^2 - itx}\, \mathrm{d}x$$

Pour tout $x \in \mathbb{R}$, la fonction $t \mapsto \mathrm{e}^{-x^2/2\sigma^2 - itx}$ est clairement dérivable et sa dérivée est $t \mapsto -ix\mathrm{e}^{-x^2/2\sigma^2 - itx}$. On a

$$|-ix\mathrm{e}^{-x^2/2\sigma^2 - itx}| = x\mathrm{e}^{-x^2/2\sigma^2},$$

qui est une fonction intégrable sur $\mathbb{R}$. D'après le théorème de dériviation sous signe somme, on obtient que la fonction $\chi_{N(0,\sigma^2)}$ est dérivable, et

$$\chi'_{N(0,\sigma^2)}(t) = \frac{-i}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} x\mathrm{e}^{-x^2/2\sigma^2 - itx}\, \mathrm{d}x.$$

En faisant une intégration par partie avec $-x\mathrm{e}^{-x^2/2\sigma^2}$ et e^{-itx}, on obtient

$$\chi'_{N(0,\sigma^2)}(t) = \frac{i\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} it\mathrm{e}^{-x^2/2\sigma^2 - itx}\, \mathrm{d}x = -\sigma^2 t \chi_{N(0,\sigma^2)}(t),$$

d'où $\chi_{N(0,\sigma^2)}(t) = \mathrm{e}^{-\sigma^2 t^2/2}$ puisque $\chi_{N(0,\sigma^2)}(0) = 1$. On en déduit que, pour tout $\mu \in \mathbb{R}$,

$$\chi_{N(\mu,\sigma^2)} = \mathrm{e}^{i\mu t - \sigma^2 t^2/2}. \tag{3.6}$$

On désigne par $N(\mathbf{0}, \sigma^2 I_d)$ la loi $N(0, \sigma^2)^{\otimes d}$ sur $\mathbb{R}^d$. On a

$$\forall\, \boldsymbol{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d, \quad \chi_{N(\mathbf{0},\sigma^2 I_d)}(\boldsymbol{t}) = \prod_{i=1}^d \chi_{N(0,\sigma^2)}(t_i) = \mathrm{e}^{-\sigma^2(t_1^2 + \cdots + t_d^2)/2}.$$

Définition 3.5.3. — Soit μ une mesure borélienne finie sur $\mathbb{R}^d$. Si $f : \mathbb{R}^d \to \mathbb{R}$ est une mesure borélienne bornée, on définit $f * \mu$ la fonction $\mathbb{R}^d \to \mathbb{R}$ définie comme

$$\forall\, \boldsymbol{x} \in \mathbb{R}^d, \quad (f * \mu)(\boldsymbol{x}) := \int_{\mathbb{R}^d} f(\boldsymbol{x} - \boldsymbol{y})\, \mu(\mathrm{d}\boldsymbol{y})$$

Proposition 3.5.4. — *Soit μ est une mesure borélienne finie sur $\mathbb{R}^d$. Pour tout $\sigma > 0$, soit $g_{d,\sigma^2} : \mathbb{R}^d \to \mathbb{R}$ la fonction de densité de la loi $N(\mathbf{0}, \sigma^2 I_d)$:*

$$g_{d,\sigma}(t_1, \ldots, t_d) = \frac{1}{(2\pi\sigma^2)^{d/2}} \mathrm{e}^{-(t_1^2 + \cdots + t_d^2)/2\sigma^2}.$$