Faculté des Sciences
Aix-Marseille Université

CTES
Centre de Télé-Enseignement Sciences

MASTER 1 MATHEMATIQUES – INFORMATIQUE
MATHEMATIQUES GENERALES
M1...

Expédition dans la semaine n°	Etape	Code UE	N° d'envoi de l'UE
11	**M1, UE 1.3**	**MIMA-SMAAU11T**	**5**

Nom de l'UE : Mesure, Intégration, Probabilités

Le cours est essentiellement tiré d'un livre disponible sur la page web indiquée ci dessous. Pour chaque semaine, il est proposé d'étudier une partie du cours et de faire des exercices (corrigés).
Deux devoirs sont à rendre afin de bénéficier d'une note de contrôle continu.
note finale=max(note-examen, 1/3(2 note-examen + note-contrôle-continu)).

Contenu de l'envoi : Cours : Chapitres 10 et 11, Exercices, Devoir 2
Le chapitre 11 est hors programme, il est intéressant en liaison avec les autres cours

- Guide du travail à effectuer

Semaine 1 (Transformation de Fourier dans L1, S et d'une mesure) :
Etudier les paragraphes 10.1, 10.2 (L1 et S) et 10.3 (mesure signée)
Exercices proposés (avec corrigés) : 10.1, 10.2, 10.4

Semaine 2 (Transformation de Fourier dans L2) :
Etudier le paragraphe 10.4 (L2)
Exercices proposés (avec corrigés) : 10.7, 10.8

Semaine 3 (Fonctions caractéristiques) :
Etudier le début du paragraphe 10.6 (fonction caractéristique d'un vecteur aléatoire)
Exercices proposés (avec corrigés) : 10.15, 10.16, 10.17

Semaine 4 (Théorème central limite) :
Etudier la fin du paragraphes 10.6 (théorème de Paul Levy et théorème central limite)
Exercices proposés (avec corrigés) : 10.12, 10.19
Faire le devoir 2

-Coordonnées des enseignants responsables de l'envoi
E. Hillion et T. Gallouet, CMI, 39 rue Joliot Curie, 13453 marseille cedex 13
email : thierry.gallouet@univ-amu.fr, erwan.hillion@univ-amu.fr
Vous pouvez aussi consulter la page web:
http://www.cmi.univ-mrs.fr/~gallouet/tele.d/mip.d
et nous poser des questions par le forum d'ametice ou par email

Chapitre 10

Transformation de Fourier

10.1 Introduction et notations

La notion de série de Fourier permet d'analyser les fonctions définies d'un compact $[a,b]$ de $\mathbb{R}$ dans $\mathbb{R}$ (ou dans $\mathbb{C}$), et donc aussi les fonctions périodiques de $\mathbb{R}$ dans $\mathbb{R}$ (ou $\mathbb{C}$). La notion de transformée de Fourier permet d'analyser les fonctions définies de $\mathbb{R}$ (ou $\mathbb{R}^{\mathrm{N}}$) dans $\mathbb{R}$ (ou $\mathbb{C}$). Les deux notions ont une multitude d'applications en physique et en mathématiques. La transformée de Fourier est une notion employée par exemple en théorie du signal, en théorie des probabilités et pour l'analyse des équations aux dérivées partielles.

Dans toute la suite, on considérera l'espace mesuré $(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ et on notera $d\lambda_{\mathrm{N}}(x) = dx$. Soit $\mathrm{N} \geq 1$, les espaces $\mathcal{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ et $\mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ ont été définis dans la section 4.10 et les espaces $\mathcal{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ et $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ ont été définis dans la section 6.2. On rappelle aussi que si f est une fonction définie de $\mathbb{R}^{\mathrm{N}}$ dans $\mathbb{C}$, la fonction f est mesurable si et seulement si ses parties réelle et imaginaire sont mesurables (chaque ensemble, $\mathbb{R}^{\mathrm{N}}$, $\mathbb{R}$ ou $\mathbb{C}$, étant muni de sa tribu borélienne). On peut, bien sûr, aussi définir les espaces $\mathcal{L}^p_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ et $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}, \mathcal{B}(\mathbb{R}^{\mathrm{N}}), \lambda_{\mathrm{N}})$ pour tout $p \in [1,\infty]$.

Définition 10.1 (Espaces $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}},\mathcal{B}(\mathbb{R}^{\mathrm{N}}),\lambda_{\mathrm{N}})$)

Soit $\mathrm{N} \geq 1$, $p \in [1,\infty]$ *et* f *une fonction mesurable de* $\mathbb{R}^{\mathrm{N}}$ *dans* $\mathbb{C}$ *(c'est-à-dire* $f^{-1}(\mathrm{A}) \in \mathcal{B}(\mathbb{R}^{\mathrm{N}})$ *pour tout* $\mathrm{A} \in \mathcal{B}(\mathbb{C})$*).*

On dit que $f \in \mathcal{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ si $|f| \in \mathcal{L}^p_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et on définit $\|f\|_p$ par $\|f\|_p = \||f|\|_p$ où $\||f|\|_p$ est la norme de $|f|$ dans $\mathcal{L}^p_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (vue au chapitre 6).

L'espace $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ est l'espace $\mathcal{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ quotienté par le relation d'équivalence "= p.p.". C'est un espace de Banach (complexe), c'est-à-dire un e.v.n. (sur $\mathbb{C}$) complet.

Remarque 10.2 Soit $N \geq 1$, $p \in [1, \infty]$ et f une fonction définie de $\mathbb{R}^N$ dans $\mathbb{C}$. Il est facile de voir que $f \in \mathcal{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ si et seulement si ses parties réelle et imaginaire sont dans $\mathcal{L}^p_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$.

10.2 Transformation de Fourier dans L^1

10.2.1 Définitions et premières propriétés

Soit $N \geq 1$. Pour $x = (x_1, \ldots, x_N)^t \in \mathbb{R}^N$ et $t = (t_1, \ldots, t_N)^t \in \mathbb{R}^N$, on note $x \cdot t$ le produit scalaire euclidien de x et t, c'est-à-dire $x \cdot t = \sum_{i=1}^N x_i t_i$. Dans ce chapitre, On note aussi $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^N)$ ou parfois simplement $\mathrm{L}^p_{\mathbb{C}}$ l'espace $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, pour $p \in [1, \infty]$.

Définition 10.3 (Transformée de Fourier dans L^1) *Soit $N \geq 1$ et $f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^N)$. Pour $t \in \mathbb{R}^N$, l'application $x \mapsto e^{-ix\cdot t} f(x)$ (définie de $\mathbb{R}^N$ dans $\mathbb{C}$) appartient à $\mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^N)$. On définit alors $\widehat{f}(t)$ par :*

$$\widehat{f}(t) = (2\pi)^{-\frac{N}{2}} \int f(x) e^{-ix\cdot t} dx.$$

La fonction $\widehat{f}$ (définie de $\mathbb{R}^N$ dans $\mathbb{C}$) s'appelle transformée de Fourier de f, qu'on notera également $\mathcal{F}(f)$.

On note $C_0(\mathbb{R}^N, \mathbb{C}) = \{g \in C(\mathbb{R}^N, \mathbb{C})$ t.q. $g(t) \to 0$ quand $|t| \to +\infty\}$. On rappelle que $C_0(\mathbb{R}^N, \mathbb{C})$ est un espace de Banach quand il est muni de la norme de la convergence uniforme, c'est-à-dire :

$$\|\varphi\|_u = \max_{x \in \mathbb{R}^N} |\varphi(x)|.$$

Proposition 10.4 (Linéarité de la transformée de Fourier) *Soit $N \geq 1$. L'application $\mathcal{F}$ qui à f (appartenant à $\mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^N)$) associe sa transformée de Fourier est linéaire continue de $\mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^N)$ dans $C_0(\mathbb{R}^N, \mathbb{C})$.*

DÉMONSTRATION – Le théorème de continuité sous le signe $\int$ (théorème 4.52) appliqué à la fonction $(x,t) \mapsto e^{-ix.t}f(x)$ entraîne immédiatement que $\widehat{f}$ est continue. Montrons que $\widehat{f} \in C_0(\mathbb{R},\mathbb{R})$.

Cas N = 1 . On remarque que pour $t \neq 0$, on a, comme $e^{i\pi} = -1$,

$$\widehat{f}(t) = -(2\pi)^{-\frac{1}{2}} \int e^{-i(x-\frac{\pi}{t})t} f(x)dx,$$

et donc avec un changement de variable simple,

$$\widehat{f}(t) = -(2\pi)^{-\frac{1}{2}} \int e^{-iyt} f(y+\frac{\pi}{t})dy.$$

On en déduit que

$$2\widehat{f}(t) = (2\pi)^{-\frac{1}{2}} \int e^{-ixt}(f(x) - f(x+\frac{\pi}{t}))dx$$

et donc que $|\widehat{f}(t)| \leq \frac{1}{2}(2\pi)^{-\frac{1}{2}}\|f(\cdot) - f(\cdot + \frac{\pi}{t})\|_1$. Le théorème de continuité en moyenne dans L^1 (théorème 5.21) donne alors le fait que $\widehat{f}(t) \to 0$ quand $|t| \to \infty$.

Cas N > 1 . On reprend la même méthode.

Pour $t \neq 0$, $t = (t_1, \ldots, t_N)^t$, il existe $j \in \{1,\ldots,N\}$ t.q. $|t_j| = \max_{k=1,\ldots,N}|t_k|$. On a alors, comme $e^{i\pi} = -1$, en notant e_j le j–ième vecteur de base de $\mathbb{R}^N$,

$$\widehat{f}(t) = -(2\pi)^{-\frac{N}{2}} \int e^{-i(x-\frac{\pi}{t_j}e_j)\cdot t} f(x)dx,$$

et donc avec un changement de variable simple,

$$\widehat{f}(t) = -(2\pi)^{-\frac{N}{2}} \int e^{-iy\cdot t} f(y+\frac{\pi}{t_j}e_j)dy.$$

On en déduit que $|\widehat{f}(t)| \leq \frac{1}{2}(2\pi)^{-\frac{N}{2}}\|f(\cdot) - f(\cdot + \frac{\pi}{t_j}e_j)\|_1$. Le théorème de continuité en moyenne dans L^1 (théorème 8.5) donne alors le fait que $\widehat{f}(t) \to 0$ quand $|t| \to \infty$.

■

La transformée de Fourier a la propriété intéressante de transformer la convolution en produit. Ceci est montré dans la proposition suivante.

Proposition 10.5 (Transformée de Fourier d'un produit) *Soient* f *et* $g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$, *alors* $\widehat{f * g} = (2\pi)^{\frac{N}{2}}\widehat{f}\widehat{g}$.

DÉMONSTRATION – Par la proposition 7.22, on a $f * g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ et pour p.p. $x \in \mathbb{R}^N$, $f * g(x) = \int f(x-y)g(y)dy$. On a donc, pour tout $t \in \mathbb{R}^N$,

$$\widehat{f * g}(t) = (2\pi)^{-\frac{N}{2}} \int \left(\int f(x-y)g(y)dy \right) e^{-ix\cdot t}dx =$$

$$(2\pi)^{-\frac{N}{2}} \int \left(\int f(x-y)g(y)e^{-i(x-y)\cdot t}e^{-iy\cdot t}dy \right) dx.$$

En appliquant le théorème de Fubini (théorème 7.12) à la fonction

$$(x,y) \mapsto f(x-y)g(y)e^{-i(x-y)\cdot t}e^{-iy\cdot t}$$

(qui est bien intégrable sur $\mathbb{R}^{2\mathrm{N}}$ car son module est la fonction $(x,y) \mapsto |f(x-y)g(y)|$ dont l'intégrale sur $\mathbb{R}^{2\mathrm{N}}$ est égale à $\|f\|_1\|g\|_1$), on obtient :

$$\widehat{f*g}(t) = (2\pi)^{-\frac{\mathrm{N}}{2}} \int \left(\int f(x-y)e^{-i(x-y)\cdot t}dx \right) g(y)e^{-iy\cdot t}dy.$$

Comme, pour tout $y \in \mathbb{R}^{\mathrm{N}}$, $\int f(x-y)e^{-i(x-y)\cdot t}dx = \int f(z)e^{-iz\cdot t}dz = (2\pi)^{\frac{\mathrm{N}}{2}}\widehat{f}(t)$, on en déduit :

$$\widehat{f*g}(t) = \widehat{f}(t) \int g(y)e^{-iy\cdot t}dy = (2\pi)^{\frac{\mathrm{N}}{2}}\widehat{f}(t)\widehat{g}(t),$$

ce qui est le résultat annoncé. ■

10.2.2 Théorème d'inversion

Il est naturel de se poser les deux questions suivantes :

La transformée de Fourier d'une fonction f caractérise-t-elle la fonction f (c'est-à-dire si $\widehat{f} = \widehat{g}$, a-t-on $f = g$ p.p.) ?

Peut-on retrouver la fonction à partir de sa transformée de Fourier ?

Les réponses à ces questions sont fournies par le théorème d'inversion de Fourier :

Théorème 10.6 (Inversion partielle de la transformée de Fourier) *Soit* $\mathrm{N} \geq 1$ *et* $f \in \mathrm{L}^1_{\mathbb{C}}$ *t.q.* $\widehat{f} \in \mathrm{L}^1_{\mathbb{C}}$. *On a alors* $f = \widehat{\widehat{f}}(-.)$ *p.p., c'est-à-dire :*

$$f(t) = (2\pi)^{-\frac{\mathrm{N}}{2}} \int \widehat{f}(x)e^{ixt}dx, \textit{ pour presque tout } t \in \mathbb{R}^{\mathrm{N}}.$$

La démonstration fait l'objet de l'exercice 10.1.

Une conséquence de ce théorème est l'injectivité de l'application $\mathcal{F}$, qui fournit donc une réponse positive à la première question. En effet, soient f et $g \in \mathrm{L}^1_{\mathbb{C}}$ t.q. $\widehat{f} = \widehat{g}$; alors par linéarité, $\widehat{f-g} = 0$ et donc $\widehat{f-g} \in \mathrm{L}^1_{\mathbb{C}}$. En appliquant le théorème d'inversion, on a donc $f = g$ p.p..

Ce théorème apporte aussi une réponse partielle à la deuxième : on peut calculer f à partir de $\widehat{f}$ dès que $\widehat{f} \in \mathrm{L}^1_{\mathbb{C}}$. Il faut remarquer à ce propos que $\mathrm{L}^1_{\mathbb{C}}$ n'est pas stable par transformation de Fourier (voir exercice 10.2).

10.2.3 Régularité et décroissance à l'infini

Proposition 10.7 (Différentiabilité, dimension 1)

1. *Soit* $f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}) \cap \mathrm{C}^1(\mathbb{R}, \mathbb{C})$ *t.q.* $f' \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R})$ *(où* f' *est la dérivée de* f*). Alors, pour tout* $t \in \mathbb{R}$, $\mathcal{F}(f')(t) = (it)\mathcal{F}(f)(t)$.
2. *Soit* $f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R})$ *t.q.* $(\cdot)f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R})$ *(où* $(\cdot)f$ *est l'application qui à* $x \in \mathbb{R}$ *associe* $xf(x)$*). Alors,* $\mathcal{F}(f) \in \mathrm{C}^1(\mathbb{R}, \mathbb{C})$ *et, pour tout* $t \in \mathbb{R}$, $\mathcal{F}(f)'(t) = \mathcal{F}((-i\cdot)f)(t)$.

DÉMONSTRATION – La démonstration du premier item consiste à faire une intégration par parties sur l'intervalle $[-n, n]$ puis à faire tendre n vers l'infini (en remarquant que $f(\pm n) \to 0$, quand $n \to +\infty$, voir l'exercice 5.6).

Le deuxième item est une conséquence immédiate du théorème 4.53 (théorème de dérivation sous le signe $\int$). ■

La transformation de Fourier transforme donc la dérivation en multiplication par la fonction $(i\cdot)$, et la multiplication par $(-i\cdot)$ en dérivation. Cette propriété est utilisée, par exemple, pour la résolution d'équations différentielles (qui sont ainsi transformées en équations algébriques).

Cette propriété se généralise au cas de la dimension N et pour un ordre k de dérivation quelconque. On introduit pour ce faire les notations suivantes : soient $\alpha = (\alpha_1, \ldots, \alpha_{\mathrm{N}})^t \in \mathbb{N}^{\mathrm{N}}$ un multi-indice et f une fonction de $\mathbb{R}^{\mathrm{N}}$ dans $\mathbb{C}$. On définit $|\alpha| = \alpha_1 + \alpha_2 + \ldots \alpha_{\mathrm{N}}$ et

$$\mathrm{D}^\alpha f = \Big(\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_{\mathrm{N}}}}{\partial x_{\mathrm{N}}^{\alpha_{\mathrm{N}}}}\Big) f$$

lorsque cette dérivée existe.

Proposition 10.8 (Différentiabilité, dimension N)

1. *Soit* $\mathrm{N} \geq 1$ *et* $k \geq 1$. *Soit* $f \in \mathrm{C}^k(\mathbb{R}^{\mathrm{N}}, \mathbb{C})$ *t.q.* $\mathrm{D}^\alpha f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ *pour tout* $\alpha \in \mathbb{N}^{\mathrm{N}}$ *t.q.* $|\alpha| \leq k$. *Alors,* $\widehat{\mathrm{D}^\alpha f}(t) = (it)^\alpha \widehat{f}(t)$ *pour tout* $t \in \mathbb{R}$ *et pour tout* $\alpha \in \mathbb{N}^{\mathrm{N}}$ *t.q.* $|\alpha| \leq k$, *avec* $(it)^\alpha = (it_1)^{\alpha_1}(it_2)^{\alpha_2} \ldots (it_{\mathrm{N}})^{\alpha_{\mathrm{N}}}$.
2. *Soit* f *t.q.* $(.)^\alpha f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ *pour tout* $\alpha \in \mathbb{N}^{\mathrm{N}}$ *tel que* $|\alpha| \leq k$. *Alors,* $\widehat{f} \in \mathrm{C}^k(\mathbb{R}^{\mathrm{N}}, \mathbb{C})$ *et* $\mathrm{D}^\alpha \widehat{f} = \widehat{(-i\cdot)^\alpha f}$ *pour tout* $\alpha \in \mathbb{N}^{\mathrm{N}}$ *tel que* $|\alpha| \leq k$.

La proposition 10.8 montre que la dérivabilité de f entraîne la décroissance de $\widehat{f}$ à l'infini (plus f est dérivable, plus $\widehat{f}$ décroît vite à l'infini), et réciproquement. Cette remarque incite à définir l'espace des fonctions à décroissance rapide (souvent appelé espace de Schwartz),

Définition 10.9 (Espace de Schwartz) *Soit* $N \geq 1$, *on appelle espace de Schwartz, noté* $\mathcal{S}(\mathbb{R}^N, \mathbb{C})$ *ou* $\mathcal{S}_N$ *ou même* $\mathcal{S}$, *l'espace des fonctions dites à décroissance rapide, défini par :*

$$\mathcal{S}_N = \{f \in C^\infty(\mathbb{R}^N, \mathbb{C}) \text{ t.q. pour tout } \alpha \text{ et } \beta \in \mathbb{N}^N, \sup_{x \in \mathbb{R}^N} |(x)^\alpha D^\beta f(x)| < +\infty\}.$$

On va montrer l'invariance par transformation de Fourier de cet espace. On commence par remarquer que $\mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$. En effet, si $f \in \mathcal{S}_N$, en prenant des choix convenables de α et β dans la définition de $\mathcal{S}_N$, on remarque qu'il existe $C \in \mathbb{R}$ t.q. $(1 + |x|^{N+1})|f(x)| \leq C$. On en déduit que $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. Plus généralement, si $f \in \mathcal{S}_N$, on remarque que $(\cdot)^\beta D^\alpha f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ pour tout $a, \beta \in \mathbb{R}^N$. On obtient alors la proposition 10.10.

Proposition 10.10 (Transformée de Fourier et dérivation dans $\mathcal{S}$) *Soit* $N \geq 1$ *et* $f \in \mathcal{S}(\mathbb{R}^N, \mathbb{C})$. *Pour tout* $\alpha, \beta \in \mathbb{N}^N$ *on a :*

$$\mathcal{F}\Big(D^\alpha((-i\cdot)^\beta f)\Big) = (i\cdot)^\alpha \widehat{(-i\cdot)^\beta f} = (i\cdot)^\alpha D^\beta \widehat{f}, \tag{10.1}$$

$$\mathcal{F}[(-i\cdot)^\alpha D^\beta f] = D^\alpha \widehat{D^\beta f} = D^\alpha[(i\cdot)^\beta \widehat{f}]. \tag{10.2}$$

La démonstration est une adaptation simple de celle de la proposition 10.8.

La proposition 10.10 et le théorème 10.6 nous permettent alors de remarquer que la transformée de Fourier est une bijection de $\mathcal{S}_N$ dans $\mathcal{S}_N$.

Proposition 10.11 (Transformée de Fourier dans S_N) *Soit* $N \geq 1$. *L'application* $\mathcal{F}$ *qui à* f *associe sa transformée de Fourier est une bijection de* $\mathcal{S}_N$ *dans* $\mathcal{S}_N$. *De plus, pour tout* $f \in \mathcal{S}_N$, *on a* $f = \widehat{\widehat{f}}(-\cdot)$.

DÉMONSTRATION – En utilisant la proposition 10.10, on montre que $\mathcal{F}$ envoie $\mathcal{S}_N$ dans $\mathcal{S}_N$. Le théorème d'inversion (théorème 10.6) donne alors que f est injective et que $f = \widehat{\widehat{f}}(-\cdot)$ pour tout $f \in \mathcal{S}_N$. De cette dernière formule, on déduit que $\mathcal{F}$ est surjective (et donc bijective) de $\mathcal{S}_N$ dans $\mathcal{S}_N$. ■

10.3 Transformée de Fourier d'une mesure signée

Il est facile d'étendre la définition de la transformation de Fourier à une mesure signée sur les boréliens de $\mathbb{R}^N$. Plus précisément, si m est une mesure signée sur les boréliens de $\mathbb{R}^N$ ($N \geq 1$), la définition 10.12 définit la fonction $\widehat{m}$ (continue et bornée de $\mathbb{R}^N$ dans $\mathbb{C}$) et, si $m = f\lambda_N$, avec $f \in L^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (c'est-à-dire que m est la mesure signée de densité f par rapport à la mesure de Lebesgue sur les boréliens de $\mathbb{R}^N$), on a $\widehat{m}(t) = \widehat{f}(t)$ pour tout $t \in \mathbb{R}^N$.

De même, si m est une mesure à valeurs complexes, on a alors $m = m_1 + im_2$, où m_1 et m_2 sont des mesures signées (ce sont les parties réelle et imaginaire de m). On peut alors définir $\widehat{m}$. C'est la fonction continue bornée de $\mathbb{R}^N$ dans $\mathbb{C}$ donnée par $\widehat{m}(t) = \widehat{m}_1(t) + i\widehat{m}_2(t)$ pour tout $t \in \mathbb{R}^N$.

Définition 10.12 (Transformée de Fourier d'une mesure signée) *Soit* $N \geq 1$ *et* m *une mesure signée sur les boréliens de* $\mathbb{R}^N$. *Soit* $t \in \mathbb{R}^N$, *l'application* $x \mapsto e^{-ix\cdot t}$ *(définie de* $\mathbb{R}^N$ *dans* $\mathbb{C}$*) appartient à* $L^1_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), m)$ *(car elle est continue, donc borélienne, et bornée). On définit alors* $\widehat{m}(t)$ *par :*

$$\widehat{m}(t) = (2\pi)^{-\frac{N}{2}} \int e^{-ix\cdot t} dm(x).$$

La fonction $\widehat{m}$ *(définie de* $\mathbb{R}^N$ *dans* $\mathbb{C}$*) s'appelle transformée de Fourier de* m.

On rappelle que

$$C_b(\mathbb{R}^N, \mathbb{C}) = \{g \in C(\mathbb{R}^N, \mathbb{C}) \text{ t.q. } \sup_{t \in \mathbb{R}^N} |g(t)| < \infty\}$$

et que $C_b(\mathbb{R}^N, \mathbb{C})$ est un espace de Banach quand il est muni de la norme de la convergence uniforme.

Proposition 10.13 *Soit* $N \geq 1$. *Soit* m *une mesure signée sur les boréliens de* $\mathbb{R}^N$. *La fonction* $\widehat{m}$ *appartient à* $C_b(\mathbb{R}^N, \mathbb{C})$.

DÉMONSTRATION – Le fait que $\widehat{m}$ est bornée est simple. On utilise la décomposition de Hahn (proposition 2.33), c'est-à-dire le fait que $m = m^+ - m^-$, où $m^\pm$ sont deux mesure finies étrangères. On remarque alors que, pour tout $t \in \mathbb{R}^N$, on a $|\widehat{m}(t)| \leq |m|(\mathbb{R}^N) < +\infty$, où $|m| = m^+ + m^-$.

La fait que $\widehat{m}$ est continue est une conséquence immédiate du théorème de continuité sous le signe $\int$ (théorème 4.52) appliqué à la fonction $(x,t) \mapsto e^{-ix.t}$ (et avec les mesures finies $m^\pm$). ∎

Comme pour la transformation de Fourier dans L^1, on peut se demander si $\widehat{m}$ caractérise la mesure signée m et si on peut retrouver m à partir de $\widehat{m}$. La proposition suivante s'intéresse à ces deux questions.

Proposition 10.14 *Soit* $\mathrm{N} \geq 1$.

1. *Soit* m *et* μ *deux mesures signées sur les boréliens de* $\mathbb{R}^{\mathrm{N}}$. *On suppose que* $\widehat{m} = \widehat{\mu}$. *alors,* $m = \mu$.
2. *Soit* m *une mesure signée sur les boréliens de* $\mathbb{R}^{\mathrm{N}}$. *On suppose que* $\widehat{m} \in \mathcal{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. *Alors,* $m = f\lambda_{\mathrm{N}}$ *avec* $f = \widehat{\widehat{m}}(-\cdot)$ *p.p. (c'est-à-dire p.p. pour la mesure* λ_{N}*).*

La démonstration de cette proposition fait l'objet de l'exercice 10.6.

Nous avons vu dans la section précédente qu'il y avait un lien entre les propriétés de décroissance de f à l'infini et les propriétés de régularité de $\widehat{f}$ (propositions 10.7, 10.8 et 10.10). Dans le même esprit, on peut remarquer que, si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}_+$, borélienne et intégrable, la continuité de $\widehat{f}$ en 0 donne une précision sur la convergence vers 0, quand $a \to +\infty$, de $\int_{]-a,a[^c} |f(x)| dx$. Nous donnons dans le lemme 10.15 cette précision dans le cas plus général d'une mesure (positive) finie. Ce lemme permet de démontrer le théorème 10.22 (lui-même utile pour démontrer le théorème central limite, théorème 10.24).

Lemme 10.15 *Soit* m *une mesure (positive) finie (et non nulle) sur* $\mathcal{B}(\mathbb{R})$. *On pose* $\mathrm{M} = m(\mathbb{R})$ *et, pour* $t \in \mathbb{R}$,

$$\psi(t) = \sqrt{2\pi}\frac{\widehat{m}(t) + \widehat{m}(-t)}{2}.$$

(La fonction ψ *prend donc ses valeurs dans* $\mathbb{R}$*, et même dans* $[-\mathrm{M}, \mathrm{M}]$*, est continue en* 0 *et* $\psi(0) = \mathrm{M}$*.) On a alors, pour tout* $a > 0$,

$$m(\{x, |x| \geq a\}) \leq a \int_0^{2/a} (\mathrm{M} - \psi(t)) dt.$$

DÉMONSTRATION – On peut supposer $\mathrm{M} = 1$ (il suffit, si $\mathrm{M} \neq 1$, de remplacer la mesure m par la mesure m/M). On remarque tout d'abord que pour tout $t \in \mathbb{R}$ on a

$$\widehat{m}(t) + \widehat{m}(-t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (e^{-ixt} + e^{ixt}) dm(x) = \frac{2}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(xt) dm(x).$$

et donc

$$\psi(t) = \int_{\mathbb{R}} \cos(xt) dm(x).$$

On en déduit bien que $\psi(t) \in \mathbb{R}$, $|\psi(t)| \leq 1$ et que $\psi(0) = 1$.

Soit $a > 0$, on pose $\varepsilon = 2/a$ et $\mathrm{T} = \int_0^{\varepsilon} (1 - \psi(t)) dt$. En utilisant le théorème de Fubini-Tonelli (théorème 7.7) on obtient

$$\begin{aligned} \mathrm{T} = \int_0^{\varepsilon} \Big(\int_{\mathbb{R}} (1 - \cos(xt)) dm(x) \Big) dt &= \int_{\mathbb{R}} \Big(\int_0^{\varepsilon} (1 - \cos(xt)) dt \Big) dm(x) \\ &= \int_{\mathbb{R}} (\varepsilon - \frac{\sin(\varepsilon x)}{x}) dm(x). \end{aligned}$$

Mais, $(\varepsilon - \frac{\sin(\varepsilon x)}{x}) = \varepsilon(1 - \frac{\sin(\varepsilon x)}{\varepsilon x}) \geq \varepsilon/2$ si $|\varepsilon x| \geq 2$, ce qui est vrai si $|x| \geq a$ car $\varepsilon a = 2$.
On a donc (comme $(\varepsilon - \frac{\sin(\varepsilon x)}{x}) \geq 0$ pour tout x)

$$T = \int_{\mathbb{R}} (\varepsilon - \frac{\sin(\varepsilon x)}{x}) dm(x) \geq \frac{\varepsilon}{2} m(\{x, |x| \geq a\}).$$

On obtient donc, finalement,

$$m(\{x, |x| \geq a\}) \leq \frac{2T}{\varepsilon} \leq \frac{2}{\varepsilon} \int_0^{\varepsilon} (1 - \psi(t)) dt = a \int_0^{2/a} (1 - \psi(t)) dt.$$

Ce qui est le résultat annoncé. ■

10.4 Transformation de Fourier dans L^2

On aimerait ici définir la transformée de Fourier d'un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N) = L^2_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. On rappelle que l'espace $L^2_{\mathbb{C}}(\mathbb{R}^N)$ est un espace de Hilbert et que le produit scalaire sur $L^2_{\mathbb{C}}(\mathbb{R}^N)$ est défini par (en notant $dt = d\lambda_N(t)$) :

$$(f \mid g)_2 = \int f(t)\overline{g(t)} dt.$$

Il est clair que la définition de $\widehat{f}$ qu'on a donnée pour $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ ne s'applique pas pour un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Pour définir la transformée de Fourier d'un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, on va utiliser la densité de $\mathcal{S}_N$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (on peut montrer que $\mathcal{S}_N \subset L^2_{\mathbb{C}}(\mathbb{R}^N)$, comme on a montré que $\mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$). On va d'abord remarquer que la transformée de Fourier envoie $\mathcal{S}_N$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et que c'est une isométrie pour la norme de $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On utilisera ensuite la densité de $\mathcal{S}_N$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ pour définir la transformée de Fourier des éléments de $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Proposition 10.16 *Soit* $N \geq 1$ *et* $f, g \in \mathcal{S}_N$ *(on a donc, en particulier,* $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$*). Alors* $\widehat{f}$ *et* $\widehat{g}$ *sont des éléments de* $L^2_{\mathbb{C}}(\mathbb{R}^N)$ *et* $(f \mid g)_2 = (\widehat{f} \mid \widehat{g})_2$*. En particulier,* $\|f\|_2 = \|\widehat{f}\|_2$.

Démonstration – Soit $f, g \in \mathcal{S}_N$. On a alors aussi $\widehat{f}, \widehat{g} \in \mathcal{S}_N$. Comme $\mathcal{S}_N \subset L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a donc $f, g, \widehat{f}, \widehat{g} \in L^2_{\mathbb{C}}(\mathbb{R}^N)$. On va montrer que $(f \mid g)_2 = (\widehat{f} \mid \widehat{g})_2$.

Comme $f, \widehat{f} \in \mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$, on peut appliquer le théorème d'inversion (théorème 10.6) pour transformer le produit scalaire de f avec g.

Le théorème d'inversion donne $f = \widehat{\widehat{f}}(-\cdot)$ et donc :

$$(f \mid g)_2 = \int \widehat{\widehat{f}}(-t)\bar{g}(t) dt = (2\pi)^{-\frac{N}{2}} \int \left(\int e^{ix\cdot t} \widehat{f}(x) dx \right) \bar{g}(t) dt.$$

On utilise maintenant le théorème de Fubini (théorème 7.12). Il s'applique car $\widehat{f}, \bar{g} \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. On obtient :

$$(f \mid g)_2 = (2\pi)^{-\frac{\mathrm{N}}{2}} \int \left(\int e^{ix\cdot t} \bar{g}(t) dt \right) \widehat{f}(x) dx = \int \bar{\widehat{g}}(t) \widehat{f}(t) dt = (\widehat{f} \mid \widehat{g})_2,$$

ce qui termine la démonstration. ∎

La proposition 10.16 permet de définir, par un argument de densité, la transformée de Fourier dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$.

Théorème 10.17 (Transformée de Fourier dans L^2, Plancherel) *Soit $\mathrm{N} \geq 1$. Il existe une application linéaire continue $\tilde{\mathcal{F}}$ de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ t.q. :*

1. *Si $f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}}) \cap \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, on a alors $\tilde{\mathcal{F}}(f) = \widehat{f}$ p.p..*
2. *Pour tout $f, g \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, on a $(f \mid g)_2 = (\tilde{\mathcal{F}}(f) \mid \tilde{\mathcal{F}}(g))_2$.*
3. *Pour tout $f \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, on a $f = \tilde{\mathcal{F}}(\tilde{\mathcal{F}}(f))(-\cdot)$.*
4. *$\tilde{\mathcal{F}}$ est une bijection de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$.*

Pour $f \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, $\tilde{\mathcal{F}}(f)$ s'appelle la transformée de Fourier de f. Compte tenu du premier item, on notera en général, $\widehat{f}$ la transformée de Fourier de f si $f \in \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ (et alors $\widehat{f} = \mathcal{F}(f) \in \mathrm{C}_0(\mathbb{R}^{\mathrm{N}}, \mathbb{C})$) ou si $f \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ (et alors $\widehat{f} = \tilde{\mathcal{F}}(f) \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$).

Démonstration – L'application $f \mapsto \widehat{f}$ est définie sur $\mathcal{S}_{\mathrm{N}}$, qui est un sous espace de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, et prend ses valeurs dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ (car $\mathcal{S}_{\mathrm{N}} \subset \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, en confondant, comme d'habitude, un élément de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ avec l'un de ses représentants). Comme cette application est linéaire, continue pour la norme de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, et que $\mathcal{S}_{\mathrm{N}}$ est dense dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ (car $\mathcal{S}_{\mathrm{N}} \supset \mathrm{C}^\infty_c(\mathbb{R}^{\mathrm{N}}, \mathbb{C})$ et $\mathrm{C}^\infty_c(\mathbb{R}^{\mathrm{N}}, \mathbb{C})$ est dense dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$, voir le théorème 8.14), on en déduit que cette application se prolonge (de manière unique) en une application, notée $\tilde{\mathcal{F}}$, linéaire continue de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$.

Plus précisément, soit $f \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. Il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_{\mathrm{N}}$ t.q. $f_n \to f$ dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ quand $n \to +\infty$. La suite $(f_n)_{n\in\mathbb{N}}$ est donc de Cauchy dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. La proposition 10.16 donne alors que la suite $(\widehat{f_n})_{n\in\mathbb{N}}$ est donc aussi de Cauchy dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. Elle converge donc dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. On aimerait définir $\tilde{\mathcal{F}}(f)$ comme étant la limite (dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$) de la suite $(\widehat{f_n})_{n\in\mathbb{N}}$. Ceci est possible à condition que cette limite ne dépende que de f et pas du choix de la suite $(f_n)_{n\in\mathbb{N}}$ convergeant dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ vers f. Or, ce dernier point est facile car si $(g_n)_{n\in\mathbb{N}}$ est une autre suite convergeant dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ vers f, on a $\|\widehat{f_n} - \widehat{g_n}\|_2 = \|f_n - g_n\|_2 \to 0$ quand $n \to +\infty$. On a ainsi défini $\tilde{\mathcal{F}}$ de $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ dans lui même.

La linéarité de $\tilde{\mathcal{F}}$ découle immédiatement du fait que l'application $f \mapsto \widehat{f}$ est linéaire de $\mathcal{S}_{\mathrm{N}}$ dans $\mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$. Enfin, soit $f \in \mathrm{L}^2_{\mathbb{C}}(\mathbb{R}^{\mathrm{N}})$ et $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_{\mathrm{N}}$ t.q. $f_n \to f$ dans

$L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.16 donne que $\|\widehat{f_n}\|_2 = \|f_n\|_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quand $n \to +\infty$, on en déduit $\|\tilde{\mathcal{F}}(f)\|_2 = \|f\|_2$. Ce qui prouve la continuité de $\tilde{\mathcal{F}}$ de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On montre maintenant les 4 items du théorème.

1. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}^N) \cap L^2_{\mathbb{C}}(\mathbb{R}^N)$. En reprenant la démonstration du théorème 8.14, il est facile de voir qu'il existe une suite $(f_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^N, \mathbb{C})$ t.q. $f_n \to f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $L^1_{\mathbb{C}}(\mathbb{R}^N)$ lorsque $n \to +\infty$ (car dans la démonstration du théorème 8.14, la suite construite pour converger vers f dans L^p ne dépend pas de p). On en déduit que $\widehat{f_n} \to \widehat{f}$ uniformément sur $\mathbb{R}^N$ lorsque $n \to +\infty$ (car $f_n \to f$ dans $L^1_{\mathbb{C}}(\mathbb{R}^N)$) et que $\widehat{f_n} \to \tilde{\mathcal{F}}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ lorsque $n \to +\infty$ (car $f_n \to f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) et donc que, après extraction éventuelle d'une sous-suite, on peut supposer que $\widehat{f_n} \to \tilde{\mathcal{F}}(f)$ p.p. quand $n \to +\infty$. On en déduit bien que $\widehat{f} = \tilde{\mathcal{F}}(f)$ p.p..
2. Soit $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$. Il existe deux suites $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_N$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \to f$, $g_n \to g$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.16 donne $(\widehat{f_n} \mid \widehat{g_n})_2 = (f_n \mid g_n)_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quand $n \to +\infty$ on obtient bien $(\tilde{\mathcal{F}}(f) \mid \tilde{\mathcal{F}}(g))_2 = (f \mid g)_2$.
3. Soit $f \in L^2$. Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \to f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \to +\infty$. On a donc $\widehat{f_n} \to \tilde{\mathcal{F}}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, ce qui donne aussi $\widehat{f_n}(-\cdot) \to \tilde{\mathcal{F}}(f)(-\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et donc $\widehat{\widehat{f_n}}(-\cdot) \to \tilde{\mathcal{F}}(\tilde{\mathcal{F}}(f))(-\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \to +\infty$. La proposition 10.11 donne $f_n = \widehat{\widehat{f_n}}(-\cdot)$ pour tout $n \in \mathbb{N}$. On en déduit donc (par unicité de la limite dans L^2) que $f = \tilde{\mathcal{F}}(\tilde{\mathcal{F}}(f))(-\cdot)$ p.p..
4. L'injectivité de $\tilde{\mathcal{F}}$ (de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) découle du fait que $\|\tilde{\mathcal{F}}(f)\|_2 = \|f\|_2$ et que $\tilde{\mathcal{F}}$ est linéaire. La surjectivité est une conséquence immédiate du troisième item.

■

Remarque 10.18

Voici deux remarques à propos de la transformée de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

1. Pour définir la transformée de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, on aurait pu utiliser la densité de $C_c^\infty(\mathbb{R}^N, \mathbb{C})$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et ne pas introduire l'espace $\mathcal{S}_N$ (voir l'exercice 10.5 pour le cas $N = 1$).
2. Si $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ mais $f \notin L^1_{\mathbb{C}}(\mathbb{R}^N)$. On ne peut pas utiliser la formule de la définition 10.3 pour calculer $\tilde{\mathcal{F}}(f)$, mais on peut utiliser cette formule avec une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions appartenant à $L^1_{\mathbb{C}}(\mathbb{R}^N)$ et convergeant vers f dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Par exemple, on peut prendre $f_n = f 1_{B_n}$, où B_n est la boule de $\mathbb{R}^N$ de centre 0 et de rayon n. On a alors $\hat{f}_n \to \tilde{\mathcal{F}}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, quand $n \to +\infty$, et $\hat{f}_n$ se calcule avec la formule de la définition 10.3.

10.5 Résolution d'une équation différentielle ordinaire ou d'une équation aux dérivées partielles

On donne ici deux exemples simples d'utilisation de la transformation de Fourier pour la résolution d'une équation différentielle (souvent notée EDO pour Équation Différentielle Ordinaire) ou d'une équation aux dérivées partielles (souvent notée EDP pour Équation aux Dérivées Partielles).

Soit $\mathrm{N} \geq 0$, $(a_0, \ldots, a_\mathrm{N}) \in \mathbb{R}^{\mathrm{N}+1}$ et $g \in \mathcal{S}_1$ (donnés). On cherche $f \in \mathcal{S}_1$ qui vérifie :

$$a_\mathrm{N} f^{(\mathrm{N})}(x) + \ldots + a_1 f'(x) + a_0 f(x) = g(x), \text{ pour tout } x \in \mathbb{R}. \tag{10.3}$$

Si $f \in \mathcal{S}_1$ vérifie (10.3), elle vérifie nécessairement, par transformation de Fourier :

$$a_\mathrm{N} \widehat{f^{(\mathrm{N})}}(t) + \ldots + a_1 \widehat{f'}(t) + a_0 \widehat{f}(t) = \widehat{g}(t), \text{ pour tout } t \in \mathbb{R}.$$

c'est-à-dire :

$$a_\mathrm{N} (it)^\mathrm{N} \widehat{f}(t) + \ldots + a_1 it \widehat{f}(t) + a_0 \widehat{f}(t) = \widehat{g}(t), \text{ pour tout } t \in \mathbb{R}.$$

En posant $p(t) = a_\mathrm{N}(it)^\mathrm{N} + \ldots + a_1 it + a_0$ et en supposant que p ne s'annule pas sur $\mathbb{R}$, on a alors :

$$\widehat{f}(t) = \frac{\widehat{g}(t)}{p(t)}.$$

Comme $g \in \mathcal{S}_1$ et que p ne s'annule par sur $\mathbb{R}$, on peut montrer que $\frac{\widehat{g}}{p} \in \mathcal{S}_1$. En utilisant maintenant le théorème d'inversion, ou la proposition 10.11, on obtient :

$$f(t) = \widehat{\widehat{f}}(-t) = \widehat{\left(\frac{\widehat{g}}{p}\right)}(-t), \text{ pour tout } t \in \mathbb{R}. \tag{10.4}$$

On a donc montré que f est nécessairement donnée par (10.4). Réciproquement, il est facile de voir que la fonction donnée par (10.4) est solution de (10.3), c'est donc l'unique solution dans $\mathcal{S}_1$ de (10.3) (en supposant que p ne s'annule par sur $\mathbb{R}$).

Soit $\mathrm{N} \geq 1$, on considère une fonction $u : \mathbb{R}^\mathrm{N} \to \mathbb{R}$, de classe C^2 (c'est-à-dire deux fois continûment dérivable), dont on peut donc définir le Laplacien $\Delta u = \sum_{i=1}^{\mathrm{N}} \frac{\partial^2 u}{\partial x_i^2}$.
On cherche à déterminer les fonctions harmoniques de $\mathcal{S}_\mathrm{N}$, c'est-à-dire les fonctions $u \in \mathcal{S}_\mathrm{N}$ telles que :

$$-\Delta u(x) = 0, \text{ pour tout } x \in \mathbb{R}^\mathrm{N}. \tag{10.5}$$

Cherchons $u \in \mathcal{S}_\mathrm{N}$ (et donc $\Delta u \in \mathcal{S}_\mathrm{N}$) solution de (10.5). On a donc $\widehat{\Delta u} = 0$ (partout sur $\mathbb{R}^\mathrm{N}$), c'est-à-dire $|t|^2 \widehat{u}(t) = 0$ et donc $\widehat{u}(t) = 0$, pour tout $t \neq 0$. Comme $\widehat{u}$ est continue, ceci entraîne que $\widehat{u} = 0$ (partout sur $\mathbb{R}^N$), et donc $u = 0$. La fonction identiquement égale à 0 est donc la seule solution de (10.5) dans $\mathcal{S}_N$.

On peut effectuer un raisonnement analogue si on cherche u de classe C^2 et dans $L^2_{\mathbb{R}}(\mathbb{R}^N)$ (on peut aussi le faire si u est seulement dans $L^2_{\mathbb{R}}(\mathbb{R}^N)$, il faut alors convenablement définir Δu). On obtient encore que la seule solution de (10.5) est $u = 0$.

Par contre, ce résultat d'unicité n'est plus vrai si on ne demande pas à u d'être de carré intégrable. En effet, les fonctions constantes sont toutes solutions de (10.5) (et on peut montrer que ce sont les seules fonctions, de classe C^2et bornées, solutions de (10.5), c'est le théorème de Liouville en analyse complexe).

10.6 Fonction caractéristique d'un vecteur aléatoire

La transformée de Fourier a son équivalent en probabilités : il s'agit de la fonction caractéristique d'une variable aléatoire (rien à voir avec la fonction caractéristique d'un ensemble !).

Définition 10.19 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d. *On appelle fonction caractéristique de* X *la fonction de* $\mathbb{R}^d$ *dans* $\mathbb{C}$ *définie par :*

$$\varphi_X(u) = \int_\Omega e^{iX\cdot u} dP = E(e^{iX\cdot u}), \text{ pour } u \in \mathbb{R}^d.$$

Dans la définition précédente, la fonction $e^{iX\cdot u}$ est bien intégrable sur Ω (pour tout $u \in \mathbb{R}^d$) car la fonction $x \mapsto e^{ix\cdot u}$ est continue bornée de $\mathbb{R}^d$ dans $\mathbb{C}$. En notant p_X la loi de X, on remarque également que $\varphi_X = (2\pi)^{d/2}\widehat{p_X}(-\cdot)$. La section 10.3 donne alors quelques propriétés élémentaires de la fonction caractéristique d'un v..a..

Proposition 10.20 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* X *un v.a. de dimension* d. *La fonction caractéristique de* X *vérifie alors les propriétés suivantes :*

1. $\varphi_X \in C_b(\mathbb{R}^d, \mathbb{C})$.
2. *La loi de* X *est entièrement déterminée par* φ_X *(c'est-à-dire que si* Y *est un autre v.a. de dimension* d *et que* $\varphi_X = \varphi_Y$, *on a nécessairement* $p_X = p_Y$*).*
3. *Si* $\varphi_X \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$, *la loi de* X *a une densité par rapport à la mesure de Lebesgue sur* $\mathcal{B}(\mathbb{R}^d)$, *c'est-à-dire* $p_X = f\lambda_d$, *et :*

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix\cdot u}\varphi_X(u)du, \;\; \lambda_d\text{-p.s. en } x \in \mathbb{R}^d.$$

DÉMONSTRATION – Comme $\varphi_X = (2\pi)^{d/2}\widehat{p_X}(-\cdot)$, le premier item est donnée par la proposition 10.13 (qui donne $\widehat{p_X} \in C_b(\mathbb{R}^d, \mathbb{C})$).

Le deuxième item est donnée par le premier item de la proposition 10.14.

Pour le troisième item, on remarque que le deuxième item de la proposition 10.14 donne $p_X = f\lambda_d$ avec $f = \widehat{\widehat{p_X}}(-\cdot)$, ce qui donne λ_d-p.s. en $x \in \mathbb{R}^d$:

$$f(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{ix\cdot t}\widehat{p_X}(t)dt = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix\cdot t}\varphi_X(t)dt.$$

■

Les fonctions caractéristiques peuvent être utilisées pour montrer l'indépendance de v.a.r. (ou de v.a.). On donne un exemple dans la proposition suivante.

Proposition 10.21 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d > 1$ *et* $X_1, \ldots, X_d$ *d v.a.r.. Alors, les v.a.r* $X_1, \ldots, X_d$ *sont indépendantes si et seulement si on a*

$$\varphi_X(u) = \prod_{j=1}^{d} \varphi_{X_j}(u_j)$$

pour tout $u = (u_1, \ldots, u_d)^t \in \mathbb{R}^d$*, où* X *est le v.a. de composantes* $X_1, \ldots, X_d$*.*

DÉMONSTRATION – D'après le théorème 9.28 les v.a.r. $X_1, \ldots, X_d$ sont indépendantes si et seulement $p_X = p_{X_1} \otimes \ldots \otimes p_{X_d}$. Par la proposition 10.14, ces deux mesures sont égales si et seulement si leurs transformées de Fourier sont égales, c'est-à-dire si et seulement si :

$$\varphi_X(u) = \int_{\mathbb{R}^d} e^{ix\cdot u} d(p_{X_1} \otimes \ldots \otimes p_{X_d}) \text{ pour tout } u \in \mathbb{R}^d.$$

Comme $e^{ix\cdot u} = \prod_{j=1}^{d} e^{ix_j u_j}$ pour tout $x = (x_1, \ldots, x_d)^t$ et tout $u = (u_1, \ldots, u_d)^t$, la définition de la mesure produit donne

$$\int_{\mathbb{R}^d} e^{ix\cdot u} d(p_{X_1} \otimes \ldots \otimes p_{X_d}) = \prod_{j=1}^{d} \varphi_{X_j}(u_j),$$

ce qui termine la démonstration de cette proposition. ■

Il est intéressant aussi de connaître le lien entre convergence en loi et convergence simple des fonctions caractéristiques. Nous donnons ce lien dans le deuxième item du théorème 10.22 (dû à Paul Lévy).

Théorème 10.22 (Convergence en loi et fonctions caractéristiques, Paul Levy) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* $(X_n)_{n\in\mathbb{N}^*}$ *un suite de v.a. de dimension* d*.*

1. *On suppose que la fonction caractéristique de X_n converge simplement (quand $n \to +\infty$) vers une fonction φ continue en 0. Il existe alors un v.a. X t.q. $X_n \to X$ en loi, quand $n \to +\infty$.*
2. *Soit X un v.a. de dimension d. Alors, $X_n \to X$ en loi, quand $n \to +\infty$, si et seulement si $\varphi_{X_n}(u) \to \varphi_X(u)$, quand $n \to +\infty$, pour tout $u \in \mathbb{R}^d$.*

DÉMONSTRATION – Le deuxième item du théorème est une conséquence facile du premier item. En effet, si $X_n \to X$ en loi, on a, bien sûr, $\varphi_{X_n}(u) \to \varphi_X(u)$ pour tout $u \in \mathbb{R}^d$ (car la fonction $x \mapsto e^{ix\cdot u}$ est continue et bornée de $\mathbb{R}$ dans $\mathbb{C}$, pour tout $u \in \mathbb{R}^d$). Réciproquement, on suppose que $\varphi_{X_n}(u) \to \varphi_X(u)$ pour tout $u \in \mathbb{R}^d$. Comme la fonction φ_X est continue (et donc, en particulier continue en 0), le premier item donne que X_n converge en loi vers un v.a. Y. Mais ceci donne que X et Y ont même fonction caractéristique et donc même loi. On en déduit bien que X_n tend vers X en loi.

Nous démontrons maintenant le premier item du théorème 10.22. Cette démonstration se fait en deux étapes. Dans la première étape on montre que la suite des lois des X_n est tendue (on utilise ici le lemme 10.15 et la continuité de φ). Puis dans la seconde étape on conclut grâce à la proposition 9.21.

Etape 1. On montre dans cette étape que la suite des lois des X_n est tendue. Pour cela, il suffit de montrer que la suite des lois de chaque composante des X_n est tendue. Soit donc $i \in \{1,\ldots,d\}$ et $(X_n^{(i)})_{n\in\mathbb{N}}$ la suite des i–ème composantes des X_n. On note m_n la loi de $X_n^{(i)}$ et, pour tout $t \in \mathbb{R}$,

$$\psi_n(t) = \frac{1}{2}\varphi_{X_N^{(i)}}(t) + \varphi_{X_N^{(i)}}(-t) = \sqrt{2\pi}\frac{\widehat{m}_n(t) + \widehat{m}_n(-t)}{2}$$

Le lemme 10.15 donne alors, pour tout $a > 0$ et $n \in \mathbb{N}$.

$$m_n(\{x, |x| \geq a\}) \leq a\int_0^{2/a}(1 - \psi_n(t))dt. \tag{10.6}$$

On a $\varphi_{X_N^{(i)}}(t) = \varphi_{X_n}(te_i)$ (ou e_i est le i-ème vecteur de la base canonique de $\mathbb{R}^d$), la fonction $\psi_n(t)$ converge donc pour tout $t \in \mathbb{R}$ et sa limite est une fonction ψ (de $\mathbb{R}$ dans $\mathbb{R}$) continue en 0. Comme $\psi_n(0) = 1$ pour tout $n \in \mathbb{N}$, on a aussi $\psi(0) = 1$.

Soit $\varepsilon > 0$, comme ψ est continue en 0 et $\psi(0) = 1$, il existe $a_0 > 0$ t.q. $\max_{t\in[0,2/a_0]}(1-\psi(t)) \leq \varepsilon$, et donc

$$a_0\int_0^{2/a_0}(1 - \psi(t))dt \leq 2\varepsilon.$$

La fonction ψ_n converge simplement vers ψ et $0 \leq (1 - \psi_n) \leq 1$ le théorème de convergence dominée donne donc

$$\lim_{n\to+\infty} a_0\int_0^{2/a_0}(1 - \psi_n(t))dt = a_0\int_0^{2/a_0}(1 - \psi(t))dt.$$

Il existe donc n_0 t.q.

$$n \geq n_0 \Rightarrow \lim_{n\to+\infty} a_0 \int_0^{2/a_0} (1-\psi_n(t))dt \leq 3\varepsilon.$$

Avec (10.6) on a donc

$$n \geq n_0 \Rightarrow m_n(\{x, |x| \geq a_0\}) \leq 3\varepsilon.$$

D'autre part, en utilisant la continuité décroissante d'une mesure, pour tout $n \in \{1, \ldots, n_0 - 1\}$, il existe a_n t.q. $m_n(\{x, |x| \geq a_n\}) \leq 3\varepsilon$.
En prenant $a = \max\{a_0, a_1, \ldots, a_{n_0-1}\}$ on a donc

$$m_n(\{x, |x| \geq a\}) \leq 3\varepsilon \text{ pour tout } n \in \mathbb{N}^*.$$

Ceci montre bien que la suite $(m_n)_{n\in\mathbb{N}^*}$ est tendue et termine la première étape.

Etape 2. Dans cette seconde étape il suffit d'utiliser la proposition 9.21. Cette proposition nous donne l'existence d'une sous-suite de la suite $(X_n)_{n\in\mathbb{N}^*}$ et d'un v.a. X tel que cette sous-suite tende vers X en loi. On en déduit, en particulier que $\varphi_X = \varphi$. Il reste à montrer que toute la suite $(X_n)_{n\in\mathbb{N}^*}$ tend vers X en loi (et pas seulement une sous-suite). Pour le montrer, on raisonne par l'absurde. Si la suite $(X_n)_{n\in\mathbb{N}^*}$ ne converge pas en loi vers X, il existe $\phi \in C_b(\mathbb{R}^d, \mathbb{R})$ t.q. $E(\phi(X_n)) \not\to E(\phi(X))$. Il existe donc $\varepsilon > 0$ et une sous-suite de la suite $(X_n)_{n\in\mathbb{N}^*}$, que nous noterons aussi $(X_n)_{n\in\mathbb{N}^*}$ (pour ne pas alourdir les notations), t.q.

$$|E(\phi(X_n)) - E(\phi(X))| \geq \varepsilon. \tag{10.7}$$

Mais, la proposition 9.21 nous donne l'existence d'une sous-suite de cette sous-suite et d'un v.a. Y t.q. la sous-suite de la sous-suite tende vers Y en loi. Comme la convergence en loi donne la convergence simple des fonctions caractéristiques (et que $\varphi_{X_n} \to \varphi$ simplement), on en déduit que $\varphi_Y = \varphi = \varphi_X$. On a donc $P_X = P_Y$. La sous-suite de la sous-suite tend donc vers X en loi. En contradiction avec (10.7). On a bien ainsi montré finalement que $X_n \to X$ en loi. ■

On donne maintenant la fonction caractéristique d'un vecteur gaussien.

Proposition 10.23 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé.*

1. *Soit* X *une v.a.r. gaussienne. On note* m *son espérance* ($m = E(X) \in \mathbb{R}$) *et* σ^2 *sa variance* ($\sigma^2 = E((X-m)^2) \geq 0$) *de sorte que* $X \sim \mathcal{N}(m, \sigma^2)$. *On a alors :*

$$\varphi_X(u) = e^{imu} e^{-\frac{\sigma^2}{2}u^2}, \text{ pour tout } u \in \mathbb{R}.$$

2. *Soit* $d \geq 1$ *et* X *une v.a gaussien de dimension* d. *On note* m *son espérance* ($m = E(X) \in \mathbb{R}^d$) *et* D *sa matrice de covariance* (D *est une matrice symétrique, de taille* $d \times d$*, semi-définie positive, son terme à la ligne* j *et la colonne* k *est donné par la covariance des composantes d'indices* j *et* k *de* X) *On a alors :*

$$\varphi_X(u) = e^{im\cdot u} e^{-\frac{1}{2}Du\cdot u}, \text{ pour tout } u \in \mathbb{R}^d.$$

Ceci montre que la loi de X *ne dépend de* m *et* D *(de sorte que* $X \sim \mathcal{N}(m, D)$ *si* D *est s.d.p., voir la proposition 9.33).*

DÉMONSTRATION – Soit X une v.a.r. gaussienne, $X \sim \mathcal{N}(m, \sigma^2)$. On suppose tout d'abord que $\sigma^2 = 0$, on a alors $X = m$ p.s. et donc, pour tout $u \in \mathbb{R}$, $\varphi_X(u) = e^{imu}$, ce qui est bien la formule annoncée. On suppose maintenant que $\sigma > 0$, on a alors, pour tout $u \in \mathbb{R}$:

$$\varphi_X(u) = \int_{\mathbb{R}} e^{ixu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx.$$

Avec le changement de variable $y = \frac{x-m}{\sigma}$, on obtient :

$$\varphi_X(u) = e^{imu} \int_{\mathbb{R}} e^{iy\sigma u} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy,$$

ce qui donne $\varphi_X(u) = e^{imu} \frac{1}{\sqrt{2\pi}} \psi(\sigma u)$ avec $\psi(t) = \int_{\mathbb{R}} e^{iyt} e^{\frac{-y^2}{2}} dy$ pour tout $t \in \mathbb{R}$. Comme la fonction $y \mapsto e^{y^2}$ est paire, on a aussi :

$$\psi(t) = \int_{\mathbb{R}} \cos(yt) e^{\frac{-y^2}{2}} dy, \text{ pour tout } t \in \mathbb{R}.$$

Pour calculer $\psi(t)$, on remarque que le théorème de dérivabilité sous le signe intégrale s'applique (théorème 4.53) et donne que ψ est de classe C^1 et

$$\psi'(t) = \int_{\mathbb{R}} (-y) \sin(yt) e^{\frac{-y^2}{2}} dy.$$

En intégrant par parties cette dernière intégrale (en fait, on intègre par parties sur $[-n, n]$ puis on fait tendre n vers l'infini), on obtient :

$$\psi'(t) = -\int_{\mathbb{R}} t \cos(yt) e^{\frac{-y^2}{2}} dy = -t\psi(t), \text{ pour tout } t \in \mathbb{R},$$

ce qui donne $\psi(t) = \psi(0) e^{\frac{-t^2}{2}}$. Comme $\psi(0) = \int_{\mathbb{R}} e^{\frac{-y^2}{2}} dy = \sqrt{2\pi}$, on en déduit que $\psi(t) = \sqrt{2\pi} e^{\frac{-t^2}{2}}$ (pour tout $t \in \mathbb{R}$) et donc que $\varphi_X(u) = e^{imu} e^{\frac{-\sigma^2 u^2}{2}}$, ce qui est bien la formule annoncée.

On suppose maintenant que $d > 1$ et que X est un v.a. gaussien. Soit $u \in \mathbb{R}^d$. On a $\varphi_X(u) = E(e^{iX\cdot u}) = \varphi_{X\cdot u}(1)$. Pour connaître $\varphi_X(u)$, il suffit donc de connaître la fonction caractéristique de la v.a.r. $X \cdot u$. Comme X est un v.a. gaussien, la v.a.r. $X \cdot u$ est une v.a.r. gaussienne. Sa moyenne est $E(X \cdot u) = E(X) \cdot u = m \cdot u$ et sa variance est (voir l'exercice 9.6) :

$$\sigma^2 = E((X \cdot u - m \cdot u)^2) = E(u^t (X-m)(X-m)^t u) = u^t \mathrm{Cov}(X) u = u^t D u.$$

On a donc, d'après la première partie de cette démonstration (c'est-à-dire le cas $d = 1$),

$$\varphi_X(u) = \varphi_{X\cdot u}(1) = e^{im\cdot u} e^{-\frac{u^t D u}{2}},$$

ce qui est bien la formule annoncée (car $u^t D u = D u \cdot u$).

Enfin, comme φ_X détermine complètement la loi de X, la loi de X ne dépend que de m et D.

■

On montre enfin le théorème central limite, qui nous dit, en particulier, qu'une somme de v.a.r.i.i.d. se répartit de manière "gaussienne" autour de sa moyenne.

Théorème 10.24 (Théorème central limite) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $d \geq 1$ *et* $(X_n)_{n\in\mathbb{N}^*}$ *une suite de v.a. de dimension d i.i.d.. On suppose que* $E(|X_1|^2) < \infty$. *On pose* $E(X_1) = m$, $D = Cov(X_1)$ *et, pour* $n \in \mathbb{N}^*$,

$$Y_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - m).$$

La suite $(Y_n)_{n\in\mathbb{N}^*}$ *converge alors en loi vers tout v.a.* Y *dont la loi est la loi normale multidimensionnelle* $\mathcal{N}(0, D)$. *(Voir la définition 9.5 et la proposition 9.33 pour la définition de cette loi normale multidimensionnelle.)*

Remarque 10.25 *Dans le cas* $d = 1$, *le théorème 10.24 donne donc la convergence en loi de la suite* $(Y_n)_{n\in\mathbb{N}^*}$ *vers* Y *où* $Y \sim \mathcal{N}(0, \sigma^2)$, σ^2 *étant la variance de* X_1. *Si* $\sigma^2 \neq 0$, *Il est intéressant de noter que cette convergence en loi est équivalente à la convergence simple de la fonction de répartition de* Y_n *vers celle de* Y *(quand* $n \to +\infty$*). Ceci est dû au fait que la fonction de répartition de* Y *est continue en tout point* $\mathbb{R}$, *voir l'exercice 6.66.*

DÉMONSTRATION – D'après le théorème 10.22 et la proposition 10.23, il suffit de montrer que, pour tout $u \in \mathbb{R}^d$,

$$\lim_{n\to+\infty} \varphi_{Y_n}(u) = e^{-\frac{u^t D u}{2}},$$

où φ_{Y_n} est la fonction caractéristique du v.a. Y_n.

Soit $u \in \mathbb{R}^d$. On a

$$\varphi_{Y_n}(u) = E(e^{iY_n\cdot u}) = E(e^{\frac{i}{\sqrt{n}}\sum_{p=1}^{n}(X_p - m)\cdot u}) = E(\prod_{p=1}^{n} e^{\frac{i}{\sqrt{n}}(X_p - m)\cdot u}).$$

Les v.a. X_p $(p = 1, \ldots, n)$ étant indépendants et identiquement distribués, on a donc, par la proposition 9.27 (cette proposition est écrite pour des fonctions à valeurs réelles mais elle s'étend à des fonctions à valeurs complexes en décomposant les fonctions en parties réelles et imaginaires),

$$\varphi_{Y_n}(u) = E(e^{\frac{i}{\sqrt{n}}(X_1 - m)\cdot u})^n. \tag{10.8}$$

On pose $z_n = E(e^{\frac{i}{\sqrt{n}}(X_1 - m)\cdot u})$ et on calcule maintenant z_n en faisant un développement limité des fonctions cos et sin. Il existe deux fonctions ε_1 et ε_2 appartenant à $C_b(\mathbb{R}, \mathbb{R})$ t.q. $\lim_{x\to 0} \varepsilon_i(x) = 0$, $i = 1, 2$ et

$$\cos(x) = 1 - \frac{x^2}{2} + x^2\varepsilon_1(x) \text{ et } \sin(x) = x + x^2\varepsilon_2(x).$$

On en déduit que

$$\int_\Omega \cos(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP = 1-\frac{1}{2n}\int_\Omega ((X_1-m)\cdot u)^2 dP$$
$$+\frac{1}{n}\int_\Omega ((X_1-m)\cdot u)^2\varepsilon_1(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP,$$

et

$$\int_\Omega \sin(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP = \frac{1}{\sqrt{n}}\int_\Omega (X_1-m)\cdot u dP$$
$$+\frac{1}{n}\int_\Omega ((X_1-m)\cdot u)^2\varepsilon_2(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP.$$

Comme $E(|X_1|^2) < +\infty$, le théorème de convergence dominée donne

$$\lim_{n\to+\infty}\int_\Omega ((X_1-m)\cdot u)^2\varepsilon_i(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP = 0 \text{ pour } i=1,2.$$

D'autre part, on a

$$\int_\Omega ((X_1-m)\cdot u)^2 dP = \int_\Omega u^t(X_1-m)(X_1-m)^t u dP$$
$$= u^t\Big(\int_\Omega (X_1-m)(X_1-m)^t dP\Big)u = u^t D u.$$

et

$$\int_\Omega (X_1-m)\cdot u dP = \Big(\int_\Omega (X_1-m)dP\Big)\cdot u = 0.$$

En posant $a = \dfrac{u^t D u}{2}$, on a donc

$$\int_\Omega \cos(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP = 1-\frac{b_n}{n},\ \int_\Omega \sin(\frac{1}{\sqrt{n}}(X_1-m)\cdot u)dP = \frac{c_n}{n},$$

avec $\lim_{n\to+\infty} b_n = a$ et $\lim_{n\to+\infty} c_n = 0$. En posant $a_n = b_n + ic_n$, on a donc $z_n = (1-\dfrac{a_n}{n})$ avec $\lim_{n\to+\infty} a_n = a$. Avec (10.8), ceci donne

$$\varphi_{Y_n}(u) = z_n^n = (1-\frac{a_n}{n})^n.$$

Comme $a\in\mathbb{R}$, le lemme 10.26 donne le résultat, c'est-à-dire $\lim_{n\to+\infty}\varphi_{Y_n}(u) = e^{-a} = e^{-\frac{u^tDu}{2}}$. ■

Lemme 10.26

Soit $a\in\mathbb{R}$ et $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes t.q. $\lim_{n\to+\infty} a_n = a$. On a alors

$$\lim_{n\to+\infty}(1-\frac{a_n}{n})^n = e^{-a}.$$

DÉMONSTRATION – Ce lemme est bien connu si $a_n\in\mathbb{R}$ pour tout $n\in\mathbb{N}$. Il suffit de remarquer que $a_n < n$ pour n assez grand et que $\lim_{n\to+\infty} n\ln(1-\frac{a_n}{n}) = -\lim_{n\to+\infty} a_n = -a$. La difficulté est ici que a_n peut ne pas être un nombre réel.

On pose $z_n = 1 - \frac{a_n}{n}$ et $y_n = 1 - \frac{a}{n}$. Comme la suite $(a_n)_{n\in\mathbb{N}}$ est bornée (car elle est convergente), il existe $\mathrm{M} \in \mathbb{R}$ t.q. $|a_n| \le \mathrm{M}$ pour tout $n \in \mathbb{N}$. On a donc $|z_n| \le 1 + \frac{\mathrm{M}}{n}$ (et aussi $|y_n| \le 1+\frac{\mathrm{M}}{n}$). On a alors (en écrivant $\varphi(1)-\varphi(0) = \int_0^1 \varphi'(t)dt$ avec $\varphi \in C^\infty(\mathbb{R},\mathbb{C})$ définie par $\varphi(t) = (tz_n + (1-t)y_n)^n$)

$$z_n^n - y_n^n = n(z_n - y_n)\int_0^1 (tz_n + (1-t)y^n)^{n-1}dt.$$

Comme $|tz_n + (1-t)y_n| \le t|z_n| + (1-t)|y_n| \le 1 + \dfrac{\mathrm{M}}{n} \le e^{\frac{\mathrm{M}}{n}}$ (pour tout $n \in \mathbb{N}$), on en déduit que

$$|z_n^n - y_n^n| \le n|z_n - y_n|e^{\mathrm{M}} = |a - a_n|e^{\mathrm{M}}.$$

On a donc $\lim_{n\to+\infty} z_n^n = \lim_{n\to+\infty} y_n^n = \lim_{n\to+\infty}(1-\frac{a}{n})^n = e^{-a}$. ■

10.7 Exercices

10.7.1 Transformation de Fourier

Exercice 10.1 (Résultat partiel d'inversion de Fourier dans $L^1_{\mathbb{C}}$) Soit $\mathrm{H}(t) = e^{-|t|}, t \in \mathbb{R}$. On pose, pour $\lambda > 0$:

$$h_\lambda(x) = (2\pi)^{-\frac{1}{2}}\int_{\mathbb{R}} \mathrm{H}(\lambda t)e^{itx}dt, x \in \mathbb{R}.$$

1. Montrer que $h_\lambda(x) = (\dfrac{2}{\pi})^{\frac{1}{2}}\dfrac{\lambda}{\lambda^2 + x^2}$, et $\displaystyle\int_{\mathbb{R}} h_\lambda(x)dx = (2\pi)^{\frac{1}{2}}$.

Corrigé – *Soit $x \in \mathbb{R}$. Comme H est paire, on a $(2\pi)^{\frac{1}{2}}h_\lambda(x) = \int_{\mathbb{R}} e^{-|\lambda t|}\cos(tx)dt$ et donc*

$$(2\pi)^{\frac{1}{2}}h_\lambda(x) = 2\lim_{n\to+\infty}\int_0^n e^{-\lambda t}\cos(tx)dt.$$

En intégrant deux fois par parties, on remarque que

$$\int_0^n e^{-|\lambda t|}\cos(tx)dt = \frac{1}{\lambda} - (\frac{x}{\lambda})^2\int_0^n e^{-|\lambda t|}\cos(tx)dt + a_n,$$

avec $\lim_{n\to+\infty} a_n = 0$. Ceci donne

$$\int_0^n e^{-|\lambda t|}\cos(tx)dt = \frac{\lambda}{\lambda^2 + x^2}(1 + \lambda a_n),$$

et donc $h_\lambda(x) = (\frac{2}{\pi})^{\frac{1}{2}}\frac{\lambda}{\lambda^2+x^2}$.

Pour calculer $\int_{\mathbb{R}} h_\lambda(x)dx$, on utilise le changement de variable $x = \lambda y$, on obtient

$$\int_{\mathbb{R}} h_\lambda(x)dx = (\frac{2}{\pi})^{\frac{1}{2}}\int_{\mathbb{R}}\frac{1}{1+x^2}dy = (2\pi)^{\frac{1}{2}}.$$

2. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que pour tout $x \in \mathbb{R}$, on a :

$$f * h_\lambda(x) = \int_{\mathbb{R}} \mathrm{H}(\lambda t)\widehat{f}(t)e^{ixt}dt.$$

Corrigé – *Noter que λ désigne ici à la fois le paramètre λ introduit au début de l'énoncé et la mesure de Lebesgue. Cette maladresse de notation semble toutefois sans gravité.*

Comme $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $h_\lambda \in L^\infty_{\mathbb{R}}(\mathbb{R}, T, \lambda)$, $f * h_\lambda(x)$ est défini pour tout $x \in \mathbb{R}$ et on a :

$$f * h_\lambda(x) = \int_{\mathbb{R}} f(t)h_\lambda(x-t)dt = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} f(t)\left(\int_{\mathbb{R}} \mathrm{H}(\lambda y)e^{i(x-t)y}dy\right)dt.$$

Comme $f \in L^1_{\mathbb{C}}(\mathbb{R}, T, \lambda)$ et $\mathrm{H}(\lambda\cdot) \in L^\infty_{\mathbb{R}}(\mathbb{R}, T, \lambda)$, on peut utiliser le théorème de Fubini (théorème 7.12) pour inverser l'ordre d'intégration et obtenir :

$$\begin{aligned} f * h_\lambda(x) &= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \mathrm{H}(\lambda y)e^{ixy}\left(\int_{\mathbb{R}} f(t)e^{-ity}dt\right)dy \\ &= \int_{\mathbb{R}} \mathrm{H}(\lambda y)e^{ixy}\widehat{f}(y)dy. \end{aligned}$$

3. Soit g une fonction mesurable bornée de $\mathbb{R}$ dans $\mathbb{C}$, continue en 0. Montrer que $g * h_\lambda(0) \to \sqrt{2\pi}g(0)$ quand $\lambda \to 0$. [Utiliser 1. et le théorème de convergence dominée.]

Corrigé – *On utilise maintenant le fait que $h_\lambda \in L^1_{\mathbb{R}}(\mathbb{R}, T, \lambda)$ et $g \in L^\infty_{\mathbb{C}}(\mathbb{R}, T, \lambda)$ pour remarquer que $g * h_\lambda(x)$ est défini pour tout $x \in \mathbb{R}$. Pour $x = 0$, on a :*

$$g * h_\lambda(0) = \int_{\mathbb{R}} g(x)h_\lambda(x)dx = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_{\mathbb{R}} \frac{\lambda}{\lambda^2 + x^2} g(x)dx.$$

Avec le changement de variable $y = \frac{x}{\lambda}$, on obtient :

$$g * h_\lambda(0) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_{\mathbb{R}} g(\lambda y)\frac{1}{1+y^2}dy.$$

Comme $|g(\lambda y)\frac{1}{1+y^2}| \le \|g\|_u \frac{1}{1+y^2}$ (avec $\|g\|_u = \sup_{x\in\mathbb{R}} |g(x)|$) et que la fonction $y \mapsto \frac{1}{1+y^2}$ est intégrable sur $\mathbb{R}$, on peut utiliser le théorème de convergence dominée pour en déduire (grâce à la continuité de g en 0) :

$$\lim_{\lambda\to 0} g * h_\lambda(0) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} g(0) \int_{\mathbb{R}} \frac{1}{1+y^2}dy = (2\pi)^{\frac{1}{2}} g(0).$$

4. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, T, \lambda)$, montrer que :

$$\|f * h_\lambda - \sqrt{2\pi}f\|_1 \to 0 \text{ lorsque } \lambda \to 0.$$

[On pourra utiliser la continuité en moyenne et la question précédente avec $g(y) = \int |f(x+y) - f(x)|dx$.]

Corrigé – *Comme* $\int_{\mathbb{R}} h_\lambda(y)dy = \sqrt{2\pi}$, *on a :*

$$\|f * h_\lambda - \sqrt{2\pi} f\|_1 = \int_{\mathbb{R}} |\int_{\mathbb{R}} (f(x-y) - f(x))h_\lambda(y)dy|dx$$
$$\leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) - f(x)| h_\lambda(y)dy \right) dx.$$

En utilisant le théorème de Fubini-Tonelli (théorème 7.7), on en déduit :

$$\|f * h_\lambda - \sqrt{2\pi} f\|_1 \leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) - f(x)|dx \right) h_\lambda(y)dy = g * h_\lambda(0),$$

avec $g(y) = \int |f(x+y) - f(x)|dx$.

Le théorème de continuité en moyenne dans L^1 *(théorème 5.21, écrit pour de fonctions à valeurs réelles mais la généralisation est immédiate pour des fonctions à valeurs complexes) donne que* g *est continue en* 0 *et donc aussi continue de* $\mathbb{R}$ *dans* $\mathbb{R}$ *(en remarquant que* $|g(y) - g(z)| \leq |g(y-z)|$*) et donc aussi mesurable de* $\mathbb{R}$ *dans* $\mathbb{C}$. *Elle est également bornée (car* $|g(y)| \leq 2\|f\|_1$, *pour tout* $y \in \mathbb{R}$*). On peut donc utiliser la question précédente, elle donne que* $\lim_{\lambda\to 0} g * h_\lambda(0) = 0$ *et donc que* $\lim_{\lambda\to 0} \|f * h_\lambda - \sqrt{2\pi} f\|_1 = 0$.

5. Déduire de ce qui précède le théorème d'inversion de Fourier, théorème 10.6.

Corrigé – *On note* $\mathrm{L}^1 = \mathrm{L}^1_{\mathbb{C}}(\mathbb{R}, T, \lambda)$. *Soit* $f \in \mathrm{L}^1$ *(on a donc* $\widehat{f} \in C_0(\mathbb{R}, \mathbb{C})$*). On suppose que* $\widehat{f} \in \mathrm{L}^1$. *Soit* $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+^*$ *une suite t.q.* $\lim_{n\to+\infty} \lambda_n = 0$. *Comme* $f \in \mathrm{L}^1$, *la question précédente nous donne que* $f * h_{\lambda_n} \to \sqrt{2\pi} f$ *dans* L^1 *et la question 2 nous donne pour tout* $n \in \mathbb{N}$ *et tout* $x \in \mathbb{R}$ *:*

$$f * h_{\lambda_n}(x) = \int_{\mathbb{R}} \mathrm{H}(\lambda_n t) \widehat{f}(t) e^{ixt} dt.$$

On utilise maintenant le théorème de convergence dominée qui s'applique parce que $\widehat{f} \in \mathrm{L}^1$ *et que l'on a, pour tout* x *et tout* n, $|\mathrm{H}(\lambda_n t)\widehat{f}(t)e^{ixt}| \leq |\widehat{f}(x)|$. *Comme* $\lim_{n\to+\infty} \mathrm{H}(\lambda_n x) = 1$, *pour tout* $x \in \mathbb{R}$, *on a donc, pour tout* $x \in \mathbb{R}$ *:*

$$\lim_{n\to+\infty} f * h_{\lambda_n}(x) = \int_{\mathbb{R}} \widehat{f}(t) e^{ixt} dt = \sqrt{2\pi} \widehat{\widehat{f}}(-x).$$

Enfin, comme $f * h_{\lambda_n} \to \sqrt{2\pi} f$ *dans* L^1, *on peut supposer, après extraction d'une sous-suite, que* $f * h_{\lambda_n} \to \sqrt{2\pi} f$ *p.p.. On a donc, finalement (par unicité de la limite dans* $\mathbb{R}$*),* $\sqrt{2\pi} f = \sqrt{2\pi} \widehat{\widehat{f}}(-\cdot)$ *p.p., c'est-à-dire* $f = \widehat{\widehat{f}}(-\cdot)$ *p.p..*

Exercice 10.2 (La transformation de Fourier n'est ni stable ni surjective) Soit $a \in \mathbb{R}_+^*$. On note $\mathrm{L}^1_{\mathbb{C}}$ l'espace $\mathrm{L}^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. On pose $f = 1_{[-a,a]}$.

1. Calculer la transformée de Fourier de f. En déduire que $\mathrm{L}^1_{\mathbb{C}}$ n'est pas stable par transformation de Fourier.

Corrigé – *On a bien* $f \in \mathrm{L}^1_{\mathbb{C}}$. *Soit* $t \in \mathbb{R}^*$, *on a*

$$\widehat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-ixt} dx = \frac{1}{-it\sqrt{2\pi}}(e^{-iat} - e^{iat}) = \frac{1}{t\sqrt{2\pi}} 2\sin(at) = \sqrt{\frac{2}{\pi}} \frac{\sin(at)}{t}.$$

Pour $t = 0$, on a $\widehat{f}(0) = \sqrt{\frac{2}{\pi}}a$. (On peut remarquer que $\widehat{f}$ est bien continue sur $\mathbb{R}$.)

La fonction $\widehat{f}$ n'est pas intégrable sur $\mathbb{R}$ (pour la mesure de Lebesgue), ce qui montre que $L^1_{\mathbb{C}}$ n'est pas stable par Fourier.

2. Pour $n \in \mathbb{N}^*$, on pose $g_n = 1_{[-n,n]}$. Calculer $f * g_n$, et montrer qu'il existe $h_n \in L^1_{\mathbb{C}}$ t.q. $\widehat{h_n} = f * g_n$. Montrer que la suite $f * g_n$ est bornée dans $C_0(\mathbb{R},\mathbb{R})$ alors que la suite h_n n'est pas bornée dans $L^1_{\mathbb{C}}$. En déduire que la transformée de Fourier n'est pas surjective de $L^1_{\mathbb{C}}$ dans $C_0(\mathbb{R},\mathbb{C})$.

Corrigé – *Soit $n \in \mathbb{N}^*$. La fonction $f * g_n$ est définie sur tout $\mathbb{R}$ et pour $x \in \mathbb{R}$ on a*

$$f * g_n(x) = \int_{\mathbb{R}} f(y)g_n(x-y)dy = \lambda([-a,a] \cap [x-n, x+n]), \tag{10.9}$$

car $f(y)g_n(x-y) = 1$ si $y \in [-a,a] \cap [x-n,x+n]$ et 0 sinon.

On peut eventuellement expliciter $f * g_n$ (mais ce n'est pas nécessaire pour la suite). Pour $n > a$, on obtient :

$$f(x) = \begin{cases} 0 & \text{si } x \geq a+n, \\ a-x+n & \text{si } -a+n \leq x < a+n \\ 2a & \text{si } a-n \leq x < -a+n \\ x+n+a & \text{si } -n-a \leq x < a-n \\ 0 & \text{si } x < -n-a. \end{cases}$$

Comme f et g_n appartiennent à $L^1_{\mathbb{C}}$, on sait (par le cours) que $f * g_n \in L^1_{\mathbb{C}}$ et $\widehat{f * g_n} = \sqrt{2\pi}\widehat{f}\widehat{g_n}$. On obtient donc, pour tout $t \in \mathbb{R}$ (en posant $\sin(t)/t = 1$ pour $t = 0$),

$$\widehat{f * g_n}(t) = h_n(t), \text{ avec } h_n(t) = 2\sqrt{\frac{2}{\pi}}\frac{\sin(at)\sin(nt)}{t^2}.$$

La fonction h_n est continue sur $\mathbb{R}$ et $|h_n(t)| \leq \sqrt{8/\pi}(1/t^2)$ pour tout t, on en déduit que $h_n \in L^1_{\mathbb{C}}$, on peut donc appliquer le théorème d'inversion de Fourier. Il donne

$$f * g_n(-\cdot) = \widehat{\widehat{f * g_n}} = \widehat{h_n}.$$

Comme $f * g_n(-\cdot) = f * g_n$ (par (10.9)), on a bien finalement $f * g_n = \widehat{h_n}$ avec $h_n \in L^1_{\mathbb{C}}$.

Comme f et g_n sont de carré intégrable, on sait que $f * g_n \in C_0(\mathbb{R},\mathbb{R})$) (voir l'exercice 8.7). Ceci peut aussi se voir directement sur la formule pour $f * g_n$. La formule (10.9) donne aussi $|f * g_n(x)| \leq 2a$ pour tout x et tout n. La suite $(f * g_n)_{n\in\mathbb{N}}$ est donc bornée dans $C_0(\mathbb{R},\mathbb{R})$ (et aussi dans $C_0(\mathbb{R},\mathbb{C})$). On montre maintenant que la suite $(h_n)_{n\in\mathbb{N}}$ n'est pas bornée dans $L^1_{\mathbb{C}}$.

Soit $\alpha > 0$ t.q. $\sin(at)/(at) \geq 1/2$ pour tout $t \in [0,\alpha]$. On a alors pour tout $n \in \mathbb{N}^*$

$$\sqrt{\frac{\pi}{8}}\int_{\mathbb{R}} |h_n(t)|dt \geq \frac{a}{2}\int_0^{\alpha} |\frac{\sin(nt)}{t}|dt = \frac{a}{2}\int_0^{n\alpha} |\frac{\sin(z)}{z}|dz,$$

Ce qui prouve que $(h_n)_{n\in\mathbb{N}}$ n'est pas bornée dans $L^1_{\mathbb{C}}$ car la fonction $z \mapsto \sin(z)/z$ n'est pas intégrable sur $\mathbb{R}$.

Si l'application $\varphi \mapsto \widehat{\varphi}$ *était surjective de* L^1_C *dans* $C_0(\mathbb{R},\mathbb{C})$*, elle serait alors bijective et continue (car on sait déjà qu'elle est injective et continue). Le théorème de Banach* [1] *donnerait alors que son inverse est aussi continue, ce qui est faux car la suite* $(\widehat{h_n})_{n\in\mathbb{N}}$ *est bornée dans* $C_0(\mathbb{R},\mathbb{C})$ *alors que la suite* $(h_n)_{n\in\mathbb{N}}$ *est non bornée dans* $L^1_{\mathbb{C}}$.

Exercice 10.3 (Une condition donnant $\widehat{f} \in L^1$)

On note L^p l'espace $L^p_{\mathbb{C}}(\mathbb{R}, T, \lambda)$.

1. Soient $f, g \in L^1$, montrer que $f\widehat{g} \in L^1$, $g\widehat{f} \in L^1$ et $\int f\widehat{g}d\lambda = \int g\widehat{f}d\lambda$.
2. Soit $B = [-\frac{1}{2}, \frac{1}{2}]$. Montrer que $1_B * 1_B(t) = (1-|t|)^+$ pour tout $t \in \mathbb{R}$.
3. On pose $\theta_n(t) = (1 - \frac{|t|}{n})^+$ pour tout $n \in \mathbb{N}^*$ et pour tout $t \in \mathbb{R}$.
   (a) Calculer $\widehat{\theta_1}(y)$ pour tout $y \in \mathbb{R}$.
   [On pourra remarquer que $\theta_1 = 1_B * 1_B$ avec $B = [-\frac{1}{2}, \frac{1}{2}]$.]
   (b) Soit $n \in \mathbb{N}^*$ et $y \in \mathbb{R}^\star$, montrer que

$$\widehat{\theta_n}(y) = \frac{4}{\sqrt{2\pi}} \frac{\sin^2(\frac{ny}{2})}{ny^2}.$$

4. Soit $f \in L^1 \cap L^\infty$ t.q. $\widehat{f}(t) \in \mathbb{R}_+$ pour tout $t \in \mathbb{R}$. On se propose de montrer que $\widehat{f} \in L^1$ (et donc que le théorème d'inversion s'applique). On utilise la fonction θ_n de la question précédente.
   (a) On note $\varphi_n = \theta_n \widehat{f}$; montrer que $\varphi_n \uparrow \widehat{f}$ et $\int \varphi_n d\lambda \uparrow \int \widehat{f} d\lambda$ lorsque $n \to +\infty$.
   (b) Montrer qu'il existe $\alpha \geq 0$ indépendant de n tel que $\int \widehat{\theta_n}(y)dy = \alpha$ pour tout $n \in \mathbb{N}^*$. En déduire que $\widehat{f} \in L^1$.

Exercice 10.4 (Transformée de Fourier inverse) Soit $f \in L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. On suppose que la transformée de Fourier de f, notée $\widehat{f}$, est aussi intégrable (pour la mesure de Lebesgue).

1. En utilisant le fait que f prend ses valeurs dans $\mathbb{R}$, montrer que

$$\begin{aligned} &\mathrm{Re}(\widehat{f}(-t)) = \mathrm{Re}(\widehat{f}(t)) \text{ pour tout } t \in \mathbb{R},\\ &\mathrm{Im}(\widehat{f}(-t)) = -\mathrm{Im}(\widehat{f}(t)) \text{ pour tout } t \in \mathbb{R}, \end{aligned}$$

   où $\mathrm{Re}(\xi)$ et $\mathrm{Im}(\xi)$ désignent les parties réelle et imaginaire du nombre complexe ξ.

1. Théorème de Banach : Si T est une application linéaire bijective et continue d'un espace de Banach E dans un espace de Banach F, alors son inverse est continue.

Corrigé – *Pour tout $t \in \mathbb{R}$, on a* $\widehat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ixt} dx$.

Soit $t \in \mathbb{R}$, on a donc

$$\widehat{f}(-t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{ixt} dx.$$

En utilisant le fait que f prend ses valeurs dans $\mathbb{R}$, on obtient alors la formule suivante pour le conjugué de $\widehat{f}(-t)$ (noté $\overline{\widehat{f}(-t)}$) :

$$\overline{\widehat{f}(-t)} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \overline{f(x)}\,\overline{e^{ixt}} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ixt} dx = \widehat{f}(t).$$

En reprenant encore le conjugué, ceci donne bien $\widehat{f}(-t) = \overline{\widehat{f}(t)}$.

2. Montrer que

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \mathrm{Re}(\widehat{f}(t) e^{itx}) dt \text{ pour presque tout } x \in \mathbb{R}.$$

Corrigé – *Comme $\widehat{f}$ est intégrable, on sait $f = \widehat{\widehat{f}}(-\cdot)$ p.p. (théorème 10.1). On a donc, pour presque tout $x \in \mathbb{R}$,*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(t) e^{ixt} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \widehat{f}(t) e^{ixt} dt + \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \widehat{f}(t) e^{ixt} dt.$$

Dans l'intégrale de $-\infty$ à 0, on utilise le changement de variable $s = -t$, on obtient

$$\sqrt{2\pi} f(x) = \int_0^{+\infty} \widehat{f}(-s) e^{-ixs} ds + \int_0^{+\infty} \widehat{f}(t) e^{ixt} dt.$$

En regroupant les deux intégrales, on en déduit

$$\sqrt{2\pi} f(x) = \int_0^{+\infty} \Big(\widehat{f}(-t) e^{-ixt} + \widehat{f}(t) e^{ixt}\Big) dt.$$

La première question montre que le nombre complexe $\widehat{f}(-t)$ est le conjugué de $\widehat{f}(t)$. Comme le nombre complexe e^{-ixt} est le conjugué de e^{ixt} on en déduit que $\widehat{f}(-t)e^{-ixt}$ est le conjugué de $\widehat{f}(t)e^{ixt}$. Ceci donne alors

$$\sqrt{2\pi} f(x) = \int_0^{+\infty} 2Re(\widehat{f}(t) e^{ixt}) dt.$$

On obtient bien, finalement, pour presque tout $x \in \mathbb{R}$,

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} Re(\widehat{f}(t) e^{itx}) dt.$$

Exercice 10.5 (Transformée de Fourier pour f de classe C^2, à support compact)
On note $\mathcal{L}^p_{\mathbb{C}}$ l'espace $\mathcal{L}^p_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Soit $f \in C^2_c(\mathbb{R}, \mathbb{C})$ (c'est-à-dire que f est une fonction de $\mathbb{R}$ dans $\mathbb{C}$, de classe C^2, et qu'il existe K compact de $\mathbb{R}$ t.q. $f = 0$ sur K^c).

1. Montrer que $f \in \mathcal{L}^1_{\mathbb{C}} \cap \mathcal{L}^2_{\mathbb{C}}$.

2. Montrer que $\hat{f} \in \mathcal{L}^1_{\mathbb{C}} \cap \mathcal{L}^2_{\mathbb{C}}$. [On pourra utiliser la proposition 10.7.]

3. Montrer que $\|f\|_2 = \|\hat{f}\|_2$.

N.B. Comme $C^2_c(\mathbb{R},\mathbb{C})$ est dense dans $L^2_{\mathbb{C}}(\mathbb{R},\mathcal{B}(\mathbb{R}),\lambda)$, cet exercice permet de définir la transformée de Fourier dans $L^2_{\mathbb{C}}$ sans utiliser l'espace $\mathcal{S}_1$.

Exercice 10.6 (Caractérisation de m par $\widehat{m}$) Soit $d \geq 1$.

1. Soit m et μ deux mesures signées sur les boréliens de $\mathbb{R}^d$. On suppose que $\widehat{m} = \widehat{\mu}$.

(a) Soit $\varphi \in L^1_{\mathbb{C}}(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\lambda_d)$. Montrer que $\int \widehat{\varphi} dm = \int \widehat{\varphi} d\mu$.

Corrigé – *On remarque que* $\int \widehat{\varphi} dm = (2\pi)^{-\frac{d}{2}} \int \left(\int e^{-ix\cdot t} \varphi(x) dx \right) dm(t)$. *(Les intégrales sont toutes sur* $\mathbb{R}^d$*). La mesure signée* m *peut se décomposer en différence de deux mesures positives étrangères* $m = m^+ - m^-$ *(décomposition de Hahn, proposition 2.33). Comme*

$$\int \int |e^{-ix\cdot t}\varphi(x)| dx dm^{\pm}(t) = \|\varphi\|_1 m^{\pm}(\mathbb{R}^d) < +\infty,$$

on peut utiliser le théorème de Funini-Tonelli (théorème 7.7) avec les mesures λ_d *et* m^+ *et les mesures* λ_d *et* m^-*. On obtient ainsi :*

$$\int \widehat{\varphi} dm = (2\pi)^{-\frac{d}{2}} \int \left(\int e^{-ix\cdot t} dm(t) \right) \varphi(x) dx = \int \widehat{m}(x)\varphi(x) dx.$$

Le même raisonnement donne $\int \widehat{\varphi} d\mu = \int \widehat{\mu}(x)\varphi(x)dx$*. Comme* $\widehat{m}(x) = \widehat{\mu}(x)$ *pour tout* $x \in \mathbb{R}^d$*, on en déduit bien* $\int \widehat{\varphi} dm = \int \widehat{\varphi} d\mu$*.*

(b) Montrer que $\int \varphi dm = \int \varphi d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d,\mathbb{C})$ (et donc pour tout $\varphi \in C^\infty_c(\mathbb{R}^d,\mathbb{R})$).

Corrigé –
Comme $\mathcal{S}(\mathbb{R}^d,\mathbb{C}) \subset L^1_{\mathbb{C}}(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\lambda_d)$*, la question précédente donne* $\int \widehat{\varphi} dm = \int \widehat{\varphi} d\mu$ *pour tout* $\varphi \in \mathcal{S}(\mathbb{R}^d,\mathbb{C})$*. Or, l'application* $\varphi \mapsto \widehat{\varphi}$ *est une bijection dans* $\mathcal{S}(\mathbb{R}^d,\mathbb{C})$ *(proposition 10.11). On a donc* $\int \varphi dm = \int \varphi d\mu$ *pour tout* $\varphi \in \mathcal{S}(\mathbb{R}^d,\mathbb{C})$*.*

(c) Montrer que $m = \mu$ (On rappelle qu'une fonction de $\varphi \in C_c(\mathbb{R}^d,\mathbb{R})$ est limite uniforme de fonctions de $\varphi \in C^\infty_c(\mathbb{R}^d,\mathbb{R})$).

Corrigé – *Soit* $\varphi \in C_c(\mathbb{R}^d,\mathbb{R})$ *et* $(\varphi_n)_{n\in\mathbb{N}} \subset C^\infty_c(\mathbb{R}^d,\mathbb{R})$ *t.q.* $\varphi_n \to \varphi$*, uniformément sur* $\mathbb{R}^d$*, quand* $n \to +\infty$*. La question précédente donne* $\int \varphi_n dm = \int \varphi_n d\mu$ *pour tout* $n \in \mathbb{N}$*. On utilise alors le théorème de convergence dominée (ce qui est possible car les mesures* $m^\pm$ *et* $\mu^\pm$ *sont des mesures finies), il donne* $\int \varphi dm = \int \varphi d\mu$*.*

La proposition 5.8 donne alors $m = \mu$*.*

2. Soit m une mesure signée sur les boréliens de $\mathbb{R}^d$. On suppose que $\widehat{m} \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que m est la mesure de densité f par rapport à la mesure de Lebesgue avec $f = \widehat{\widehat{m}}(-\cdot)$.

Exercice 10.7 (Transformée de Fourier du produit de fonctions)

Pour $p \in [1, +\infty]$, On note L^p l'espace $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et on désigne par $\|\cdot\|_p$ la norme dans L^p.

On note $\mathrm{C}_c^\infty(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions de classe C^∞ et à support compact de $\mathbb{R}$ dans $\mathbb{C}$. Enfin, on note $\mathrm{C}_b(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions continues bornées de $\mathbb{R}$ dans $\mathbb{C}$.

Si $f \in \mathrm{L}^1$, on désigne par $\widehat{f}$ la transformée de f (on a donc $\widehat{f} \in \mathrm{C}_b(\mathbb{R}, \mathbb{C})$).

Si $f \in \mathrm{L}^2$, on désigne par $\tilde{\mathcal{F}}(f)$ la transformée de Fourier de f (on a donc $\tilde{\mathcal{F}}(f) \in \mathrm{L}^2$).

On rappelle que si $f \in \mathrm{L}^1 \cap \mathrm{L}^2$, on a $\widehat{f} = \mathrm{F}(f)$ p.p.. Dans ce cas, on confond, en général, $\widehat{f}$ et $\tilde{\mathcal{F}}(f)$.

1. (Convolution $\mathrm{L}^2 - \mathrm{L}^2$). Soit $u, v \in \mathrm{L}^2$. Montrer que pour tout $t \in \mathbb{R}$ la fonction $s \mapsto u(t-s)v(s)$ est intégrable et donc que la fonction $u * v$ est définie sur tout $\mathbb{R}$ par la formule

$$u * v(t) = \int_{\mathbb{R}} u(t-s)v(s)ds, \text{ pour tout } t \in \mathbb{R}.$$

Montrer que $u * v \in \mathrm{C}_b(\mathbb{R}, \mathbb{C})$ et que $\|u * v\|_\infty \le \|u\|_2 \|v\|_2$.

Corrigé – *On peut choisir pour u et v des représentants et donc considérer que $u, v \in \mathcal{L}^2_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. La question est alors corrigée dans l'exercice 8.7 (voir la première question de cet exercice en prenant $p = q = 2$). Dans l'exercice 8.7, les fonctions sont à valeurs réelles mais la démonstration est identique pour des fonctions à valeurs complexes.*

2. Soit $f, g \in \mathrm{C}_c^\infty(\mathbb{R}, \mathbb{C})$.

(a) Montrer que f, g, $fg \in \mathrm{L}^1 \cap \mathrm{L}^2$, puis que $\widehat{f}, \widehat{g} \in \mathrm{L}^1 \cap \mathrm{L}^2$.

Corrigé – *Il suffit de remarquer que f et g appartiennent à l'espace de Schwartz, noté $\mathcal{S}_1$ (voir la definition 10.9). Comme $\mathcal{S}_1 \subset \mathrm{L}^p$ pour tout $1 \le p \le +\infty$, on a donc $f, g \in \mathrm{L}^1 \cap \mathrm{L}^2$ (avec la confusion habituelle entre un élément de L^p et son représentant continu lorsqu'il existe). Puis, comme la transformée de Fourier envoie $\mathcal{S}_1$ dans $\mathcal{S}_1$ (voir la proposition 10.11), on a aussi $\widehat{f}, \widehat{g} \in \mathcal{S}_1$ et donc $\widehat{f}, \widehat{g} \in \mathrm{L}^1 \cap \mathrm{L}^2$.*

(b) Montrer que, pour tout $t \in \mathbb{R}$,

$$\widehat{fg}(t) = \frac{1}{\sqrt{2\pi}} \widehat{f} * \widehat{g}(t).$$

[On pourra, par exemple, utiliser le fait que f, $\widehat{g} \in \mathrm{L}^1$ et calculer $\widehat{f} * \widehat{g}(t)$ en utilisant la définition de $\widehat{f}$ et la transformée de Fourier inverse pour $\widehat{g}$.]

Corrigé – *Soit $t \in \mathbb{R}$. comme $\widehat{f}, \widehat{g} \in \mathrm{L}^2$, la fonction $\widehat{f} * \widehat{g}$ est bien définie au point t et on a*

$$\widehat{f} * \widehat{g}(t) = \int_{\mathbb{R}} \widehat{f}(t-x)\widehat{g}(x)dx.$$

Comme $f \in \mathrm{L}^1$, on peut utiliser la formule définissant $\widehat{f}(t-x)$. Elle donne

$$\widehat{f} * \widehat{g}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} e^{-i(t-x)y} f(y)dy \Big)\widehat{g}(x)dx.$$

On peut maintenant utiliser le théorème de Fubini car $|e^{-i(t-x)y} f(y)\widehat{g}(x)| = |f(y)\widehat{g}(x)|$ (et $\int_{\mathbb{R}}\int_{\mathbb{R}} |f(y)\widehat{g}(x)|dxdy = \|f\|_1\|\widehat{g}\|_1 < +\infty$). On obtient

$$\widehat{f} * \widehat{g}(t) = \int_{\mathbb{R}} \Big(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixy}\widehat{g}(x)dx \Big) e^{-ity} f(y)dy.$$

Comme $g, \widehat{g} \in \mathrm{L}^1$, le théorème d'inversion de Fourier (théorème 10.6) donne $g(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixy}\widehat{g}(x)dx$ et donc

$$\widehat{f} * \widehat{g}(t) = \int_{\mathbb{R}} e^{-ity} f(y)g(y)dy,$$

c'est-à-dire (noter que $fg \in \mathrm{L}^1$) $\widehat{f} * \widehat{g}(t) = \sqrt{2\pi}\widehat{fg}(t)$.

3. Soit $f, g \in \mathrm{L}^2$. Montrer que, pour tout $t \in \mathbb{R}$,

$$\widehat{fg}(t) = \frac{1}{\sqrt{2\pi}} \tilde{\mathcal{F}}(f) * \tilde{\mathcal{F}}(g)(t).$$

Corrigé – *Il suffit d'utiliser la densité de $\mathrm{C}_c^\infty(\mathbb{R}, \mathbb{C})$ dans L^2. Soit $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ deux suites de $\mathrm{C}_c^\infty(\mathbb{R}, \mathbb{C})$ telles que $f_n \to f$ dans L^2 et $g_n \to g$ dans L^2, quand $n \to +\infty$. On a alors $f_n g_n \to fg$ dans L^1 (il suffit de remarquer que $\|f_n g_n - fg\|_1 \le \|f_n - f\|_2\|g_n\|_2 + \|f\|_2\|g_n - g\|_2$). La transformée de Fourier envoie continûment L^1 dans $\mathrm{C}_b(\mathbb{R}, \mathbb{C})$. On a donc $\widehat{f_n g_n} \to \widehat{fg}$ uniformément sur $\mathbb{R}$. La transformée de Fourier envoie aussi continûment L^2 dans L^2. On a donc $\widehat{f_n} \to \tilde{\mathcal{F}}(f)$ et $\widehat{g_n} \to \tilde{\mathcal{F}}(g)$ dans L^2. La première question donne alors que $\widehat{f_n} * \widehat{g_n} \to \tilde{\mathcal{F}}(f) * \tilde{\mathcal{F}}(g)$ uniformément sur $\mathbb{R}$. Soit maintenant $t \in \mathbb{R}$, la question précédente donne $\widehat{f_n g_n}(t) = \frac{1}{\sqrt{2\pi}}\widehat{f_n} * \widehat{g_n}(t)$. En passant à la limte quand $n \to +\infty$ on en déduit bien que $\widehat{fg}(t) = \frac{1}{\sqrt{2\pi}}\tilde{\mathcal{F}}(f) * \tilde{\mathcal{F}}(g)(t)$.*

Exercice 10.8 (Caractérisation des fonctions à valeurs réelles) Soit $d \ge 1$. Pour $1 \le p \le \infty$, on note L^p l'espace $\mathrm{L}^p_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$.

1. Soit $f \in \mathrm{L}^1$. Montrer que $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$ si et seulement si $\overline{\widehat{f}}(\xi) = \widehat{f}(-\xi)$ pour tout $\xi \in \mathbb{R}^d$.

Corrigé – *Pour tout $\xi \in \mathbb{R}$, on a $\widehat{f}(\xi) = (2\pi)^{-d/2} \int f(x)e^{-ix\cdot\xi}dx$. On a donc*

$$\overline{\widehat{f}}(\xi) = (2\pi)^{-d/2} \int \overline{f}(x)e^{+ix\cdot\xi}dx = \widehat{\overline{f}}(-\xi). \tag{10.10}$$

On en déduit que

$$\overline{\widehat{\overline{f}}}(\xi) - \widehat{f}(-\xi) = \widehat{\overline{f} - f}(-\xi). \tag{10.11}$$

Si $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$, on a $\overline{f} - f = 0$ p.p. et donc $\widehat{\overline{f} - f}(\xi) = 0$ pour tout $\xi \in \mathbb{R}^d$. Avec (10.11)*, on en déduit bien que $\overline{\widehat{f}}(\xi) = \widehat{f}(-\xi)$ pour tout $\xi \in \mathbb{R}^d$.*

Réciproquement, on suppose maintenant que $\overline{\widehat{f}}(\xi) = \widehat{f}(-\xi)$ pour tout $\xi \in \mathbb{R}^d$, on déduit alors de (10.11) *que $\widehat{\overline{f} - f}(\xi) = 0$ pour tout $\xi \in \mathbb{R}^d$, et donc par le théorème d'inversion (théorème 10.6) que $\overline{f} - f = 0$ p.p., c'est à dire que $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$.*

2. Soit $f \in L^2$. On désigne par $\tilde{\mathcal{F}}(f)$ la transformée de Fourier de f (on a donc $\tilde{\mathcal{F}}(f) \in L^2$). Montrer que $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$ si et seulement si $\overline{\tilde{\mathcal{F}}(f)}(\xi) = \tilde{\mathcal{F}}(f)(-\xi)$ pour presque tout $\xi \in \mathbb{R}^d$.

   Corrigé – *Il suffit de démontrer que l'égalité* (10.11) *est toujours vraie p.p. en remplaçant la transformée de Fourier dans L^1 par la transformée de Fourier dans L^2. Pour cela, on choisit une suite $(f_n)_{n\in\mathbb{N}}$ avec $f_n \in L^1 \cap L^2$ (pour tout n) et $f_n \to f$ dans L^2 quand $n \to +\infty$ (une telle suite existe, on peut même supposer que $f_n \in C_c^\infty(\mathbb{R}^N, \mathbb{C})$). On a donc aussi $\overline{f}_n \to \overline{f}$ dans L^2 et, en utilisant la définition de $\tilde{\mathcal{F}}$, on en déduit que*

   $$\widehat{f_n} \to \tilde{\mathcal{F}}(f) \text{ et } \widehat{\overline{f}_n} \to \tilde{\mathcal{F}}(\overline{f}) \text{ dans } L^2, \text{ quand } n \to +\infty.$$

   D'après (10.10)*, on a $\overline{\widehat{f_n}} = \widehat{\overline{f}_n}(-\cdot)$. Les deux termes de cette égalité ont une limite dans L^2 quand $n \to +\infty$. On obtient donc, quand $n \to +\infty$, $\overline{\tilde{\mathcal{F}}(f)} = \tilde{\mathcal{F}}(\overline{f})(-\cdot)$ p.p. et donc (par linéarité de $\tilde{\mathcal{F}}$)*

   $$\overline{\tilde{\mathcal{F}}(f)} - \tilde{\mathcal{F}}(f)(-\cdot) = \tilde{\mathcal{F}}(\overline{f} - f)(-\cdot) \text{ p.p.}. \tag{10.12}$$

   Pour conclure, on procède comme dans la question précédente. Si $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$, on a $\overline{f} - f = 0$ p.p. et donc $\tilde{\mathcal{F}}(\overline{f} - f) = 0$ p.p.. Avec (10.12)*, on en déduit bien que $\overline{\tilde{\mathcal{F}}(f)} = \tilde{\mathcal{F}}(f)(-\cdot)$ p.p..*

   Réciproquement, on suppose maintenant que $\overline{\tilde{\mathcal{F}}(f)} = \tilde{\mathcal{F}}(f)(-\cdot)$ p.p.. On déduit alors de (10.12) *que $\tilde{\mathcal{F}}(\overline{f} - f) = 0$ p.p., et donc par l'injectivité de $\tilde{\mathcal{F}}$ que $\overline{f} - f = 0$ p.p., c'est à dire que $f(x) \in \mathbb{R}$ pour presque tout $x \in \mathbb{R}^d$.*

Exercice 10.9 (Fourier, série et transformée)

On rappelle que l'espace $\mathcal{S}$ est défini par

$$\mathcal{S} = \{\varphi \in C^\infty(\mathbb{R}, \mathbb{C}) \text{ telle que pour tout } \alpha \in \mathbb{N} \text{ et tout } n \in \mathbb{N}, \sup_{x\in\mathbb{R}} |x|^\alpha |\varphi^{(n)}(x)| < +\infty\}.$$

Soit $\varphi \in \mathcal{S}$ (On rappelle que $\widehat{\varphi}$ est alors aussi dans l'espace $\mathcal{S}$ et que $\widehat{\widehat{\varphi}} = \varphi(-\cdot)$).

1. Soit $t \in \mathbb{R}$.
   Montrer que la série $\sum_{k\in\mathbb{Z}} \widehat{\varphi}(t+2k\pi)$ est absolument convergente (dans $\mathbb{C}$).
   Pour $t \in \mathbb{R}$, on définit $f(t)$ par $f(t) = \sum_{k\in\mathbb{Z}} \widehat{\varphi}(t+2k\pi)$.
2. Montrer que $f \in C^1(\mathbb{R}, \mathbb{C})$ et que f est 2π-périodique.
   On rappelle que ceci donne que f est somme de sa série de Fourier, c'est-à-dire que pour tout t dans $\mathbb{R}$ on a
   $$f(t) = \sum_{n\in\mathbb{Z}} c_n(f)e^{int},$$
   avec $c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx$.
3. Soit $n \in \mathbb{Z}$, montrer que
   $$c_n(f) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\varphi}(x)e^{-inx}dx = \frac{1}{\sqrt{2\pi}}\widehat{\widehat{\varphi}}(n) = \frac{1}{\sqrt{2\pi}}\varphi(-n).$$
   En déduire que pour tout t dans $\mathbb{R}$
   $$\sum_{k\in\mathbb{Z}} \widehat{\varphi}(t+2k\pi) = \frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}} \varphi(-n)e^{int}.$$

Exercice 10.10 ($f \in C_c^\infty$ et $\widehat{f} = 0$ sur un ouvert non vide implique $f = 0$)

Soit $f \in C_c^\infty(\mathbb{R}, \mathbb{C})$ et U un ouvert non vide de $\mathbb{R}$. On suppose que $\widehat{f} = 0$ sur U. Montrer que $f = 0$ sur tout $\mathbb{R}$. [On pourra commencer par montrer que l'application $z = \xi + i\eta \in \mathbb{C} \mapsto \int f(x)e^{-ix(\xi+i\eta)dx}$ est bien définie et dérivable sur $\mathbb{C}$.]

Exercice 10.11 (Théorème d'injection de Sobolev)

Soit $d \geq 1$. Pour $p \in [1, \infty]$, on note L^p l'espace l'espace $L^p_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Si $f \in L^2$, on note $\widehat{f}$ la transformée de Fourier de f.

Pour $s \in \mathbb{R}$, $s \geq 0$, on note $H^s(\mathbb{R}^d) = \{f \in L^2 \text{ t.q. } (1+|\cdot|^2)^{\frac{s}{2}}\widehat{f} \in L^2\}$ et $\|f\|_{H^s} = \|(1+|\cdot|^2)^{\frac{s}{2}}\widehat{f}\|_{L^2}$.

On rappelle que $C_0(\mathbb{R}^d, \mathbb{C})$ est un sous espace vectoriel fermé de $C_b(\mathbb{R}^d, \mathbb{C})$ muni de la norme $\|f\|_u = \sup_{x\in\mathbb{R}^d} |f(x)|$. Avec cette norme, $C_b(\mathbb{R}^d, \mathbb{C})$ et $C_0(\mathbb{R}^d, \mathbb{C})$ sont des espaces de Banach.

Soit $s > \frac{d}{2}$.

1. Soit $f \in H^s(\mathbb{R}^d)$. Montrer que $\widehat{f} \in L^1$. En déduire que $f \in C_0(\mathbb{R}^d, \mathbb{C})$ (au sens "il existe $g \in C_0(\mathbb{R}^2)$ telle que $f = g$ p.p."; on confond alors f et g).
2. Montrer que $H^s(\mathbb{R}^d) \subset C_0(\mathbb{R}^d, \mathbb{C})$ et qu'il existe C_s, ne dépendant que de s et d, t.q. :
   $$\|f\|_u \leq C_s\|f\|_{H^s} \text{ pour tout } f \in H^s(\mathbb{R}^d).$$

3. On s'intéresse maintenant au cas $d = 2$.

(a) Soit $s > 1$ et $u \in \mathrm{H}^s(\mathbb{R}^2)$. Montrer que

$$\|f\|_u \le \frac{1}{2\sqrt{\pi(s-1)}}\|f\|_{\mathrm{H}^s}.$$

(b) Soit $1 < s < 2$ et $f \in \mathrm{H}^2(\mathbb{R}^2)$. Montrer que

$$\|f\|_{\mathrm{H}^s} \le \|u\|_{\mathrm{H}^1}^{2-s}\|f\|_{\mathrm{H}^2}^{s-1}.$$

[Utiliser l'inégalité de Hölder.]

(c) On pose $\mathrm{C} = \sqrt{\frac{e}{2\pi}}$. Montrer que

$$f \in \mathrm{H}^2(\mathbb{R}^2),\ \|f\|_{\mathrm{H}^1} = 1 \ \Rightarrow\ \|f\|_u \le \mathrm{C}\sqrt{\ln(1+\|f\|_{\mathrm{H}^2})}.$$

[Pour $a > 1$, on pourra chercher le minimum pour $s \in]1,2[$ de la fonction $s \mapsto \frac{a^{s-1}}{\sqrt{s-1}}$, et distinguer selon les valeurs de a.]

Soit $\beta > 0$, montrer qu'il existe C_β, ne dépendant que de β, t.q. :

$$f \in \mathrm{H}^2(\mathbb{R}^2),\ \|f\|_{\mathrm{H}^1} \le \beta \ \Rightarrow\ \|f\|_u \le \mathrm{C}_\beta\sqrt{\ln(1+\|f\|_{\mathrm{H}^2})}.$$

Exercice 10.12 (Comportement à l'infini de f et régularité en 0 de $\hat{f}$)

On note $\mathcal{L}^1$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, intégrables pour la mesure de Lebesgue sur les boréliens de $\mathbb{R}$.

Pour $f \in \mathcal{L}^1$, on note $\hat{f}$ la transformée de Fourier de f.

Pour $a > 0$, on note $\mathrm{A}_a = \{x \in \mathbb{R};\ |x| \ge a\}$.

1. Soit $f \in \mathcal{L}^1$, $f \ge 0$ p.p.. Pour tout $t \in \mathbb{R}$, on pose $\psi(t) = \frac{\sqrt{2\pi}}{2}(\hat{f}(t) + \hat{f}(-t))$.

(a) Soit $t \in \mathbb{R}$. Montrer que $\psi(t) \in \mathbb{R}$ et $0 \le \psi(0) - \psi(t) = \displaystyle\int f(x)(1-\cos(xt))dx$.

Corrigé – *On a $\psi(t) = (1/2)\int f(x)(e^{ixt} + e^{-ixt})dx = \int f(x)\cos(xt)dx$. Comme f prend ses valeurs dans $\mathbb{R}$, on a donc $\psi(t) \in \mathbb{R}$. Puis, comme $\psi(0) = \int f(x)dx$. On a bien*

$$\psi(0) - \psi(t) = \int f(x)(1-\cos(xt))dx.$$

Comme $f(x)(1-\cos(xt)) \ge 0$ pour presque tout $x \in \mathbb{R}$, on a bien $\psi(0) - \psi(t) \ge 0$.

(b) Soit $a > 0$. Montrer que

$$\int_0^{2/a}\Big(\int f(x)(1-\cos(xt))dx\Big)dt = \int f(x)(\frac{2}{a}-\frac{1}{x}\sin(2\frac{x}{a}))dx.$$

Corrigé – *La fonction $(x,t)\mapsto f(x)(1-\cos(xt))$ est mesurable positive de $\mathbb{R}\times]0,2a[\to\mathbb{R}$, on peut donc appliquer le théorème de Fubini-Tonelli, il donne*

$$\int_0^{2/a}\Big(\int f(x)(1-\cos(xt))dx\Big)dt = \int\Big(\int_0^{2/a}(1-\cos(xt))dt\Big)f(x)dx$$
$$= \int\Big(\frac{2}{a}-\frac{\sin(2x/a)}{x}\Big)dx.$$

(c) Soit $a > 0$. Montrer que $\int_{A_a} f(x)dx \le \int f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx$. [On pourra dans l'intégrale de droite utiliser $\mathbb{R} = A_a \cup A_a^c$.]

En déduire (avec les deux questions précédentes) que

$$\int_{A_a} f(x)dx \le a\int_0^{2/a}(\psi(0)-\psi(t))dt.$$

Corrigé –

$$\int f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx$$
$$= \int_{A_a} f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx + \int_{A_a^c} f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx.$$

Comme $|\sin(y)|\le|y|$ pour tout $y\in\mathbb{R}$, on a $|\frac{a}{x}\sin(2\frac{x}{a}))|\le 2$ pour tout $x\in A_a^c$ (et même tout $x\in\mathbb{R}^\star$) et donc $\int_{A_a^c} f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx\ge 0$.

Si $x\in A_a$, on a $|\frac{a}{x}\sin(2\frac{x}{a}))|\le|\sin(2\frac{x}{a}))|\le 1$ et donc $\int_{A_a} f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx \ge \int_{A_a} f(x)dx$. On en déduit bien

$$\int f(x)(2-\frac{a}{x}\sin(2\frac{x}{a}))dx \ge \int_{A_a} f(x)dx.$$

Avec la question précédente, on a donc

$$\int_{A_a} f(x)dx \le a\int_0^{2/a}\Big(\int f(x)(1-\cos(xt))dx\Big)dt,$$

et la question 1(a) donne alors

$$\int_{A_a} f(x)dx \le a\int_0^{2/a}(\psi(0)-\psi(t))dt. \qquad (10.13)$$

(d) On suppose que $\hat{f}$ est de classe C^2. Montrer qu'il existe $C\in\mathbb{R}$, ne dépendant que de f, tel que $\int_{A_a} f(x)dx \le \frac{C}{a^2}$, pour tout $a > 0$.

Corrigé – *La fonction* ψ *est de classe* C^2. *Comme* $\psi(t) \leq \psi(0)$ *pour tout* t, *on a* $\psi'(0) = 0$. *On pose* $C_1 = \max\{|\psi''(t)|, t \in [-2, 2]\}$. *La formule de Taylor-Lagrange à l'ordre* 2 *donne alors, pour tout* $t \in [-2, 2]$,

$$\psi(0) - \psi(t) \leq C_1 \frac{t^2}{2}.$$

On en déduit, avec (10.13), *que, pour tout* a *tel que* $|a| \geq 1$, *on a*

$$\int_{A_a} f(x)dx \leq aC_1 \int_0^{2/a} \frac{t^2}{2} dt = \frac{4C_1}{3a^2}.$$

En posant $C = \max\{4C_1/3, \int f(x)dx\}$, *on a donc* $\int_{A_a} f(x)dx \leq \frac{C}{a^2}$, *pour tout* $a > 0$.

2. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathbb{R}$ dans $\mathbb{R}$, intégrables et positives p.p..
   On suppose que $\hat{f}_n$ converge simplement vers une fonction φ de $\mathbb{R}$ dans $\mathbb{R}$ continue en 0.
   Pour $n \in \mathbb{N}$ et $t \in \mathbb{R}$, on pose $\psi_n(t) = \frac{\sqrt{2\pi}}{2}(\hat{f}_n(t) + \hat{f}_n(-t))$.
   On note $\tilde{\psi}$ la limite simple des fonctions ψ_n.

(a) Monter que $\tilde{\psi}$ est continue en 0.

Corrigé – *On a* $\tilde{\psi}(t) = \frac{\sqrt{2\pi}}{2}(\varphi(t)+\varphi(-t))$ *pour tout* $t \in \mathbb{R}$. *Comme* φ *est continue en* 0, *on en déduit que* $\tilde{\psi}$ *est continue en* 0.

(b) Montrer que la suite $(\psi_n)_{n\in\mathbb{N}}$ est uniformément bornée sur $\mathbb{R}$.

Corrigé – *Pour tout* $t \in \mathbb{R}$ *et tout* $n \in \mathbb{N}$, *on a* $\psi_n(t) = \int f_n(x)\cos(xt)dx$ *et donc*

$$|\psi_n(t)| \leq \int f_n(x)dx = \psi_n(0).$$

La suite $(\psi_n(0))_{n\in\mathbb{N}}$ *est convergente (dans* $\mathbb{R}$*), elle donc bornée. On pose* $M = \sup_{n\in\mathbb{N}}\{\psi_n(0)\}$. *On a* $M \in \mathbb{R}$ *et* $|\psi_n(t)| \leq M$ *pour tout* $n \in \mathbb{N}$ *et* $t \in \mathbb{R}$.

(c) Montrer que, pour tout $a > 0$, $\lim_{n\to+\infty} \int_0^{2/a} (\psi_n(0) - \psi_n(t))dt = \int_0^{2/a} (\tilde{\psi}(0) - \tilde{\psi}(t))dt$.

Corrigé – *On a* $\lim_{n\to+\infty}(\psi_n(0) - \psi_n(t)) = (\tilde{\psi}(0) - \tilde{\psi}(t))$ *pour tout* $t \in [0, 2/a]$ *et* $|\psi_n(0) - \psi_n(t)| \leq 2M$ *pour tout* $t \in [0, 2/a]$ *et tout* $n \in \mathbb{N}$, *où* M *est définie à la question précédente. On peut donc appliquer le théorème de convergence dominée. Il donne bien* $\lim_{n\to+\infty} \int_0^{2/a} (\psi_n(0) - \psi_n(t))dt = \int_0^{2/a} (\tilde{\psi}(0) - \tilde{\psi}(t))dt$.

(d) Montrer que $\int_{A_a} f_n(x)dx \to 0$, quand $a \to +\infty$, uniformément par rapport à n.

Corrigé – *Pour tout $n \in \mathbb{N}$, on a $\int_{A_a} f_n(x)dx \to 0$, quand $a \to +\infty$ (en utilisant, par exemple, le théorème de convergence dominée). Le problème est donc de montrer que cette convergence est uniforme par rapport à n.*
Soit $\varepsilon > 0$. Comme cela vient d'être dit, pour tout $n \in \mathbb{N}$, il existe a_n tel que

$$a \geq a_n \Rightarrow \int_{A_a} f_n(x)dx \leq \varepsilon.$$

Mais, on ne peut pas conclure car on pourrait avoir $\sup_{n\in\mathbb{N}} a_n = +\infty$. Pour conclure on va utiliser la continuité de $\tilde{\psi}$ en 0. Comme $\tilde{\psi}$ est continue en 0, il existe $\bar{a} > 0$ tel que

$$\bar{a}\int_0^{2/\bar{a}} (\tilde{\psi}(0) - \tilde{\psi}(t))dt \leq \bar{a}\frac{2}{\bar{a}} \max\{|\tilde{\psi}(0) - \tilde{\psi}(t)|, t \in [0, \frac{2}{\bar{a}}]\} \leq 2\varepsilon.$$

la question précédente donne $\lim_{n\to+\infty} \bar{a}\int_0^{2/\bar{a}}(\psi_n(0) - \psi_n(t))dt = \bar{a}\int_0^{2/\bar{a}}(\tilde{\psi}(0) - \tilde{\psi}(t))dt \leq 2\varepsilon$. Il existe donc n_0 tel que

$$n \geq n_0 \Rightarrow \bar{a}\int_0^{2/\bar{a}} (\psi_n(0) - \psi_n(t))dt \leq 3\varepsilon.$$

La question 1(c) donne alors, pour $n \geq n_0$,

$$\int_{A_{\bar{a}}} f_n(x)dx \leq \bar{a}\int_0^{2/\bar{a}} (\psi_n(0) - \psi_n(t))dt \leq 3\varepsilon.$$

Il suffit maintenant de prendre $\tilde{a} = \max\{\bar{a}, b\}$ avec $b = \max_{n=0,\ldots,n_0} a_n$ et on obtient bien (en utilisant la positivité de f_n), pour tout $n \in \mathbb{N}$,

$$a \geq \tilde{a} \Rightarrow \int_{A_a} f_n(x)dx \leq 3\varepsilon.$$

10.7.2 Fonction Caractéristique d'une v.a.r.

Exercice 10.13 (Calcul de fonctions caractéristiques) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a. réelle. Calculer la fonction caractéristique φ_X de X dans les cas suivants :

1. $X = a$ p.s. ($a \in \mathbb{R}$).
2. $X \sim \mathcal{B}(p)$ (loi de Bernoulli de paramètre p : $P(\{X = 1\}) = p = 1 - P(\{X = 0\})$).
3. X suit une loi exponentielle de paramètre λ (avec $\lambda > 0$).

Exercice 10.14 (Vecteurs gaussiens et densité) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $d \geq 1$ et $X = (X_1, \ldots, X_d)$ un v.a. de dimension d.

On suppose que X est un vecteur gaussien (c'est-à-dire que $\sum_{i=1}^d a_i X_i$ suit une loi gaussienne pour tout $a_1, \ldots, a_d \in \mathbb{R}$). On note m la moyenne de X et D la matrice de covariance de X. Montrer que la loi de X est de densité par rapport à la mesure de Lebesgue (sur $\mathbb{R}^d$) si et seulement si D est inversible.

Exercice 10.15 (V.a. gaussiens, indépendance, covariance) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $d \geq 1$ et $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_d)$ un v.a. de dimension d.

1. On suppose ici que $d = 2$.

(a) On suppose que X_1 et X_2 suivent des lois gaussiennes et sont indépendantes, montrer que X est un vecteur gaussien et que $\mathrm{Cov}(\mathrm{X}_1, \mathrm{X}_2) = 0$.

Corrigé – *Comme X_1 et X_2 suivent des lois gaussiennes, il existe $m_1, m_2 \in \mathbb{R}$ et $\sigma_1, \sigma_2 \in \mathbb{R}_+$ t.q. $\mathrm{X}_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$. On a donc $\varphi_{\mathrm{X}_k}(u) = e^{ium_k} e^{-\frac{\sigma_k^2 u^2}{2}}$ pour $k = 1, 2$ et pour tout $u \in \mathbb{R}$.*

Soit $a_1, a_2 \in \mathbb{R}$. On calcule alors la fonction caractéristique de la v.a.r. $a_1\mathrm{X}_1 + a_2\mathrm{X}_2$. Soit $u \in \mathbb{R}$, on a :

$$\varphi_{a_1\mathrm{X}_1+a_2\mathrm{X}_2}(u) = \int_\Omega e^{i(a_1\mathrm{X}_1+a_2\mathrm{X}_2)u} d\mathrm{P} = \int_\Omega e^{ia_1\mathrm{X}_1 u} e^{ia_2\mathrm{X}_2 u} d\mathrm{P}$$

En utilisant l'indépendance de X_1 et X_2, on en déduit :

$$\varphi_{a_1\mathrm{X}_1+a_2\mathrm{X}_2}(u) = \int_\Omega e^{ia_1\mathrm{X}_1 u} d\mathrm{P} \int_\Omega e^{ia_2\mathrm{X}_2 u} d\mathrm{P} = \varphi_{\mathrm{X}_1}(a_1 u)\varphi_{\mathrm{X}_2}(a_2 u).$$

Ce qui donne $\varphi_{a_1\mathrm{X}_1+a_2\mathrm{X}_2}(u) = e^{iu(a_1m_1+a_2m_2)} e^{-\frac{u^2(\sigma_1^2 a_1^2+\sigma_2^2 a_2^2)}{2}}$. Comme la loi d'une v.a.r. est entièrement déterminée par sa fonction caractéristique (voir la proposition 10.20), on en déduit que $a_1\mathrm{X}_1 + a_2\mathrm{X}_2 \sim \mathcal{N}(m, \sigma^2)$ avec $m = a_1m_1 + a_2m_2$ et $\sigma = \sqrt{\sigma_1^2 a_1^2 + \sigma_2^2 a_2^2}$. Ceci prouve bien que X est un vecteur gaussien.

L'indépendance de X_1 et X_2 donne $Cov(\mathrm{X}_1, \mathrm{X}_2) = 0$. Le fait que X_1 et X_2 suivent des lois gaussiennes est inutile ici.

(b) On suppose que X est un vecteur gaussien et que $\mathrm{Cov}(\mathrm{X}_1, \mathrm{X}_2) = 0$. Montrer que X_1 et X_2 sont indépendantes.

Corrigé – *D'après la proposition 10.21, pour montrer que X_1 et X_2 sont indépendantes, il suffit de montrer que $\varphi_\mathrm{X}(u) = \prod_{j=1}^2 \varphi_{\mathrm{X}_j}(u_j)$ pour tout $u = (u_1, u_2)^t \in \mathbb{R}^2$. Ceci est une conséquence facile du fait que $Cov(\mathrm{X}_1, \mathrm{X}_2) = 0$. En effet, la matrice de covariance de X est alors diagonale et on a bien (grâce à la proposition 10.23), en reprenant les notations de la question précédente (c'est-à-dire $\mathrm{X}_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$),*

$$\varphi_\mathrm{X}(u) = e^{i(u_1m_1+u_2m_2)} e^{-\frac{u_1^2\sigma_1^2+u_2^2\sigma_2^2}{2}} = \varphi_{\mathrm{X}_1}(u_1)\varphi_{\mathrm{X}_2}(u_2),$$

c'est-à-dire $\varphi_\mathrm{X}(u) = \prod_{j=1}^2 \varphi_{\mathrm{X}_j}(u_j)$ pour tout $u = (u_1, u_2)^t \in \mathbb{R}^2$.

2. On suppose toujours $d = 2$. Donner un exemple pour lequel X_1 et X_2 sont gaussiennes mais X n'est pas un vecteur gaussien. [On pourra, par exemple, choisir $(\Omega, \mathcal{A}, \mathrm{P})$ et X_1, X_2 de manière à avoir $\mathrm{Cov}(\mathrm{X}_1, \mathrm{X}_2) = 0$ sans que X_1, X_2 soient indépendantes, voir l'exercice 4.48.]

Corrigé – *On considère les deux v.a.r.* S *et* X *de l'exercice 4.48 et on prend* $X_1 = SX$ *et* $X_2 = X$. *Les v.a.r.* X_1 *et* X_2 *sont gaussiennes (on a* $X_1 \sim \mathcal{N}(0,1)$ *et* $X_2 \sim \mathcal{N}(0,1)$*), elles sont dépendantes et on a* $Cov(X_1, X_2) = 0$ *(voir l'exercice 4.48). On en déduit que le v.a.* $X = (X_1, X_2)$ *n'est pas gaussien (sinon les v.a.r. seraient indépendantes, d'après la question précédente).*

3. On suppose que X est un vecteur gaussien et que les composantes de X sont indépendantes deux à deux. Montrer que $X_1, \ldots, X_d$ sont indépendantes.

Corrigé – *La démonstration est ici similaire à celle de la question 1(b). D'après la proposition 10.21, pour montrer que les v.a.r.* $X_1, \ldots, X_d$ *sont indépendantes, il suffit de montrer que* $\varphi_X(u) = \prod_{j=1}^d \varphi_{X_j}(u_j)$ *pour tout* $u = (u_1, \ldots, u_d)^t \in \mathbb{R}^d$. *Ceci est une conséquence facile du fait que les v.a.r.* $X_1, \ldots, X_d$ *sont indépendantes deux à deux. En effet, cette hypothèse d'indépendance deux à deux donne que la matrice de covariance de* X *est diagonale et on a alors (grâce à la proposition 10.23), avec* $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ *pour* $i = 1, \ldots, d$,

$$\varphi_X(u) = e^{i\sum_{k=1}^d u_k m_k} e^{-\frac{\sum_{k=1}^d u_k^2 \sigma_k^2}{2}} = \varphi_{X_1}(u_1) \ldots \varphi_{X_d}(u_d),$$

c'est-à-dire $\varphi_X(u) = \prod_{j=1}^d \varphi_{X_j}(u_j)$ *pour tout* $u = (u_1, \ldots, u_d)^t \in \mathbb{R}^d$.

Exercice 10.16 (Suite de v.a.r.i.i.d. de Poisson) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a. de Poisson de paramètre λ ($\lambda > 0$). On rappelle que, $P(\{X = k\}) = e^{-\lambda}\frac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$ et que $E(X) = \lambda$, $Var(X) = \lambda$.

1. Calculer la fonction caractéristique φ_X de X.

Corrigé – *Soit* $u \in \mathbb{R}$, *on a*

$$\varphi_X(u) = \int_\Omega e^{iXu} dP = \sum_{k \in \mathbb{N}} e^{iku} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k \in \mathbb{N}} \frac{(\lambda e^{iu})^k}{k!} = e^{\lambda(e^{iu}-1)}.$$

2. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. indépendantes et de Poisson de paramètre λ.

(a) Soit $n > 1$. Déduire de la question 1 la loi de la v.a. $Y_n = \sum_{p=1}^n X_p$.

Corrigé – *Soit* $u \in \mathbb{R}$. *On a* $\varphi_{Y_n}(u) = \int_\Omega e^{iY_n u} dP = \int_\Omega \prod_{p=1}^n e^{iX_p u} dP$. *Comme les v.a.r.* $X_1, \ldots, X_n$ *sont indépendantes, on en déduit que*

$$\varphi_{Y_n}(u) = \prod_{p=1}^n \int_\Omega e^{iX_p u} dP = e^{n\lambda(e^{iu}-1)}.$$

Comme la loi d'une v.a.r. est entièrement déterminée par sa fonction caractéristique (proposition 10.20), on en déduit que Y_n *est une v.a. de Poisson de paramètre* $n\lambda$.

(b) Utiliser le théorème central limite pour démontrer que

$$e^{-n} \sum_{k=0}^n \frac{n^k}{k!} \to \frac{1}{2} \text{ quand } n \to +\infty.$$

Corrigé – *On suppose maintenant que* $\lambda = 1$. *la suite* $(X_n)_{n\in\mathbb{N}^*}$ *est une suite de v.a.r.i.i.d. de carrés intégrables et on a* $E(X_1) = 1$ *et* $Var(X_1) = 1$. *Pour* $n \in \mathbb{N}^*$, *on pose*

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - 1) = \frac{1}{\sqrt{n}} (Y_n - n).$$

Le théorème central limite (théorème 10.24) donne que la suite $(P_{Z_n})_{n\in\mathbb{N}^*}$ *converge étroitement vers la loi normale* $\mathcal{N}(0,1)$. *Comme une loi normale est diffuse (c'est-à-dire qu'elle ne charge pas les points), on en déduit que* $P(Z_n \le 0)$ *tend vers* $1/2$ *quand* $n \to +\infty$ *(voir l'exercice 6.65 et noter que* $P(Z_n \le 0) = P_{Z_n}(]-\infty, 0])$*). Or,* $P(Z_n \le 0) = P(Y_n \le n)$ *et, comme* Y_n *est une v.a. de Poisson de paramètre n, on a :*

$$P(Y_n \le n) = \sum_{k=0}^{n} P(Y_n = k) = \sum_{k=0}^{n} e^{-n} \frac{n^k}{k!}.$$

On a donc $e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \to 1/2$, *quand* $n \to +\infty$.

Exercice 10.17 (Sur les lois des grands nombres) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a.r. dont la loi est la loi de Cauchy c'est-à-dire que $P_X = f\lambda$, avec $f(x) = \frac{1}{\pi(1+x^2)}$ pour $x \in \mathbb{R}$ (on rappelle que λ désigne la mesure de Lebesgue sur les boréliens de $\mathbb{R}$).

1. Montrer que X n'est pas une v.a.r. intégrable.

Corrigé – *On a*

$$\int_\Omega |X| dP = \int_{\mathbb{R}} \frac{|x|}{\pi(1+x^2)} dx \ge \frac{1}{2\pi} \int_1^{+\infty} \frac{1}{x} dx = +\infty.$$

Ce qui prouve que X *n'est pas intégrable.*

2. Soit $n \in \mathbb{N}^*$, montrer que $P(\{|X| > n\}) \ge \frac{1}{n\pi}$, où $\{|X| > n\} = \{\omega \in \Omega,\ |X(\omega)| > n\}$. [On pourra remarquer que $\frac{2}{1+x^2} \ge \frac{1}{x^2}$ pour tout $x \ge 1$.]

Corrigé – *On a*

$$P(\{|X| > n\}) = \int_{|x|>n} \frac{1}{\pi(1+x^2)} dx \ge \frac{1}{\pi} \int_n^{+\infty} \frac{1}{x^2} dx = \frac{1}{n\pi}.$$

3. Pour $x \in \mathbb{R}$, on pose $g(x) = e^{-|x|}$. Montrer que

$$\frac{2}{1+u^2} = \int_{\mathbb{R}} e^{iux} g(x) dx = \sqrt{2\pi}\,\widehat{g}(u) \ \text{ pour tout } u \in \mathbb{R}.$$

En déduire que la fonction caractéristique de X, notée φ_X vérifie $\varphi_X(u) = e^{-|u|}$ pour tout $u \in \mathbb{R}$. [On pourra utiliser de théorème d'inversion de Fourier qui donne $\widehat{\widehat{g}} = g(-\cdot)$ p.p. si $g, \widehat{g} \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.]

Corrigé – *Soit $u \in \mathbb{R}$. La fonction $x \mapsto e^{iux}g(x)$ est bien intégrable sur $\mathbb{R}$. Comme la fonction g est paire on a*

$$\int_{\mathbb{R}} e^{iux} g(x)dx = 2\int_0^{+\infty} \cos(ux)e^{-x}dx.$$

En intégrant deux fois par parties, on montre que

$$\int_0^{+\infty} \cos(ux)e^{-x}dx = -\int_0^{+\infty} u\sin(ux)e^{-x}dx + 1 = -\int_0^{+\infty} u^2\cos(ux)e^{-x}dx.$$

On en déduit que

$$(1+u^2)\int_0^{+\infty} \cos(ux)e^{-x}dx = 1,$$

et donc que $\int_{\mathbb{R}} e^{iux}g(x)dx = \frac{2}{1+u^2}$. On remarque aussi que, grâce à la parité de g,

$$\widehat{g}(u) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-iux}g(x)dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{iux}g(x)dx.$$

Enfin, la définition de fonction caractéristique de X donne

$$\varphi_X(u) = \sqrt{2\pi}\widehat{f}(u).$$

Comme $f(x) = \frac{1}{2\pi}\frac{2}{1+x^2} = \frac{1}{\sqrt{2\pi}}\widehat{g}(x)$ (pour tout $x \in \mathbb{R}$), on a donc (en utilisant $\widehat{\widehat{g}} = g(-\cdot)$ et la parité de g)

$$\widehat{f} = \frac{1}{\sqrt{2\pi}}\widehat{\widehat{g}} = \frac{1}{\sqrt{2\pi}}g(-\cdot) = \frac{1}{\sqrt{2\pi}}g,$$

ce qui donne, finalement, $\varphi_X = g$.

On se donne maintenant une suite $X_1, X_2, \ldots, X_n, \ldots$ de v.a.r.i.i.d. et on suppose que la loi de X_1 (et donc de tous les X_n) est la loi de Cauchy. Pour $n \in \mathbb{N}^*$, on pose :

$$Z_n = \frac{1}{n}(\sum_{p=1}^{n} X_p).$$

4. Soit $n \in \mathbb{N}^*$.

(a) Montrer que

$$\varphi_{Z_n}(u) = \varphi_{X_1}^n(\frac{u}{n}) \text{ pour tout } u \in \mathbb{R}.$$

Corrigé – *Soit $u \in \mathbb{R}$, on a, grâce à l'indépendance des X_p,*

$$\varphi_{Z_n}(u) = E(e^{iuZ_n}) = E(\prod_{p=1}^{n} e^{i\frac{u}{n}X_p}) = \prod_{p=1}^{n} E(e^{i\frac{u}{n}X_p}).$$

Les v.a.r. X_p étant identiquement distribuées, on a, pour tout $p \in \mathbb{N}^*$ et tout $v \in \mathbb{R}$, $E(e^{iuX_p}) = e^{iuX_1} = \varphi_{X_1}(u)$ et donc

$$\varphi_{Z_n}(u) = \varphi_{X_1}^n(\frac{u}{n}).$$

(b) Donner la loi de Z_n.

Corrigé – *Comme* $\varphi_{X_1} = g$ *et que* $g(x)^n = g(nx)$, *on a* $\varphi_{Z_n} = \varphi_{X_1}$. *La loi de* Z_n *est donc la même que la loi de* X_1, *c'est-à-dire la loi de Cauchy (car deux v.a.r. qui ont même fonction caractéristique ont même loi).*

5. Pour $n \in \mathbb{N}^*$, on pose $A_n = \{|X_n| > n\}$ et $B_n = \bigcup_{p \geq n} A_p$. On pose aussi $B = \bigcap_{n \in \mathbb{N}^*} B_n$.

(a) Soit $n \in \mathbb{N}^*$. En utilisant la question 2, montrer que

$$P(B_n^c) = \prod_{p \geq n} P(A_p^c) \leq \prod_{p \geq n} (1 - \frac{1}{p\pi}).$$

En déduire que $P(B_n^c) = 0$ et donc que $P(B_n) = 1$.

Corrigé – *Comme* $B_n^c = \bigcap_{p=n}^{+\infty} A_p^c$, *on a aussi*

$$B_n^c = \bigcap_{k=n}^{+\infty} C_k \textit{ avec } C_k = \bigcap_{p=n}^{k} A_p^c.$$

Comme $C_{k+1} \subset C_k$ *pour tout* $k \geq n$, *la continuité décroissante de* P *donne*

$$P(B_n^c) = \lim_{k \to +\infty} P(C_k).$$

Pour tout p, *on a* $A_p^c \in \sigma(X_p)$. *Comme les v.a.r.* X_p *sont indépendantes, on a donc*

$$P(C_k) = \prod_{p=n}^{k} P(A_p^c).$$

La question 2 donne $P(A_p^c) = 1 - P(A_p) = 1 - \dfrac{1}{p\pi}$ *et donc*

$$P(C_k) = \prod_{p=n}^{k} (1 - \frac{1}{p\pi}).$$

De l'égalité précédente on déduit que

$$\ln(P(C_k)) = \sum_{p=n}^{k} \ln(1 - \frac{1}{p\pi}).$$

Comme $\lim_{p \to +\infty} p\pi \ln(1 - \frac{1}{p\pi}) = -1$, *le terme de droite de cette égalité tend vers* $-\infty$ *quand* $k \to \infty$. *On a donc* $\lim_{k \to +\infty} P(C_k) = 0$ *ce qui donne*

$$P(B_n^c) = 0 \textit{ et donc } P(B_n) = 1.$$

(b) Montrer que $P(B) = 1$. [Cette question est une conséquence de la question (a) mais elle peut aussi être faite directement en appliquant le lemme de Borel-Cantelli et la question 2.]

Corrigé – *La suite* $(B_n)_{n\in\mathbb{N}^*}$ *est décroissante (c'est-à-dire* $B_{n+1} \subset B_n$ *pour tout* n*). On peut donc appliquer la propriété de continuité décroissante de* P*. On obtient*

$$P(B) = \lim_{n\to+\infty} P(B_n) = 1.$$

(c) Soit $\omega \in B$. Montrer que $\frac{X_n(\omega)}{n} \not\to 0$ quand $n \to +\infty$.

Corrigé – *Pour tout* $n \in \mathbb{N}^*$*, on a* $\omega \in B_n$*. Il existe donc* $p \geq n$ *t.q.* $\omega \in A_p$*, c'est-à-dire* $\frac{|X_p(\omega)|}{p} > 1$*, ceci montre bien que bien que* $\frac{X_n(\omega)}{n} \not\to 0$ *quand* $n \to +\infty$*.*

(d) Soit $\omega \in \Omega$. Montrer que si la suite $(Z_n(\omega))_{n\in\mathbb{N}^*}$ converge vers une limite finie on a alors

$$\lim_{n\to+\infty} \frac{X_n(\omega)}{n} = 0.$$

[Écrire X_n en fonction de Z_n et Z_{n-1}.]

Corrigé – *Pour tout* $n \in \mathbb{N}^*$*, on a* $X_n(\omega) = nZ_n(\omega) - (n-1)Z_{n-1}(\omega)$ *et donc*

$$\frac{X_n(\omega)}{n} = Z_n(\omega) - \frac{n-1}{n} Z_{n-1}(\omega).$$

Si la suite $(Z_n(\omega))_{n\in\mathbb{N}^*}$ *converge (dans* $\mathbb{R}$*), on en déduit bien*

$$\lim_{n\to+\infty} \frac{X_n(\omega)}{n} = 0.$$

(e) Déduire des trois questions précédentes que la suite $(Z_n)_{n\in\mathbb{N}^*}$ ne converge pas p.s. vers une limite finie quand $n \to +\infty$.

Corrigé – *La conséquence des questions précédentes est que pour tout* $\omega \in B$*, la suite* $(Z_n(\omega))_{n\in\mathbb{N}^*}$ *ne converge pas dans* $\mathbb{R}$*. Comme* $P(B) = 1$*, on a donc presque sûrement une non convergence de la suite* $(Z_n)_{n\in\mathbb{N}^*}$*.*

6. Pour tout $n \in \mathbb{N}^*$, montrer que

$$Z_{2n} - Z_n = \frac{U_n + V_n}{2},$$

où U_n et V_n sont deux v.a.r. indépendantes, de loi de Cauchy. En déduire que la suite $(Z_n)_{n\in\mathbb{N}^*}$ ne converge pas en probabilité.

Corrigé – *Pour tout* $n \in \mathbb{N}^*$*, on a*

$$Z_{2n} - Z_n = -\frac{1}{2n}\sum_{i=1}^{n} X_i + \frac{1}{2n}\sum_{i=n+1}^{2n} X_i = \frac{U_n}{2} + \frac{V_n}{2},$$

en posant

$$U_n = -Z_n \text{ et } V_n = \frac{1}{n}\sum_{i=n+1}^{2n} X_i.$$

Les v.a.r. U_n et V_n sont indépendantes (voir la proposition 9.23) et ont pour loi la loi de Cauchy (c'est une conséquence de la question 4(b)). La question 4(b) donne aussi que la v.a.r. $(1/2)(U_n + V_n)$ a pour loi la loi de Cauchy. On obtient donc finalement que la v.a.r. $Z_{2n} - Z_n$ a pour loi la loi de Cauchy et la question 2 donne alors

$$P(|Z_{2n} - Z_n| > 1) \geq \frac{1}{\pi}. \tag{10.14}$$

On en déduit que la suite $(Z_n)_{n\in\mathbb{N}^*}$ ne peut pas converger en probabilités. En effet, si $Z_n \to Z$ en probabilité (quand $n \to +\infty$), on a, par exemple,

$$\{|Z_{2n} - Z_n| > 1\} \subset \{|Z_{2n} - Z| > \frac{1}{2}\} \cup \{|Z_n - Z| > \frac{1}{2}\},$$

et donc

$$P(\{|Z_{2n} - Z_n| > 1\}) \leq P(\{|Z_{2n} - Z| > \frac{1}{2}\}) + P(\{|Z_n - Z| > \frac{1}{2}\}).$$

Comme $Z_n \to Z$ en probabilité, ceci donne

$$\lim_{n\to+\infty} P(\{|Z_{2n} - Z_n| > 1\}) = 0,$$

en contradiction avec (10.14).

7. Pour quelles raisons ne peut-on appliquer, à la suite $(Z_n)_{n\in\mathbb{N}^*}$, les lois forte et faible des grands nombres ?

Corrigé – *La loi forte des grands nombres ne s'applique pas car la suite $(X_n)_{n\in\mathbb{N}}$ n'est pas une suite de v.a.r. intégrables. La loi faible des grands nombres ne s'applique pas car la suite $(X_n)_{n\in\mathbb{N}}$ n'est pas une suite de v.a.r. de carrés intégrables.*

Exercice 10.18 (Sur la loi d'un vecteur aléatoire) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X un v.a. de dimension d. Montrer que la loi de X est uniquement déterminée par la donnée des lois de toutes les v.a.r. $a \cdot X$, $a \in \mathbb{R}^d$, $|a| = 1$. [On pourra utiliser la fonction caractéristique de X.]

Exercice 10.19 (Limite en loi de Gaussiennes) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, X une v.a.r. et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. gaussiennes. On suppose que X_n tend en loi vers X, quand $n \to +\infty$. On note m_n l'espérance de X_n et σ_n^2 la variance de X_n (avec $\sigma_n \geq 0$). Si $\sigma_n > 0$, la v.a.r. X_n a donc pour loi une loi dont la densité par rapport à la mesure de Lebesgue est

$$f_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-m_m)^2}{2\sigma_n^2}} \text{ pour } x \in \mathbb{R}.$$

Si $\sigma_n = 0$, la loi de X_n est δ_{m_n}.

Si Z est une v.a.r., on note φ_Z la fonction caractéristique de Z (on rappelle que φ_Z est une fonction continue de $\mathbb{R}$ dans $\mathbb{C}$ et que $\varphi_Z(0) = 1$). On rappelle que la fonction caractéristique de X_n est

$$\varphi_{X_n}(t) = e^{-\frac{\sigma_n^2}{2}t^2} e^{im_n t} \text{ pour } t \in \mathbb{R}.$$

1. Soit Z une v.a.r.. Montrer qu'il existe $t \in \mathbb{R}^*$ t.q. $|\varphi_Z(t)| > 0$.

   Corrigé – *La fonction φ_Z est continue sur $\mathbb{R}$, elle donc continue en 0. Comme $\varphi_Z(0) = 1$, la fonction φ_Z est donc non nulle au voisinage de 0. Il existe donc $t \neq 0$ t.q. $|\varphi_Z(t)| \neq 0$.*

2. Montrer que la suite $(\sigma_n^2)_{n\in\mathbb{N}}$ est convergente dans $\mathbb{R}$. [On pourra utiliser la convergence de $|\varphi_{X_n}(t)|$ vers $|\varphi_X(t)|$ pour t bien choisi.]

   Corrigé – *On choisit $t \neq 0$ t.q. $|\varphi_X(t)| > 0$ (ce qui est possible grâce à la première question). Comme $X_n \to X$ en loi quand $n \to +\infty$, on a $\lim_{n\to+\infty} \varphi_{X_n}(t) = \varphi_X(t)$ et donc*

$$e^{-\frac{\sigma_n^2}{2}t^2} = |\varphi_{X_n}(t)| \to |\varphi_X(t)| \text{ quand } n \to +\infty.$$

   Comme la fonction $\ln$ est continue sur $\mathbb{R}_+^*$, on en déduit que

$$\lim_{n\to+\infty} \frac{\sigma_n^2}{2} t^2 = -\ln(|\varphi_X(t)|)$$

   et donc

$$\lim_{n\to+\infty} \sigma_n^2 = \sigma^2 \text{ avec } \sigma^2 = -\frac{2\ln(|\varphi_X(t)|}{t^2} \in \mathbb{R}_+.$$

3. Calculer $P(\{X_n \geq m_n\})$ et $P(\{X_n \leq m_n\})$ et en déduire que la suite $(m_n)_{n\in\mathbb{N}}$ est bornée. [On pourra utiliser le fait que la suite $(P_{X_n})_{n\in\mathbb{N}}$ est tendue, d'après la proposition 9.20.]

   Corrigé – *Soit $n \in \mathbb{N}$. Si $\sigma_n = 0$, on a $P_{X_n} = \delta_{m_n}$ et $P(\{X_n \geq m_n\}) = P(\{X_n \leq m_n\}) = 1$.*
   Si $\sigma_n \neq 0$, On a

$$P(\{X_n \geq m_n\}) = \int_{m_n}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-m_m)^2}{2\sigma_n^2}} dx = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{2}.$$

   On a de même $P(\{X_n \leq m_n\}) = 1/2$.

   Comme $X_n \to X$ en loi, la proposition 9.20 donne que la suite $(P_{X_n})_{n\in\mathbb{N}}$ est tendue. Il existe donc $a > 0$ (ne dépendant de n) t.q. $P(\{X_n \geq a\}) < 1/2$ et $P(\{X_n \geq -a\}) < 1/2$. On en déduit que $-a < m_n < a$. La suite $(m_n)_{n\in\mathbb{N}}$ est donc bornée.

4. Montrer que la suite $(m_n)_{n\in\mathbb{N}}$ est convergente. En déduire que X est une v.a.r. gaussienne.

   Corrigé – *Soit $t \in \mathbb{R}$, comme $\lim_{n\to+\infty} \varphi_{X_n}(t) = \varphi_X(t)$ et que $\lim_{n\to+\infty} \sigma_n^2 = \sigma^2$ (donné à la deuxième question) on a*

$$\lim_{n\to+\infty} e^{im_n t} = e^{\frac{\sigma^2}{2}t^2} \varphi_X(t).$$

   Si m et μ sont deux valeurs d'adhérence de la suite $(m_n)_{n\in\mathbb{N}}$, on a donc $e^{imt} = e^{i\mu t}$ pour tout $t \in \mathbb{R}$. Ceci est impossible si $m \neq \mu$ (car, si $m \neq \mu$, en prenant $t = \pi/(m-\mu)$ on a $e^{i(m-\mu)t} = e^{i\pi} = -1 \neq 1$). On a donc $m = \mu$. La suite $(m_n)_{n\in\mathbb{N}}$ est donc une *suite (réelle) bornée ayant une seule valeur d'adhérence, ce qui montre qu'elle est convergente dans* $\mathbb{R}$.

On pose $m = \lim_{n\to+\infty} m_n$, *on a alors* $\varphi_X(t) = e^{-\frac{\sigma^2}{2}t^2}e^{imt}$ *pour tout* $t \in \mathbb{R}$. *Ceci montre que* X *est une v.a.r. gaussienne et* $X \sim \mathcal{N}(m, \sigma^2)$.

Exercice 10.20 (Caractérisation de $X = 0$ **p.s.)** Soit $(\Omega, \mathcal{A}, p)$ un espace probabilisé. Si Z est une v.a.r., on note φ_z la fonction caractéristique de Z, c'est-à-dire $\varphi_Z(t) = E(e^{iZt})$ pour tout $t \in \mathbb{R}$.

1. Soit X une variable aléatoire réelle et $a > 0$. On suppose que $\varphi_X(t) = 1$ pour tout $t \in [-a, a]$.

(a) Montrer que, pour tout $t \in [-a, a]$, $\displaystyle\int_\Omega [1 - \cos(tX(\omega))]dp(\omega) = 0$.

(b) Montrer que $X = 0$ p.s..

2. Soit X et Y deux variables aléatoires réelles indépendantes.

(a) Montrer que $\varphi_{X+Y} = \varphi_X\varphi_Y$.

(b) On suppose maintenant que $X + Y$ et Y ont même loi. Montrer que $X = 0$ p.s..

Exercice 10.21 (Indépendance et variables gaussiennes) Soit $(\Omega, \mathcal{A}, p)$ un espace probabilisé et X, Y deux v.a.r.i.i.d. t.q. $E(X) = 0$ et $\mathrm{Var}(X) = 1$.

On rappelle une caractérisation de l'indépendance avec les fonctions caractéristiques, donnée dans la proposition 10.21 : Soit Z_1, Z_2 deux v.a.r., on note Z le v.a. (de dimension 2) dont les composantes sont Z_1 et Z_2. Les v.a.r. Z_1 et Z_2 sont indépendantes si et seulement si les fonctions caractéristiques de Z, Z_1 et Z_2 (notées φ_Z, φ_{Z_1} et φ_{Z_2}) vérifient $\varphi_Z(u_1, u_2) = \varphi_{Z_1}(u_1)\varphi_{Z_2}(u_2)$ pour tout $u_1, u_2 \in \mathbb{R}$.

1. On suppose dans cette question que X et Y sont des v.a.r. gaussiennes. Montrer que $X + Y$ et $X - Y$ sont des v.a.r. indépendantes. [On pourra utiliser le rappel.]

   On suppose maintenant, pour toute la suite de l'exercice, que $X + Y$ et $X - Y$ sont indépendantes. Le but des questions suivantes est de démontrer que X et Y sont des variables gaussiennes. On note toujours φ_X la fonction caractéristique de X.
2. Montrer que $\lim_{t\to 0} \frac{1-\varphi_X(t)}{t^2} = \frac{1}{2}$. [Utiliser $E(X) = 0$, $\mathrm{Var}(X) = 1$ et un développement limité des fonctions cos et sin.]
3. Montrer que $\varphi_X(2t) = \varphi_X(t)^3\varphi_X(-t)$ pour tout $t \in \mathbb{R}$ [Utiliser l'indépendance de $X + Y$ et $X - Y$]. Montrer que $\varphi_X(t) \neq 0$ pour tout $t \in \mathbb{R}$.
4. Soit $\rho(t) = \frac{\varphi_X(t)}{\varphi_X(-t)}$. Montrer que $\rho(t) = \rho(t/2^k)^{2^k}$ pour tout $t \in \mathbb{R}$ et tout $k \in \mathbb{N}$. Montrer que $\rho(t) = 1$ pour tout $t \in \mathbb{R}$ [utiliser la question 2]. Montrer que φ_X prend ses valeurs dans $\mathbb{R}_+^*$.

5. Montrer que $\varphi_X(t+s)\varphi_X(t-s) = \varphi_X(t)^2\varphi_X(s)^2$ pour tout $t, s \in \mathbb{R}$. En déduire que $\varphi_X(nt) = \varphi_X(t)^{n^2}$ pour tout $t \in \mathbb{R}$ et tout $n \in \mathbb{N}$.
6. Montrer que X (et donc aussi Y) est une v.a.r. gaussienne.

Chapitre 11

Espérance conditionnelle et martingales

A l'origine, une martingale est une technique utilisée par les joueurs dans les jeux de hasard pour tenter d'infléchir le hasard en leur faveur, avec cependant peu de réussite en pratique ! Nous allons ici donner une introduction à la théorie des martingales en probabilités.

11.1 Espérance conditionnelle

Nous commençons par définir l'espérance conditionnée par une tribu.

Définition 11.1 (Espérance conditionnelle d'une v.a.r.) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* X *une v.a.r. intégrable et* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$*. On appelle espérance conditionnée par* $\mathcal{B}$ *de* X*, ou espérance conditionnelle de* X *par rapport à* $\mathcal{B}$ *ou encore espérance de* X *sachant* $\mathcal{B}$*, l'ensemble des applications* Z *de* Ω *dans* $\mathbb{R}$*,* $\mathcal{B}$*-mesurables, intégrables et t.q. :*

$$\mathrm{E}(\mathrm{ZU}) = \mathrm{E}(\mathrm{XU}) \text{ pour toute application } \mathrm{U} : \Omega \to \mathbb{R},\ \mathcal{B}\text{-mesurable, bornée.} \tag{11.1}$$

On note $\mathrm{E}(\mathrm{X}|\mathcal{B})$ *cette espérance conditionnelle (qui est donc un ensemble de fonctions, on montre plus loin que c'est un élément de* $\mathrm{L}^1_{\mathbb{R}}(\Omega, \mathcal{B}, \mathrm{P})$*). Noter que dans (11.1), les applications* ZU *et* XU *sont bien des v.a.r. intégrables.*

Cette définition peut sembler un peu abrupte. On montre dans la proposition 11.2 que, sous les hypothèses de la définition 11.1, l'espérance conditionnelle existe, c'est-à-dire que l'ensemble $\mathrm{E}(\mathrm{X}|\mathcal{B})$ est non vide, et que $\mathrm{E}(\mathrm{X}|\mathcal{B})$ est unique, ceci signifiant que si $Z_1, Z_2 \in E(X|\mathcal{B})$, on a nécessairement $Z_1 = Z_2$ p.s.. En fait, $E(X|\mathcal{B})$ est donc un élément de $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$.

L'ensemble $E(X|\mathcal{B})$, défini dans la définition 11.1, est donc un ensemble de v.a.r. (car Z $\mathcal{B}$-mesurable implique Z $\mathcal{A}$-mesurable) et, en pratique, on confond, comme d'habitude, cet ensemble avec l'un de ces éléments (on confond un élément de $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$ avec l'un de ses représentants). Si Z est une v.a.r. $\mathcal{B}$-mesurable intégrable et t.q. $E(ZU) = E(XU)$ pour toute v.a.r. U $\mathcal{B}$-mesurable bornée, on écrira donc $Z = E(X|\mathcal{B})$ p.s. au lieu d'écrire $Z \in E(X|\mathcal{B})$.

Avant de démontrer l'existence et l'unicité de l'espérance conditionnelle, donnons quelques exemples simples. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a.r. intégrable.

Cas $\mathcal{B}$ est la tribu grossière. Prenons tout d'abord $\mathcal{B} = \{\emptyset, \Omega\}$. Il est alors facile de voir (exercice 11.1) que $E(X|\mathcal{B})$ est réduit à un seul élément et que cet élément est la fonction constante et égale à $E(X)$.

Cas $\mathcal{B} = \mathcal{A}$. Si maintenant $\mathcal{B} = \mathcal{A}$, alors $E(X \mid B) = X$ p.s.. Plus précisément, $E(X|\mathcal{B})$ est ici l'ensemble des v.a.r. Z t.q. $Z = X$ p.s., c'est-à-dire X en tant que élément de $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$.

Cas où $\mathcal{B} = \{\emptyset, A, A^c, \Omega\}$, $0 < P(A) < 1$. Soit maintenant $A \in \mathcal{A}$ t.q. $0 < P(A) < 1$ et $\mathcal{B} = \{\emptyset, A, A^c, \Omega\}$ (qui est bien une tribu incluse dans $\mathcal{A}$). On peut ici montrer (exercice 11.1) que $E(X|\mathcal{B})$ est réduit à un seul élément et cet élément est la fonction Z définie par :

$$Z = \frac{E(X1_A)}{P(A)} 1_A + \frac{E(X1_{A^c})}{P(A^c)} 1_{A^c}.$$

La quantité $\frac{E(X1_A)}{P(A)}$ s'appelle espérance de X sachant A. On a ainsi fait le lien entre espérance de X sachant un événement et espérance de X par rapport à une tribu (ou selon une tribu).

Cas où $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$, $P(B) = 1$. On prend $B \in \mathcal{A}$ t.q. $P(B) = 1$ et $B^c \neq \emptyset$ (c'est le cas, par exemple, si P est une mesure diffuse, que $\mathcal{A}$ contient les singletons et que B^c est formé d'un nombre fini ou dénombrable de points de Ω). On prend encore $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$. Pour $a \in \mathbb{R}$, on pose $Z_a = E(X)1_B + a1_{B^c}$. On peut alors montrer, pour être précis, que $E(X|\mathcal{B}) = \{Z_a,\ a \in \mathbb{R}\}$ (exercice 11.2) c'est-à-dire que $E(X|\mathcal{B})$ est la fonction constante et égale à $E(X)$ en tant que élément de $L^1_{\mathbb{R}}(\Omega, \mathcal{B}, P)$, ce qu'on écrit, avec la confusion habituelle entre un élément de L^1 et l'un de ses représentants, $E(X|\mathcal{B}) = E(X)$ p.s..

On montre maintenant l'existence et l'unicité de l'espérance, conditionnée par une tribu, d'une v.a.r. intégrable.

Proposition 11.2 *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$. *Soit* X *une v.a.r. intégrable. Alors :*

(Existence) $E(X|\mathcal{B}) \neq \emptyset$.

(Unicité) $Z_1, Z_2 \in E(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$ *p.s.*.

DÉMONSTRATION – On démontre d'abord l'unicité de $E(X|\mathcal{B})$. Puis, on démontre l'existence de $E(X|\mathcal{B})$ si X est de carré intégrable, puis l'existence si X est positive (et intégrable) et enfin l'existence si X est seulement intégrable. En fait, la démonstration de l'existence si X est de carré intégrable n'est pas utilisée pour la suite de la démonstration mais elle est éventuellement intéressante pour la compréhension de l'espérance conditionnelle.

Unicité. Soit $Z_1, Z_2 \in E(X|\mathcal{B})$. On pose $U = \text{sign}(Z_1 - Z_2)$ (on rappelle que la fonction sign est définie par $\text{sign}(s) = -1$ si $s < 0$, $\text{sign}(s) = 1$ si $s > 0$ et (par exemple) $\text{sign}(0) = 0$). Comme la fonction sign est borélienne de $\mathbb{R}$ dans $\mathbb{R}$ et que $(Z_1 - Z_2)$ est $\mathcal{B}$-mesurable, la fonction U est bien $\mathcal{B}$-mesurable. Elle est aussi bornée, on a donc en utilisant la définition de l'espéérance conditionnelle (11.1) avec $Z = Z_1$ et $Z = Z_2$, on obtient

$$E(XU) = E(Z_1 U) \text{ et } E(XU) = E(Z_2 U).$$

Ceci donne $E((Z_1 - Z_2)U) = 0$ et donc $E(|Z_1 - Z_2|) = 0$. On en déduit $Z_1 = Z_2$ p.s..

Existence si X **est de carré intégrable.** On note $P_{\mathcal{B}}$ la restriction de P (qui est une mesure sur $\mathcal{A}$) à $\mathcal{B}$ (tribu incluse dans $\mathcal{A}$). La mesure $P_{\mathcal{B}}$ est donc une probabilité sur $\mathcal{B}$. On note H l'espace de Hilbert $L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$ et, pour $V \in H$, on pose :

$$T(V) = \int_\Omega XV dP.$$

Il est clair que $T(V)$ est bien définie. En étant précis, on remarque que $T(V) = \int_\Omega Xv dP$, où $v \in \mathcal{L}^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$ est un représentant de V (et cette quantité ne dépend pas du représentant choisi). C'est pour définir T que nous avons besoin que X soit de carré intégrable.

L'application T est linéaire continue de H dans $\mathbb{R}$ (et on a $\|T\| \le \|X\|_2$). On peut donc appliquer le théorème de Riesz dans les espaces de Hilbert (théorème 6.56), qui donne l'existence de $Z \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$ t.q. :

$$T(V) = \int_\Omega ZV dP \text{ pour tout } V \in H. \tag{11.2}$$

Comme $Z \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$, la fonction Z est bien $\mathcal{B}$-mesurable et intégrable (elle est même de carré intégrable). On montrer maintenant que Z vérifie la propriété (11.1) (et donc que $Z \in E(X|B)$). Soit U une application $\mathcal{B}$-mesurable bornée de Ω dans $\mathbb{R}$. On a $U \in \mathcal{L}^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$, on peut donc utiliser (11.2) avec pour V la classe de U et on obtient

$$E(XU) = T(V) = \int_\Omega ZV dP = E(ZU).$$

L'application Z vérifie donc bien la propriété (11.1), ce qui prouve que $Z \in E(X|\mathcal{B})$.

Plus précisément, un développement du raisonnement ci–avant (que les courageux peuvent faire) permet d'interpréter l'application $X \mapsto E(X|\mathcal{B})$ comme l'opérateur de projection orthogonale de $L^2(\Omega, \mathcal{A}, P)$ dans le sous–espace vectoriel fermé formé à partir de $L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$.

Existence si X **est positive et intégrable.** On utilise ici le théorème de Radon-Nikodym (théorème 6.78, qui se démontre d'ailleurs avec le théorème de Riesz dans les espaces de Hilbert, théorème 6.56). On note toujours $p_{\mathcal{B}}$ la restriction de P à $\mathcal{B}$ (de sorte que $P_{\mathcal{B}}$ est une probabilité sur $\mathcal{B}$).

Pour $B \in \mathcal{A}$, on pose $m(A) = \int_\Omega X1_A dP$. On définit ainsi une mesure finie, m, sur $\mathcal{A}$, c'est la mesure de densité X par rapport à P. On note maintenant $m_{\mathcal{B}}$ la restriction de cette mesure à la tribu $\mathcal{B}$. La mesure $m_{\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$ (car $B \in \mathcal{B}$, $P_{\mathcal{B}}(B) = 0$ implique que $P(B) = 0$ et donc $X1_B = 0$ p.s. et donc $m(B) = 0$ et donc $m_{\mathcal{B}}(B) = 0$). Le théorème de Radon-Nikodym (théorème 6.78) donne alors l'existence de Z, $\mathcal{B}$-mesurable positive, t.q. $m_B = ZP_{\mathcal{B}}$ (c'est-à-dire que $m_{\mathcal{B}}$ est la mesure sur $\mathcal{B}$ de densité Z par rapport à $P_{\mathcal{B}}$).

La fonction Z est intégrable car

$$\int_\Omega Z dP = \int_\Omega Z dP_B = m_{\mathcal{B}}(\Omega) = m(\Omega) = E(X) < +\infty \text{ (et } E(Z) = E(X)).$$

Il reste à montrer que Z vérifie la propriété (11.1) (ce qui donne que $Z \in E(X|B)$). On remarque tout d'abord que

$$E(Z1_B) = \int_\Omega Z1_B dP = m_{\mathcal{B}}(B) = \int_\Omega X1_B dP = E(X1_B) \text{ pour tout } B \in \mathcal{B}$$

. Par linéarité positive, on a donc, pour toute fonction $\mathcal{B}$-étagée positive,

$$E(ZU) = \int_\Omega ZU dP = \int_\Omega ZU dP_{\mathcal{B}} = \int_\Omega U dm = \int_\Omega UX dP = E(XU). \qquad (11.3)$$

Par convergence monotone, on en déduit alors que (11.3) est encore vraie pour toute fonctions $\mathcal{B}$-mesurable positive. Enfin, en utilisant $U = U^+ - U^-$ (et en remarquant que ZU et XU sont intégrables), on conclut que (11.3) est vraie pour toute fonction U $\mathcal{B}$-mesurable bornée de Ω dans $\mathbb{R}$. (On a ici repris un argument vu dans la remarque 6.75.) L'application Z vérifie donc la propriété (11.1), ce qui prouve que $Z \in E(X|\mathcal{B})$.

Existence si X **est seulement intégrable.** Comme les fonctions X^+ et X^- sont positives et intégrables, il existe $Z_1 \in E(X^+|\mathcal{B})$ et $Z_2 \in E(X^-|\mathcal{B})$. On pose $Z = Z_1 - Z_2$. L'application Z est $\mathcal{B}$-mesurable et intégrable (car Z_1 et Z_2 le sont) et, pour tout fonction U $\mathcal{B}$-mesurable bornée, on a :

$$E(ZU) = E(Z_1U) - E(Z_2U) = E(X^+U) - E(X^-U) = E(XU).$$

L'application Z vérifie donc la propriété (11.1), ce qui prouve que $Z \in E(X|\mathcal{B})$.

■

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $\mathcal{B}$ une tribu incluse dans $\mathcal{A}$. On a défini l'espérance, conditionnée par $\mathcal{B}$, d'une v.a.r. intégrable. On va maintenant montrer qu'on peut étendre la définition à des v.a.r. qui ne sont pas intégrables mais qui sont positives (la démonstration est déjà essentiellement dans la démonstration de la proposition 11.2). Pour cela, on va commencer par donner une p.s.-caractérisation de

$E(X|B)$ lorsque X est une v.a.r. positive et intégrable. Cette caractérisation n'utilisant pas l'intégrabilité de X on aura ainsi une définition de $E(X|\mathcal{B})$ lorsque X est une v.a.r. positive. Ceci est fait dans la proposition 11.3 et la définition 11.4.

Proposition 11.3 (Caractérisation de l'espérance conditionnelle d'une v.a.r. positive) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$.

1. *Soit* X *une v.a.r. intégrable positive. Alors,* $Z \in E(X|\mathcal{B})$ *si et seulement si* Z *est* $\mathcal{B}$*-mesurable, intégrable,* ≥ 0 *p.s. et t.q. :*

$$E(ZU) = E(XU), \tag{11.4}$$

pour toute application U *de* Ω *dans* $\mathbb{R}$*,* $\mathcal{B}$*-mesurable et positive.*

2. *Soit* X *une v.a.r. positive. On note* $\bar{E}(X|\mathcal{B})$ *l'ensemble des applications* $\mathcal{B}$*-mesurables positives vérifiant la propriété (11.4). On a alors :*

(a) *(Existence)* $\bar{E}(X|\mathcal{B}) \neq \emptyset$.

(b) *(Unicité)* $Z_1, Z_2 \in \bar{E}(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$ *p.s..*

DÉMONSTRATION – On commence par montrer le premier item. Si $Z \in E(X|\mathcal{B})$, la fonction Z est bien $\mathcal{B}$-mesurable intégrable et vérifie la propriété(11.1). Elle vérifie donc la propriété (11.4) en ajoutant l'hypothèse U bornée. Pour montrer que $Z \geq 0$ p.s., on prend $U = 1_B$ avec $B = \{Z < 0\}$ (U est bien $\mathcal{B}$-mesurable bornée). On obtient $E(ZU) = E(XU) \geq 0$. Comme $ZU \leq 0$, on a donc $ZU = 0$ p.s. et donc $Z \geq 0$ p.s.. Enfin, pour montrer que Z vérifie la propriété (11.4) (c'est-à-dire avec U $\mathcal{B}$-mesurable positive mais non nécessairement bornée), il suffit d'utiliser le théorème de convergence monotone (théorème 4.17) en introduisant $U_n = U1_{B_n}$ avec, pour $n \in \mathbb{N}$, $B_n = \{U \leq n\}$.

Réciproquement, si Z est $\mathcal{B}$-mesurable, intégrable, ≥ 0 p.s. et vérifie la propriété (11.4) ; il est facile de voir que Z vérifie (11.1). En effet, si U est $\mathcal{B}$-mesurable bornée, on utilise (11.4) avec les parties positive et négative de U pour obtenir (11.1). Donc, $Z \in E(X|\mathcal{B})$.

On montre maintenant le deuxième item de la proposition.

Existence

On reprend la démonstration de la proposition 11.2. On rappelle que $p_{\mathcal{B}}$ est la restriction de P à $\mathcal{B}$, $m = XP$ (c'est-à-dire la mesure de densité X par rapport à P) et $m_{\mathcal{B}}$ est la restriction de m à $\mathcal{B}$. La mesure $m_{\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$. La mesure m n'est pas finie si X n'est pas intégrable. On ne peut donc pas appliquer directement le théorème 6.78 (qui demande que $m_{\mathcal{B}}$ soit finie). Pour résoudre cette petite difficulté, on pose, pour $n \in \mathbb{N}$ $m_n = X1_{\{n \leq X < n+1\}}$ (qui est une mesure sur $\mathcal{A}$) et $m_{n,\mathcal{B}}$ sa restriction $\mathcal{B}$. La mesure $m_{n,\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$ et est finie. Le théorème de Radon-Nikodym donne alors l'existence de Z_n, $\mathcal{B}$-mesurable positive, t.q. $m_{n,\mathcal{B}} = Z_n P_{\mathcal{B}}$. On pose alors $Z = \sum_{n \in \mathbb{N}} Z_n$. La fonction Z est $\mathcal{B}$-mesurable positive, il reste à montrer que Z vérifie la propriété (11.4) (ce qui donnera que $Z \in \bar{E}(X|B)$). Soit U une application $\mathcal{B}$-mesurable positive de Ω dans $\mathbb{R}$. Pour tout $n \in \mathbb{N}$, on a

$$E(Z_n U) = \int_\Omega Z_n U dP = \int_\Omega Z_n U dP_{\mathcal{B}} = \int_\Omega U dm_{n,\mathcal{B}} = \int_\Omega UX1_{\{n \le X < n+1\}} dP.$$

En sommant cette dernière égalité pour $n \in \mathbb{N}$, on obtient (par le corollaire 4.18 sur les séries à termes positifs)

$$E(ZU) = E(XU).$$

L'application Z vérifie donc la propriété (11.4), ce qui prouve que $Z \in \bar{E}(X|\mathcal{B})$.

Unicité

Soit $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$. prenons $U = (\text{sign}(Z_1 - Z_2))^+$ (qui est bien $\mathcal{B}$-mesurable et positive). On a donc, par la propriété (11.4),

$$E(Z_1 U) = E(Z_2 U) = E(XU),$$

mais on ne peut rien en déduire car il est possible que $E(XU) = +\infty$. On va donc modifier légèrement U. Pour $n \in \mathbb{N}^*$, on pose

$$B_n = \{Z_1 \le n\} \cap \{Z_2 \le n\} \text{ et } U_n = U1_{B_n}.$$

L'application U_n est encore $\mathcal{B}$-mesurable et positive ; comme $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$, la propriété (11.4) donne $E(Z_1 U_n) = E(Z_2 U_n)$. On a donc

$$0 \le E(Z_1 U_n) = E(Z_2 U_n) \le n,$$

d'où l'on déduit que $E((Z_1 - Z_2)U_n) = 0$. Mais, $(Z_1 - Z_2)U_n \ge 0$. En faisant tendre n vers l'infini, le théorème de convergence monotone (théorème 4.16) donne $E((Z_1 - Z_2)U) = 0$, c'est-à-dire $E((Z_1 - Z_2)^+) = 0$ et donc $Z_1 \le Z_2$ p.s.. En changeant les rôles de Z_1 et Z_2 on a aussi $Z_2 \le Z_1$ p.s.. D'où $Z_1 = Z_2$ p.s.. ■

Définition 11.4 (Espérance conditionnelle d'une v.a.r. positive) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* X *une v.a.r. positive et* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$*. On appelle espérance conditionnée par* $\mathcal{B}$ *de* X *ou espérance conditionnelle de* X *par rapport à* $\mathcal{B}$*, l'ensemble des applications* Z *de* Ω *dans* $\mathbb{R}$*,* $\mathcal{B}$*-mesurables, positives et t.q. :*

$$E(ZU) = E(XU), \forall U : \Omega \to \mathbb{R},\ \mathcal{B} - \textit{mesurable et positive}. \tag{11.5}$$

On note $\bar{E}(X|\mathcal{B})$ *cette espérance conditionnelle (c'est donc un ensemble de fonctions). (Noter que dans (11.5), les applications* ZU *et* XU *sont bien des v.a.r. positives, leur intégrale sur* Ω *est donc bien définie et appartient à* $\bar{\mathbb{R}}_+$*).*

La proposition 11.3 nous donne l'existence et l'unicité (p.s.) de l'espérance conditionnelle lorsque X est une v.a.r. positive. Sous les hypothèses de la définition 11.4, si X est de plus intégrable, on a donc deux définitions de l'espérance conditionnelle de X par rapport à $\mathcal{B}$, notée $E(X|\mathcal{B})$ et $\bar{E}(X|\mathcal{B})$. La proposition 11.3 montre que $Z_1 \in E(X|\mathcal{B})$ et $Z_2 \in \bar{E}(X|\mathcal{B})$ implique $Z_1 = Z_2$ p.s.. En pratique, comme on confond $E(X|\mathcal{B})$ avec l'un de ces éléments et $\bar{E}(X|\mathcal{B})$ avec l'un de ces éléments, on a donc $E(X|\mathcal{B}) = \bar{E}(X|\mathcal{B})$

p.s.. Il est donc inutile de conserver la notation $\bar{\mathrm{E}}(\mathrm{X}|\mathcal{B})$ et on conservera la notation $\mathrm{E}(\mathrm{X}|\mathcal{B})$ dans les deux cas, c'est-à-dire "X v.a.r. intégrable" et "X v.a.r. positive".

Nous donnons maintenant quelques propriétés de l'espérance conditionnelle.

Remarque 11.5 (Linéarité de l'espérance conditionnelle) Un première propriété de l'espérance conditionnelle est sa linéarité. Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé, $\mathcal{B}$ une sous tribu de $\mathcal{A}$ et $\mathrm{X}_1, \mathrm{X}_2$ deux v.a.r. intégrables. On pose $\mathrm{Z}_1 = \mathrm{E}(\mathrm{X}_1|\mathcal{B})$ et $\mathrm{Z}_2 = \mathrm{E}(\mathrm{X}_2|\mathcal{B})$ (plus précisément, Z_1 et Z_2 sont des représentants de $\mathrm{E}(\mathrm{X}_1|\mathcal{B})$ et $\mathrm{E}(\mathrm{X}_2|\mathcal{B})$). Soit U une v.a.r. $\mathcal{B}$-mesurable bornée, on a

$$\mathrm{E}(\mathrm{Z}_1\mathrm{U}) = \mathrm{E}(\mathrm{X}_1\mathrm{U}) \text{ et } \mathrm{E}(\mathrm{Z}_2\mathrm{U}) = \mathrm{E}(\mathrm{X}_2\mathrm{U}).$$

Par linéarité de l'espérance, on a donc $\mathrm{E}((\mathrm{Z}_1 + \mathrm{Z}_2)\mathrm{U}) = \mathrm{E}((\mathrm{X}_1 + \mathrm{X}_2)\mathrm{U})$. Ceci prouve que $\mathrm{Z}_1 + \mathrm{Z}_2 = \mathrm{E}(\mathrm{X}_1 + \mathrm{X}_2|\mathcal{B})$ p.s. et donc que

$$\mathrm{E}(\mathrm{X}_1 + \mathrm{X}_2|\mathcal{B}) = \mathrm{E}(\mathrm{X}_1|\mathcal{B}) + \mathrm{E}(\mathrm{X}_2|\mathcal{B}) \text{ p.s..}$$

Proposition 11.6 *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$ *et* X *une v.a.r.. Soit* $p \in]1, \infty]$ *et* q *le nombre conjugué de* p *(i.e.* $q = p/(p-1)$ *si* $p < +\infty$ *et* $q = 1$ *si* $p = \infty$*). On suppose que* $\mathrm{X} \in \mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{A}, \mathrm{P})$*. Soit* $\mathrm{Z} = \mathrm{E}(\mathrm{X}|\mathcal{B})$ *p.s.. Alors,* $\mathrm{Z} \in \mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{A}, \mathrm{P})$ *et* $\mathrm{E}(\mathrm{ZU}) = \mathrm{E}(\mathrm{XU})$ *pour toute application* U *(de* Ω *dans* $\mathbb{R}$*)* $\mathcal{B}$*-mesurable telle que* $|\mathrm{U}|^q$ *soit intégrable.*

DÉMONSTRATION – La démonstration fait partie de l'exercice 11.6. En fait, le cas $p = 2$ a déjà été vu dans la démonstration de la proposition 11.2. ∎

Proposition 11.7 (Inégalité de Jensen généralisée) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $\mathcal{B}$ *une tribu incluse dans* $\mathcal{A}$ *et* X *une v.a.r. de carré intégrable. Soit* φ *une fonction convexe de* $\mathbb{R}$ *dans* $\mathbb{R}$*. On suppose que* $\varphi(\mathrm{X})$ *est intégrable. On a alors*

$$\mathrm{E}(\varphi(\mathrm{X})|\mathcal{B}) \geq \varphi(\mathrm{E}(\mathrm{X}|\mathcal{B})) \; p.s..$$

DÉMONSTRATION – D'après le lemme 11.8 donné ci–après, comme φ est convexe, il existe c, fonction croissante de $\mathbb{R}$ dans $\mathbb{R}$ (et donc fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) t.q., pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$.

Soit $\mathrm{Z} = \mathrm{E}(\mathrm{X}|\mathcal{B})$ p.s. (plus précisément, Z est un représentant de $\mathrm{E}(\mathrm{X}|\mathcal{B})$). On a donc pour tout $\omega \in \Omega$:

$$\varphi(\mathrm{X}(\omega)) - \varphi(\mathrm{Z}(\omega)) \geq c(\mathrm{Z}(\omega))(\mathrm{X}(\omega) - \mathrm{Z}(\omega)). \tag{11.6}$$

On aimerait intégrer cette inégalité sur un élément (bien choisi) de $\mathcal{B}$ mais cela n'est pas possible car les v.a.r. $\varphi(\mathrm{Z})$ et $c(\mathrm{Z})(\mathrm{X} - \mathrm{Z})$ peuvent ne pas être intégrables (bien que

Z et X soient intégrables). Pour $p \in \mathbb{N}^*$, on introduit donc $A_p = \{|Z| \leq p\}$ de sorte que les v.a.r. $1_{A_p} c(Z)(X - Z)$ et $1_{A_p} \varphi(Z)$ sont intégrables (noter que $c(Z)$ est bornée sur A_p car c est croissante). On pose aussi

$$A = \{E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0\} \text{ et } B_p = A_p \cap A.$$

Soit $p \in \mathbb{N}^*$, l'inégalité (11.6) donne $1_{B_p}(\varphi(X) - \varphi(Z)) \geq 1_{B_p} c(Z)(X - Z)$ et donc, en intégrant sur Ω :

$$\int_{B_p} (\varphi(X) - \varphi(Z)) dP \geq \int_{B_p} c(Z)(X - Z) dP. \tag{11.7}$$

Comme Z et $E(\varphi(X)|\mathcal{B})$ sont $\mathcal{B}$-mesurables, on a $B_p \in \mathcal{B}$ (et donc 1_{B_p} est $\mathcal{B}$-mesurable). On a aussi $c(Z)$ $\mathcal{B}$-mesurable (car c est borélienne) et donc $1_{B_p} c(Z)$ $\mathcal{B}$-mesurable. On en déduit :

$$\int_{B_p} c(Z)(X - Z) dP = E(1_{B_p} c(Z)(X - Z)) = 0 \text{ (car } Z \in E(X|\mathcal{B})),$$

et

$$\int_{B_p} (\varphi(X) - \varphi(Z)) dP = E(1_{B_p}(\varphi(X) - \varphi(Z))) = E\Big(1_{B_p}(E(\varphi(X)|\mathcal{B}) - \varphi(Z))\Big).$$

Avec (11.7), on en déduit :

$$\int_{B_p} (E(\varphi(X)|\mathcal{B}) - \varphi(Z)) dP \geq 0.$$

Comme $E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0$ sur B_p (car $B_p \subset A$), on a donc $P(B_p) = 0$ et donc $P(A) = P(\bigcup_{p \in \mathbb{N}^*} B_p) = 0$, ce qui donne bien $E(\varphi(X)|\mathcal{B}) \geq \varphi(Z)$ p.s.. ■

Voici maintenant le lemme utilisé dans la démonstration précédente.

Lemme 11.8 *Soit φ une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$, il existe alors c, fonction croissante de $\mathbb{R}$ dans $\mathbb{R}$ (et donc fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) t.q., pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$.*

DÉMONSTRATION – Si φ est dérivable sur $\mathbb{R}$, la fonction c existe et est unique, elle est donnée par $c = \varphi'$. L'existence de c est légèrement plus difficile si φ n'est pas dérivable sur tout $\mathbb{R}$ (et on perd l'unicité de c).

Soit $a \in \mathbb{R}$, on considère le fonction $h_a : x \mapsto \frac{\varphi(x) - \varphi(a)}{x - a}$ qui est définie sur $\mathbb{R} \setminus \{a\}$. La convexité de φ permet de montrer que h_a est croissante (c'est-à-dire que $x, y \in \mathbb{R} \setminus \{a\}$, $x > y \Rightarrow h_a(x) \geq h_a(y)$). La fonction h_a a donc une limite à gauche (et à droite) en tout point, y compris au point a. On pose (par exemple) :

$$c(a) = \lim_{x \to a, x < a} h_a(x).$$

Il est facile de vérifier que la fonction c ainsi définie est croissante de $\mathbb{R}$ dans $\mathbb{R}$ et vérifie, pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$. ■

On définit maintenant l'espérance conditionnelle par rapport à une v.a.r. ou un v.a.

Définition 11.9 (Espérance conditionnellement à une v.a.r) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X *une v.a.r. (ou un v.a. de dimension* d*,* $d \geq 1$*). Soit* Y *une v.a.r. intégrable ou une v.a.r. positive. On appelle espérance conditionnée par* X *de* Y*" ou espérance conditionnelle de* Y *par rapport à* X*" (ou espérance de* Y *sachant* X*) l'ensemble* $\mathrm{E}(\mathrm{Y}|\sigma(\mathrm{X}))$*, où* $\sigma(\mathrm{X})$ *est la tribu engendrée par* X*. On note* $\mathrm{E}(\mathrm{Y}|\mathrm{X})$ *cette espérance conditionnelle, de sorte que* $\mathrm{E}(\mathrm{Y}|\mathrm{X}) = \mathrm{E}(\mathrm{Y}|\sigma(\mathrm{X}))$*. (L'ensemble* $\mathrm{E}(\mathrm{Y}|\mathrm{X})$ *est donc un ensemble de v.a.r. et, comme d'habitude, on confond* $\mathrm{E}(\mathrm{Y}|\mathrm{X})$ *avec l'un de ces éléments.)*

Pour caractériser $\mathrm{E}(\mathrm{Y}|\mathrm{X})$ (sous les hypothèses de la définition 11.9) et pour calculer cette espérance conditionnelle, on utilise, en général, le théorème 3.31 que nous rappelons sous une forme légèrement plus précise (donnée dans la démonstration du théorème 3.31).

Théorème 11.10 (Mesurabilité d'une v.a.r. par rapport à une autre) *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X*,* Y *deux v.a.r.. On note* $\sigma(\mathrm{X})$ *la tribu engendrée par* X*. Alors :*

— *La v.a.r.* Y *est* $\sigma(\mathrm{X})$*-mesurable si et seulement s'il existe* f*, fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$*, t.q.* $\mathrm{Y} = f(\mathrm{X})$*.*

— *La v.a.r.* Y *est* $\sigma(\mathrm{X})$*-mesurable bornée si et seulement s'il existe* f*, fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*, t.q.* $\mathrm{Y} = f(\mathrm{X})$*.*

— *La v.a.r.* Y *est* $\sigma(\mathrm{X})$*-mesurable positive si et seulement s'il existe* f*, fonction borélienne positive de* $\mathbb{R}$ *dans* $\mathbb{R}$*, t.q.* $\mathrm{Y} = f(\mathrm{X})$*.*

DÉMONSTRATION – La démonstration de ce théorème est donnée dans la démonstration du théorème 3.31. ■

Voici une conséquence immédiate de ce théorème, utilisée pour calculer $\mathrm{E}(\mathrm{Y}|\mathrm{X})$

Proposition 11.11 (Calcul de $\mathrm{E}(\mathrm{Y}|\mathrm{X})$**)** *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* X*,* Y *deux v.a.r.. Soit* Z *une application de* Ω *dans* $\mathbb{R}$*.*

1. On suppose que Y *est intégrable. Alors,* $\mathrm{Z} \in \mathrm{E}(\mathrm{Y}|\mathrm{X})$ *si et seulement s'il existe* ψ *application borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$ *t.q.* $\mathrm{Z} = \psi(\mathrm{X})$*,* $\psi(\mathrm{X})$ *est intégrable et*

$$\mathrm{E}(\psi(\mathrm{X})\varphi(\mathrm{X})) = \mathrm{E}(\mathrm{Y}\varphi(\mathrm{X})), \tag{11.8}$$

pour toute application φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$*, borélienne bornée.*

2. *On suppose que* Y *est positive. Alors,* $Z \in E(Y|X)$ *si et seulement s'il existe* ψ *application borélienne positive de* $\mathbb{R}$ *dans* $\mathbb{R}$ *t.q.* $Z = \psi(X)$ *et*

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \tag{11.9}$$

pour toute application φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$*, borélienne positive.*

DÉMONSTRATION –
La démonstration est une conséquence facile du théorème 11.10. ∎

La conséquence de la proposition 11.11 est que (sous les hypothèses de la proposition) l'on cherche $E(Y|X)$ sous la forme d'une fonction $\psi(X)$ (avec ψ de $\mathbb{R}$ dans $\mathbb{R}$) vérifiant (11.8) (ou (11.9)). On raisonne, en général, par condition nécessaire sur ψ et, comme on sait que $E(X|Y)$ existe, il est même inutile de vérifier que la fonction $\psi(X)$ que l'on trouve (qui est, en général, définie p.s.) est bien intégrable (ou positive).

La proposition 11.11 montre également que (sous les hypothèses de la proposition 11.11) la fonction Y est une fonction de X (c'est-à-dire $Y = \psi(X)$ pour un certain ψ de $\mathbb{R}$ dans $\mathbb{R}$) si et seulement si $E(Y|X) = Y$. Pour montrer que la v.a.r. Y est une fonction d'une autre v.a.r. X, il suffit donc de montrer que $E(Y|X) = Y$. En comparaison, le calcul de la covariance entre X et Y (après normalisation) s'intéresse seulement à l'existence ou non d'une dépendance affine de Y en fonction de X. Voir, à ce propos, l'exercice 11.14.

Remarque 11.12 (Deux propriétés de l'espérance conditionnelle) Voici deux propriétés qui nous serons utiles dans la section suivante. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{B}$ une sous tribu de $\mathcal{A}$ et X une v.a.r. intégrable.

1. Soit V une v.a.r. $\mathcal{B}$-mesurable bornée. On a alors $E(XV|\mathcal{B}) = VE(X|B)$ p.s.. (Voir l'exercice 11.4.)
2. On suppose que $\sigma(X)$ et $\mathcal{B}$ sont des tribus indépendantes. On a alors $E(X|\mathcal{B}) = E(X)$ p.s.. En particulier, si Y est une v.a.r. indépendante de X (c'est-à-dire que $\sigma(X)$ et $\sigma(Y)$ sont des tribus indépendantes), on a alors $E(X|Y) = E(X)$ p.s.. (Voir l'exercice 11.5.)

11.2 Martingales

Définition 11.13 (Filtration et processus) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé*

1. *On appelle filtration une suite de tribus* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *t.q.* $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{A}$*, pour tout* $n \in \mathbb{N}$.

2. *On appelle processus réel une suite de v.a.r..*
3. *Soit* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *une filtration et* $(X_n)_{n\in\mathbb{N}}$ *un processus réel. On dit que* $(X_n)_{n\in\mathbb{N}}$ *est adapté à la filtration* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *si, pour tout* $n \in \mathbb{N}$, X_n *est* $\mathcal{B}_n$*-mesurable.*

Définition 11.14 (Martingale) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *une filtration et* $(X_n)_{n\in\mathbb{N}}$ *un processus réel (c'est-à-dire une suite de v.a.r.).*

(Martingale) *La suite* $(X_n)_{n\in\mathbb{N}}$ *est une martingale par rapport à la filtration* $(\mathcal{B}_n)_n$ *si on a, pour tout* $n \in \mathbb{N}$,

1. X_n *est* $\mathcal{B}_n$*-mesurable et intégrable,*
2. $E(X_{n+1}|\mathcal{B}_n) = X_n$ *p.s..*

(Sous et sur martingale) *La suite* $(X_n)_{n\in\mathbb{N}}$ *est une sous–martingale [resp. sur–martingale] par rapport à la filtration* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *si on a, pour tout* $n \in \mathbb{N}$,

1. X_n *est* $\mathcal{B}_n$*-mesurable et intégrable,*
2. $E(X_{n+1}|\mathcal{B}_n) \geq X_n$ *p.s. [resp.* $E(X_{n+1}|\mathcal{B}_n) \leq X_n$ *p.s.].*

Remarque 11.15 Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $(\mathcal{B}_n)_{n\in\mathbb{N}}$ une filtration et $(X_n)_{n\in\mathbb{N}}$ une martingale par rapport à la filtration $(\mathcal{B}_n)_n$. Pour tout $n \in \mathbb{N}$, on a, comme 1_Ω est $\mathcal{B}_n$-intégrable bornée, et $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s.,

$$E(X_{n+1}) = \int_\Omega X_{n+1} 1_\Omega dP = \int_\Omega E(X_{n+1}|\mathcal{B}_n) 1_\Omega dP = \int_\Omega X_n 1_\Omega dP = E(X_n).$$

On a donc $E(X_n) = E(X_0)$ pour tout $n \in \mathbb{N}$.

Exemple 11.16 (Exemples de Martingales) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

1. Soit $(\mathcal{B}_n)_{n\in\mathbb{N}}$ une filtration et X une v.a.r. intégrable. On pose $X_n = E(X|\mathcal{B}_n)$. La suite $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). (Voir l'exercice 11.24.)
2. Soit X_0 une v.a.r. intégrable et $(J_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. intégrable et de moyenne nulle. On suppose que la suite formée de X_0 et $(J_n)_{n\in\mathbb{N}^*}$ est une suite de v.a.r. indépendantes. On pose alors, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$ et $\mathcal{B}_n$ la tribu engendrée par $X_0, \ldots, X_n$. La suite $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). (Voir l'exercice 11.25.)

Définition 11.17 (Temps d'arrêt) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé,* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *une filtration et* T *une v.a. à valeurs dans* $\mathbb{N} \cup \{+\infty\}$ *(c'est-à-dire une application mesurable de* Ω*, muni de la tribu* $\mathcal{A}$*, dans* $\mathbb{N} \cup \{+\infty\}$*, muni de la tribu formée de l'ensemble des parties de* $\mathbb{N} \cup \{+\infty\}$*). L'application* T *s'appelle un temps d'arrêt si, pour tout* $n \in \mathbb{N}$,

$\{T = n\} \in \mathcal{B}_n$. *Si* T *est un temps d'arrêt, on note* $\mathcal{B}_T$ *la tribu définie par* $\mathcal{B}_T = \{A \in \mathcal{B}_\infty$ *t.q., pour tout* $n \in \mathbb{N}$, $A \cap \{T = n\} \in \mathcal{B}_n\}$ *où* $\mathcal{B}_\infty$ *est la plus petite tribu contenant toutes les tribus* $\mathcal{B}_n$ *(c'est-à-dire la tribu engendrée par les* $\mathcal{B}_n$*).*

Le théorème suivant montre qu'une martingale arrêtée est encore une martingale.

Théorème 11.18 (Martingale arrêtée à un temps d'arrêt) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé muni d'une filtration* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *et* $(X_n)_{n\in\mathbb{N}}$ *une martingale (par rapport à la filtration* $(\mathcal{B}_n)_n$*). Soit* ν *un temps d'arrêt. Pour* $n \in \mathbb{N}$, *on pose* $Y_n = X_{\nu\wedge n}$ *(On rappelle que* $\nu \wedge n(\omega) = \min\{\nu(\omega), n\}$ *et donc que* $X_{\nu\wedge n}(\omega) = X_{\min\{\nu(\omega),n\}}(\omega)$, *pour tout* $\omega \in \Omega$*.) Alors, la suite* $(Y_n)_{n\in\mathbb{N}}$ *est encore une martingale (par rapport à la filtration* $(\mathcal{B}_n)_n$*).*

DÉMONSTRATION – Soit $n \in \mathbb{N}$, on a $Y_n = X_n 1_{\{T>n\}} + \sum_{k=0}^n X_k 1_{\{T=k\}}$. On en déduit tout d'abord que Y_n est intégrable (comme somme finie de fonctions intégrables). Puis, on montre que Y_n est $\mathcal{B}_n$-mesurable. Pour cela, on remarque que $\{T = k\} \in \mathcal{B}_k \subset \mathcal{B}_n$, pour $k \le n$, et que

$$\{T > n\} = \Big(\bigcup_{k=0}^{n} \{T = k\}\Big)^c \in \mathcal{B}_n.$$

Enfin, on remarque que X_k est $\mathcal{B}_n$-mesurable pour tout $k \le n$. Grâce à la stabilité des fonctions mesurables par somme et produit, on obtient bien, finalement, que Y_n est $\mathcal{B}_n$-mesurable.

Il reste maintenant à montrer que $E(Y_{n+1}|\mathcal{B}_n) = Y_n$ p.s., pour tout $n \in \mathbb{N}$. Soit $n \in \mathbb{N}$, on a

$$Y_{n+1} = X_{n+1} 1_{\{T>n+1\}} + \sum_{k=0}^{n+1} X_k 1_{\{T=k\}} = X_{n+1} 1_{\{T\ge n+1\}} + \sum_{k=0}^{n} X_k 1_{\{T=k\}}.$$

Par linéarité de l'espérance conditionnelle, on a donc

$$E(Y_{n+1}|\mathcal{B}_n) = E(X_{n+1} 1_{\{T\ge n+1\}}|\mathcal{B}_n) + \sum_{k=0}^{n} E(X_k 1_{\{T=k\}}|\mathcal{B}_n).$$

Comme $\{T \ge n+1\} = \left(\bigcup_{k=0}^n \{T = k\}\right)^c \in \mathcal{B}_n$, la remarque 11.12 (et le fait que $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s.) donne

$$E(X_{n+1} 1_{\{T\ge n+1\}}|\mathcal{B}_n) = 1_{\{T\ge n+1\}} E(X_{n+1}|\mathcal{B}_n) = X_n 1_{\{T\ge n+1\}}.$$

Puis comme, pour $k \in \{0, \dots, n\}$, $X_k 1_{\{T=k\}}$ est $\mathcal{B}_n$-mesurable, on a $E(X_k 1_{\{T=k\}}|\mathcal{B}_n) = X_k 1_{\{T=k\}}$. On obtient ainsi

$$E(Y_{n+1}|\mathcal{B}_n) = X_n 1_{\{T\ge n+1\}} + \sum_{k=0}^{n} X_k 1_{\{T=k\}} = Y_n \text{ p.s..}$$

Ce qui termine la démonstration. ■

On conclut cette section par un théorème, sans démonstration, sur la convergence des martingales.

Théorème 11.19 (Convergence p.s. d'une martingale)

Soit $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé,* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *une filtration et* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *une suite de v.a.r.. On suppose que* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *est une martingale par rapport à* $(\mathcal{B}_n)_{n\in\mathbb{N}}$.

1. *On suppose que la suite* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *est bornée dans* $\mathrm{L}^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathrm{P})$. *Alors il existe une v.a.r. intégrable,* X, *t.q.* $\mathrm{X}_n \to \mathrm{X}$ *p.s., quand* $n \to +\infty$.
2. *On suppose* $\mathrm{X}_n \geq 0$ *pour tout* $n \in \mathbb{N}$. *Alors, il existe une v.a.r. intégrable,* X, *t.q.* $\mathrm{X}_n \to \mathrm{X}$ *p.s., quand* $n \to +\infty$.

On peut noter que le deuxième item du théorème 11.19 est une conséquence du premier car, pour une martingale, on a toujours $\mathrm{E}(\mathrm{X}_n) = \mathrm{E}(\mathrm{X}_0)$ pour tout $n \in \mathbb{N}$ (et les X_n sont toujours intégrales). Si $\mathrm{X}_n \geq 0$, on a donc $\|\mathrm{X}_n\|_1 = \mathrm{E}(\mathrm{X}_0) < +\infty$.

11.3 Exercices

11.3.1 Espérance conditionnelle

Exercice 11.1 (Espérance conditionnellement à une tribu) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et Y une variable aléatoire réelle intégrable. Dans les trois cas suivants, montrer que $\mathrm{E}(\mathrm{Y}|\mathcal{B})$ est réduit à un élément et déterminer $\mathrm{E}(\mathrm{Y}|\mathcal{B})$ (en fonction de Y et $\mathcal{B}$).

1. La tribu $\mathcal{B}$ la tribu grossière, c'est-à-dire $\mathcal{B} = \{\emptyset, \Omega\}$.

> **Corrigé** – *Soit* Z *une application de* Ω *dans* $\mathbb{R}$, $\mathcal{B}$*-mesurable. Soit* $a \in \mathrm{Im}(\mathrm{Z})$ *(on suppose, bien sûr,* $\Omega \neq \emptyset$*). On a alors* $\{\mathrm{Z} = a\} = \{\omega \in \Omega;\ \mathrm{Z}(\omega) = a\} \neq \emptyset$. *Comme* Z *est* $\mathcal{B}$*-mesurable, on a donc* $\{\mathrm{Z} = a\} = \Omega$. *Une application* $\mathcal{B}$*-mesurable est donc une fonction constante (réciproquement, une fonction constante est bien* $\mathcal{B}$*-mesurable). Si* $\mathrm{Z} \in \mathrm{E}(\mathrm{Y}|\mathcal{B})$, *il existe donc* $a \in \mathbb{R}$ *t.q.* $\mathrm{Z}(\omega) = a$ *pour tout* $\omega \in \Omega$. *Le réel a doit alors vérifier* $\mathrm{E}(a\mathrm{U}) = \mathrm{E}(\mathrm{UY})$ *pour tout application* U, $\mathcal{B}$*-mesurable de* Ω *dans* $\mathbb{R}$. *On a donc* $ab = \mathrm{E}(ab) = \mathrm{E}(b\mathrm{Y}) = b\mathrm{E}(\mathrm{Y})$ *pour tout* $b \in \mathbb{R}$. *La seule solution est donc* $a = \mathrm{E}(\mathrm{Y})$. *L'ensemble* $\mathrm{E}(\mathrm{Y}|\mathcal{B})$ *est donc réduit à un seul élément, la fonction constante et égale à* $\mathrm{E}(\mathrm{Y})$.

2. Soit $\mathrm{B} \in \mathcal{A}$ t.q. $0 < \mathrm{P}(\mathrm{B}) < 1$. On prend pour $\mathcal{B}$ la tribu engendrée par B.

> **Corrigé** – *Soit* Z *une application de* Ω *dans* $\mathbb{R}$, $\mathcal{B}$*-mesurable. Les parties* B *et* B^c *sont non vides (car de probabilité strictement positive). Soit* $\omega_1 \in \mathrm{B}$ *et* $a = \mathrm{Z}(\omega_1)$. *On a alors* $\{\mathrm{Z} = a\} = \{\omega \in \Omega;\ \mathrm{Z}(\omega) = a\} \neq \emptyset$. *Comme* Z *est* $\mathcal{B}$*-mesurable, on a donc* $\{\mathrm{Z} = a\} = \mathrm{B}$ *ou* Ω *et donc* $\{\mathrm{Z} = a\} \supset \mathrm{B}$. *De même, soit* $\omega_2 \in \mathrm{B}^c$ *et* $b = \mathrm{Z}(\omega_2)$, *on a* $\{\mathrm{Z} = b\} \supset \mathrm{B}^c$. *Une application* $\mathcal{B}$*-mesurable est donc une fonction constante sur* B *et* B^c. *Réciproquement, une fonction constante sur* B *et* B^c *est bien* $\mathcal{B}$*-mesurable.*

Si $Z \in E(Y|\mathcal{B})$*, il existe donc* $a, b \in \mathbb{R}$ *t.q.* $Z = a1_B + b1_{B^c}$*. Les réels* a, b *doivent alors vérifier* $E(ZU) = E(UY)$ *pour tout application* U $\mathcal{B}$*-mesurable de* Ω *dans* $\mathbb{R}$*, c'est-à-dire :*

$$a\alpha P(B) + b\beta P(B^c) = \alpha \int_B Y dP + \beta \int_{B^c} Y dP \text{ pour tout } \alpha, \beta \in \mathbb{R}.$$

Comme $P(B) > 0$ *et* $P(B^c) = 1 - P(B) > 0$*, la seule solution est donc :*

$$a = \frac{\int_B Y dP}{P(B)} \text{ et } b = \frac{\int_{B^c} Y dP}{P(B^c)}.$$

L'ensemble $E(Y|\mathcal{B})$ *est donc réduit à un seul élément, la fonction* Z *définie par*

$$Z = \frac{\int_B Y dP}{P(B)} 1_B + \frac{\int_{B^c} Y dP}{P(B^c)} 1_{B^c}.$$

3. Soit $(B_n)_{n\in\mathbb{N}^*} \subset \mathcal{A}$ t.q. $B_n \cap B_m = \emptyset$ si $n \neq m$, $\Omega = \bigcup_{n\in\mathbb{N}^*} B_n$ et $0 < P(B_n) < 1$ pour tout $n \in \mathbb{N}^*$. On prend pour $\mathcal{B}$ la tribu engendrée par $(B_n)_{n\in\mathbb{N}^*}$ (c'est-à-dire $\mathcal{B} = \{\bigcup_{n\in J} B_n, J \subset \mathbb{N}^*\}$).

Corrigé – *On reprend le même raisonnement que dans les deux questions précédentes. On remarque d'abord qu'une application* Z *de* Ω *dans* $\mathbb{R}$ *est* $\mathcal{B}$*-mesurable si et seulement si il existe une suite* $(\alpha_n)_{n\in\mathbb{N}^*} \subset \mathbb{R}$ *t.q.* $Z = \sum_{n\in\mathbb{N}^*} \alpha_n 1_{B_n}$*. (Cette série est bien convergente en tout point de* Ω *car les* B_n *sont disjoints deux à deux.)*

Si $Z \in E(Y|\mathcal{B})$*, il existe donc* $(a_n)_{n\in\mathbb{N}^*} \subset \mathbb{R}$ *t.q.* $Z = \sum_{n\in\mathbb{N}^*} a_n 1_{B_n}$*. La suite* $(a_n)_{n\in\mathbb{N}^*}$ *doit alors être telle que* Z *soit intégrable et que* $E(ZU) = E(UY)$ *pour tout application* U $\mathcal{B}$*-mesurable bornée de* Ω *dans* $\mathbb{R}$*, c'est-à-dire t.q.*

$$\sum_{n\in\mathbb{N}^*} |a_n| P(B_n) < +\infty \text{ et } \sum_{n\in\mathbb{N}^*} \alpha_n a_n P(B_n) = \sum_{n\in\mathbb{N}^*} \alpha_n \int_{B_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n\in\mathbb{N}^*} \subset \mathbb{R}$*. Comme* $P(B_n) > 0$*, la seule solution est donc :*

$$a_n = \frac{\int_{B_n} Y dP}{P(B_n)} \text{ pour tout } n \in \mathbb{N}^*.$$

Comme on sait que l'ensemble $E(Y|\mathcal{B})$ *est non vide, il est inutile de vérifier que la fonction* Z *trouvée est intégrable (puisque cette fonction est la seule fonction pouvant appartenir à* $E(Y|\mathcal{B})$*). L'ensemble* $E(Y|\mathcal{B})$ *est donc réduit à un seul élément. Cet élément est la fonction* Z *définie par*

$$Z = \sum_{n\in\mathbb{N}^*} \frac{\int_{B_n} Y dP}{P(B_n)} 1_{B_n}.$$

Exercice 11.2 (Espérance conditionnellement à une tribu (2)) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une v.a.r. intégrable.

1. Soit $B \in \mathcal{A}$ t.q. $P(B) = 1$ et $B^c \neq \emptyset$ (c'est le cas, par exemple, si P est une mesure diffuse, que $\mathcal{A}$ contient les singletons et que B^c est formé d'un nombre fini ou dénombrable de points de Ω). On pose $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$. Pour $a \in \mathbb{R}$, on pose

$$Z_a = E(X)1_B + a1_{B^c}.$$

Montrer que $E(X|\mathcal{B})$ est l'ensemble des v.a.r. Z_a avec $a \in \mathbb{R}$ (en pratique, comme on confond $E(X|\mathcal{B})$ avec l'un de ses représentants, on peut écrire $E(X|\mathcal{B}) = Z_a$ p.s. avec n'importe quel a dans $\mathbb{R}$, et donc, par exemple, $E(X|\mathcal{B}) = E(X)$ p.s..)

Corrigé – *On raisonne comme dans l'exercice 11.1. Si* $Z \in E(X|\mathcal{B})$, *il existe* $a, b \in \mathbb{R}$ *t.q.* $Z = a1_B + b1_{B^c}$ *et les réels* a, b *doivent alors vérifier* $E(ZU) = E(UY)$ *pour tout application* U $\mathcal{B}$*-mesurable de* Ω *dans* $\mathbb{R}$, *c'est-à-dire :*

$$a\alpha P(B) + b\beta P(B^c) = \alpha \int_B X dP + \beta \int_{B^c} X dP \text{ pour tout } \alpha, \beta \in \mathbb{R}.$$

Comme $P(B) = 0$ *et* $P(B^c) = 1$ *ceci donne* $a\alpha = \alpha E(X)$ *pour tout* $\alpha \in \mathbb{R}$ *et donc* $a = E(X)$. *On a donc bien* $E(X|\mathcal{B}) = \{Z_b,\ b \in \mathbb{R}\}$.

2. Soit I un ensemble fini ou dénombrable, $(B_n)_{n\in I}$ une famille de sous tribus de $\mathcal{A}$ telle que

$$B_n \cap B_m = \emptyset \text{ si } n \neq m\ \Omega = \bigcup_{n\in I} B_n.$$

On prend pour $\mathcal{B}$ la tribu engendrée par $(B_n)_{n\in I}$ (c'est-à-dire $\mathcal{B} = \{\bigcup_{n\in J} B_n,\ J \subset I\}$). Montrer que

$$E(X|\mathcal{B}) = \sum_{n\in J} \frac{\int_{B_n} X dP}{P(B_n)} 1_{B_n} \text{ p.s.},$$

où $J = \{n \in I, \text{ t.q. } P(B_n) > 0\}$.

Corrigé – *On raisonne encore comme dans l'exercice 11.1.*

On remarque d'abord qu'une application Z *de* Ω *dans* $\mathbb{R}$ *est* $\mathcal{B}$*-mesurable si et seulement si il existe une suite* $(\alpha_n)_{n\in I} \subset \mathbb{R}$ *telle que*

$$Z = \sum_{n\in I} \alpha_n 1_{B_n}.$$

Si $Z \in E(X|\mathcal{B})$, *il existe donc* $(a_n)_{n\in I} \subset \mathbb{R}$ *t.q.* $Z = \sum_{n\in I} a_n 1_{B_n}$. *La famille* $(a_n)_{n\in I}$ *doit alors être telle que* Z *soit intégrable et que* $E(ZU) = E(UY)$ *pour toute application* U $\mathcal{B}$*-mesurable bornée de* Ω *dans* $\mathbb{R}$, *c'est-à-dire telle que*

$$\sum_{n\in I} |a_n| P(B_n) < +\infty \text{ et } \sum_{n\in I} \alpha_n a_n P(B_n) = \sum_{n\in I} \alpha_n \int_{B_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n\in I} \subset \mathbb{R}$. *Comme* $P(B_n) > 0$ *si* $n \in J$ *et* $P(B_n) = 0$ *si* $n \notin J$, *ceci donne*

$$a_n = \frac{\int_{B_n} Y dP}{P(B_n)} \text{ pour tout } n \in J,$$

ce qui définit Z *p.s.. On obtient donc*

$$Z = \sum_{n\in J} \frac{\int_{B_n} Y dP}{P(B_n)} 1_{B_n} \text{ p.s.}.$$

Exercice 11.3 (Espérance conditionnellement à une v.a.r.) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, X une variable aléatoire réelle et Y une variable aléatoire réelle intégrable.

1. On suppose qu'il existe $a \in \mathbb{R}$ t.q. $X = a$ p.s.. Donner un élément de $E(Y|X)$.

Corrigé – *On utilise la proposition 11.11. Soit* $Z \in E(Y|X)$, *il existe alors* ψ, *fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$, *t.q.* $Z = \psi(X)$, $\psi(X)$ *est intégrable et, pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$,*borélienne bornée,*

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)).$$

On a donc $Z = \psi(a)$ *p.s. et en prenant pour* φ *une fonction t.q.* $\varphi(a) = 1$ *dans l'égalité précédente, on obtient* $\psi(a) = E(Y)$. *On a donc finalement* $Z = E(Y)$ *p.s.. La fonction constante et égale à* $E(Y)$ *est un élément de* $E(Y|X)$. *Plus précisément, la fonction* $\psi(X)$ *est un élément de* $E(Y|X)$, *dès que* ψ *est une fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$ *et t.q.* $\psi(a) = E(Y)$.

2. On suppose que X prend p.s. deux valeurs x_1 ou x_2 avec $x_1 \neq x_2$. Donner un élément de $E(Y|X)$.

Corrigé – *On pose*

$$A_1 = \{X = x_1\} \textit{ et } A_2 = \{X = x_2\}.$$

On suppose que $P(A_1) > 0$ *et* $P(A_2) > 0$ *(sinon, on est ramené à la question précédente). Noter que* $A_1 \cap A_2 = \emptyset$ *et* $P(A_1) + P(A_2) = 1$. *On utilise encore la proposition 11.11. Soit* $Z \in E(Y|X)$, *il existe alors* ψ, *fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$, *t.q.* $Z = \psi(X)$, $\psi(X)$ *est intégrable et, pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$,*borélienne bornée,*

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \tag{11.10}$$

On a donc $Z = \psi(x_1)1_{A_1} + \psi(x_2)1_{A_2}$ *p.s.. Soit* $\alpha_1, \alpha_2 \in \mathbb{R}$, *en prenant pour* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$ *t.q.* $\varphi(x_1) = a_1$ *et* $\varphi(x_2) = a_2$ *dans l'égalité (11.10), on obtient :*

$$\psi(x_1)\alpha_1 P(A_1) + \psi(x_2)\alpha_2 P(A_2) = \alpha_1 \int_{A_1} Y dP + \alpha_2 \int_{A_2} Y dP,$$

pour tout $\alpha_1, \alpha_2 \in \mathbb{R}$. *Comme* $P(A_i) > 0$ *pour* $i = 1, 2$, *on en déduit que* $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$, *pour* $i = 1, 2$, *et donc que*

$$Z = \frac{\int_{A_1} Y dP}{P(A_1)} 1_{A_1} + \frac{\int_{A_2} Y dP}{P(A_2)} 1_{A_2} \textit{ p.s..}$$

Ici encore, la fonction $\psi(X)$ *est un élément de* $E(Y|X)$ *dès que* ψ *est une fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$ *et t.q.* $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$ *pour* $i = 1, 2$ *(un exemple possible est donc* $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$ *pour* $i = 1, 2$ *et* $\psi(x) = 0$ *pour* $x \notin \{x_1, x_2\}$*).*

3. On suppose que X est une v.a. prenant p.s. ses valeurs dans un ensemble dénombrable $\{x_n, n \in \mathbb{N}^\star\}$ avec $P(X = x_n) \neq 0$ pour tout $n \in \mathbb{N}^\star$. Donner un élément de $E(Y|X)$.

Corrigé – *On peut supposer que les x_n sont différents deux à deux. Pour $n \in \mathbb{N}^*$, on pose $A_n = \{X = x_n\}$. Les ensembles A_n sont disjoints deux à deux, $P(A_n) > 0$ pour tout $n \in \mathbb{N}^*$ et $\sum_{n=1}^{\infty} P(A_n) = 1$.*

On utilise encore la proposition 11.11. Soit $Z \in E(Y|X)$ (Noter que, comme on sait que $E(Y|X)$ est non vide, il existe $Z \in E(Y|X)$). Il existe alors ψ, fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application φ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \tag{11.11}$$

On a donc

$$Z = \sum_{n\in\mathbb{N}^*} \psi(x_n)1_{A_n} \textit{ p.s.}.$$

Soit $(\alpha_n)_{n\in\mathbb{N}^*} \subset \mathbb{R}$ une suite bornée de $\mathbb{R}$. Dans l'égalité (11.11), on prend pour φ une fonction borélienne bornée de $\mathbb{R}$ dans $\mathbb{R}$ t.q. $\varphi(x_n) = \alpha_n$ pour tout $n \in \mathbb{N}^*$ (un tel φ existe, on peut prendre, par exemple, $\varphi(x) = 0$ si x est différent de tous les x_n), on obtient :

$$\sum_{n\in\mathbb{N}^*} \psi(x_n)\alpha_n P(A_n) = \sum_{n\in\mathbb{N}^*} \alpha_n \int_{A_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n\in\mathbb{N}^*} \subset \mathbb{R}$. Comme $P(A_n) > 0$ pour tout $n \in \mathbb{N}^*$, on en déduit que $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$, pour tout $n \in \mathbb{N}^*$, et donc que

$$Z = \sum_{n\in\mathbb{N}^*} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n} \textit{ p.s.}.$$

Enfin, ici encore, la fonction $\psi(X)$ est un élément de $E(Y|X)$ dès que ψ est une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$ et t.q. $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$ pour tout $n \in \mathbb{N}^*$. Un exemple possible est donc $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$ pour tout $n \in \mathbb{N}^*$ et $\psi(x) = 0$ pour $x \notin \{x_n, n \in \mathbb{N}^*\}$. Cet exemple donne la fonction

$$\sum_{n\in\mathbb{N}^*} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n}.$$

Exercice 11.4 (Égalité d'espérances conditionnelles) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{B}$ une sous–tribu de $\mathcal{A}$, X une v.a.r. intégrable et V une v.a.r. $\mathcal{B}$-mesurable bornée. Montrer que $E(XV|\mathcal{B}) = E(X|\mathcal{B})V$ p.s..

Corrigé – *Comme X est intégrable et que V est bornée, la v.a.r. XV est intégrable et donc $E(XV|\mathcal{B})$ est bien définie.*

On pose $Z = E(X|\mathcal{B})$ (en toute rigueur, on choisit plutôt un élément de $E(X|\mathcal{B})$). Soit U une v.a.r. $\mathcal{B}$-mesurable et bornée. La v.a.r. UV est aussi $\mathcal{B}$-mesurable et bornée, on a donc

$$E((XV)U) = E(X(UV)) = E(Z(UV)) = E((ZV)U).$$

Comme ZV *est* $\mathcal{B}$*-mesurable (et que* U *est arbitraire), ceci prouve que* $E(XV|\mathcal{B}) = ZV$ *p.s. (plus précisément,* ZV *est un élément de* $E(XV|\mathcal{B})$*). On a donc bien montré que* $E(XV|\mathcal{B}) = E(X|\mathcal{B})V$ *p.s.*.

Exercice 11.5 (Espérance conditionnelle et indépendance) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et X une v.a.r. intégrable.

1. Soit Y une v.a.r. indépendante de X. Montrer que $E(X|Y) = E(X)$ p.s..

   Corrigé – *Soit* U *une v.a.r.* $\sigma(Y)$*-mesurable bornée. Selon le théorème 11.10, il existe* φ*, fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*, t.q.* $U = \varphi(Y)$*. Comme* $\sigma(\varphi(Y)) \subset \sigma(Y)$*, les v.a.r.* X *et* $\varphi(Y)$ *sont aussi indépendantes, on a donc*
   $$E(X\varphi(Y)) = E(X)E(\varphi(Y)) = E(E(X)\varphi(Y)).$$
   Comme le fonctions constantes sont $\sigma(Y)$*-mesurables, on en déduit que* $E(X|Y) = E(X)$ *p.s. (en toute rigueur, on a démontré que la fonction constante et égale à* $E(X)$ *est un élément de* $E(X|Y)$*).*

2. Soit $\mathcal{B}$ une sous tribu de $\mathcal{A}$. On suppose que $\sigma(X)$ et $\mathcal{B}$ sont des tribus indépendantes. Montrer que $E(X|\mathcal{B}) = E(X)$ p.s..

   Corrigé – *Cette question contient la question précédente. En effet, on a* $E(X|Y) = E(X|\sigma(Y))$ *et, par définition, l'indépendance de* X *et* Y *est l'indépendance des tribus* $\sigma(X)$ *et* $\sigma(Y)$*. La démonstration est très voisine de la précédente.*

   Soit U *une v.a.r* $\mathcal{B}$*-mesurable bornée. Comme* U *est* $\mathcal{B}$*-mesurable, on a* $\sigma(U) \subset \mathcal{B}$*. On en déduit que* X *et* U *sont indépendantes. On a donc*
   $$E(XU) = E(X)E(U) = E(E(X)U).$$
   On en déduit bien que $E(X|\mathcal{B}) = E(X)$ *p.s.*.

Exercice 11.6 (Espérance conditionnelle d'une v.a.r. appartenant à $\mathcal{L}^p$)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{G}$ une sous–tribu de $\mathcal{A}$ et Y une v.a.r. intégrable. On pose $Z = E(Y|\mathcal{G})$ (plus précisément, on confond ici, comme d'habitude, la classe $E(Y|\mathcal{G})$ avec l'un de ses éléments).

1. On suppose, dans cette question, qu'il existe $M \in \mathbb{R}$ t.q. $|Y| \le M$ p.s.. Montrer que $|Z| \le M$ p.s.. et que $E(Z\Phi) = E(Y\Phi)$ pour tout Φ $\mathcal{G}$–mesurable et intégrable.

   Corrigé – *On pose*
   $$U = 1_{\{Z>M\}} - 1_{\{Z<-M\}}.$$
   La v.a.r. U *est* $\mathcal{G}$*-mesurable (car* Z *est* $\mathcal{G}$*-mesurable) bornée. On a donc* $E(ZU) = E(YU)$.
   Comme $E(ZU) = E(|ZU|)$ *et*
   $$E(YU) = \int_\Omega YU dP \le \int_\Omega |Y||U| dP \le ME(|U|),$$

on a $E(|ZU|) \le ME(|U|)$ *et donc* $E((|Z|-M)|U|) \le 0$*, c'est-à-dire*

$$\int_{|Z|>M} (|Z|-M)dP \le 0.$$

On en déduit que $P(\{|Z|>M\}) = 0$*, et donc* $|Z| \le M$ *p.s..*

Soit Φ *une v.a.r.* $\mathcal{G}$*–mesurable et intégrable. Pour* $n \in \mathbb{N}^*$*, on pose*

$$\Phi_n = T_n(\Phi) \text{ avec } T_n(s) = \max\{-n, \min\{n, s\}\} \text{ (pour } s \in \mathbb{R}).$$

La v.a.r. Φ_n *est* $\mathcal{G}$*–mesurable bornée, on a donc, pour tout* $n \in \mathbb{N}^*$*,*

$$E(Z\Phi_n) = E(Y\Phi_n). \tag{11.12}$$

Comme $\Phi_n \to \Phi$ *p.s. quand* $n \to +\infty$ *et que* $|Z\Phi_n| \le M|\Phi|$ *p.s. et* $|Y\Phi_n| \le M|\Phi|$ *p.s., le théorème de convergence dominée nous permet de passer à la limite dans (11.12) quand* $n \to +\infty$*. On obtient*

$$E(Z\Phi) = E(Y\Phi).$$

2. Soit $p \in]1, \infty[$ et $q = \frac{p}{p-1}$. On suppose que $|Y|^p$ est intégrable, montrer que $|Z|^p$ est intégrable et que $E(Z\Phi) = E(Y\Phi)$ pour tout Φ $\mathcal{G}$–mesurable et t.q. $|\Phi|^q$ soit intégrable.

Corrigé – *On utilise la fonction borélienne* T_n *définie à la question précédente et on pose* $\text{sign}(s) = 1$ *si* $s > 0$, $\text{sign}(s) = -1$ *si* $s < 0$ *et* $\text{sign}(0) = 0$ *(c'est la fonction signe définie par* (4.76)*.*.

Soit $n \in \mathbb{N}^*$*. On pose*

$$Z_n = \text{sign}(Z)|T_n(Z)|^{p-1}.$$

La v.a.r. Z_n *est* $\mathcal{G}$*-mesurable bornée, on a donc* $E(ZZ_n) = E(YZ_n)$*. En utilisant l'inégalité de Hölder avec p et q (ce qui est possible car* $|Y|^p$ *est intégrable et* $|Z_n|^q$ *est intégrable car bornée), ceci donne*

$$\int_{|Z|\le n} |Z|^p dP = E(ZZ_n) = E(YZ_n) \le \|Y\|_p \Big(\int_{|Z|\le n} |Z|^p dP \Big)^{1-\frac{1}{p}}.$$

On en déduit

$$\int_{|Z|\le n} |Z|^p dP \le \|Y\|_p^p.$$

Le théorème de convergence monotone nous permet alors de conclure que $|Z|^p$ *est intégrable et que* $\|Z\|_p \le \|Y\|_p$*.*

Soit Φ *une v.a.r.* $\mathcal{G}$*–mesurable. On suppose que* $|\Phi|^q$ *est intégrable.*

Pour $n \in \mathbb{N}^*$*, on pose* $\Phi_n = T_n(\Phi)$*. La v.a.r.* Φ_n *est* $\mathcal{G}$*–mesurable bornée, on a donc, pour tout* $n \in \mathbb{N}^*$*,*

$$E(Z\Phi_n) = E(Y\Phi_n). \tag{11.13}$$

On a $\Phi_n \to \Phi$ *p.s. quand* $n \to +\infty$ *et* $|Z\Phi_n| \le |Z\Phi|$ *p.s.,* $|Y\Phi_n| \le |Y\Phi|$ *p.s.. Les fonctions* $Z\Phi$ *et* $Y\Phi$ *sont intégrables (car* $|Z|^p, |Y|^p$ *sont intégrables et* $|\Phi|^q$ *est intégrable). Le théorème de convergence dominée nous permet alors de passer à la limite dans (11.13) quand* $n \to +\infty$*. On obtient* $E(Z\Phi) = E(Y\Phi)$*.*

Exercice 11.7 (Calcul de $E(\exp(XY)|X)$ **si** $Y \sim \mathcal{N}(0,1)$**)** Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a. réelles indépendantes. On suppose que Y suit une loi gaussienne centrée réduite et que $E(\exp(X^2/2)) < \infty$. Montrer que $\exp(XY)$ est intégrable et déterminer $E(\exp(XY)|X)$.

Corrigé – *La v.a.r.* $\exp(XY)$ *est positive. On calcule* $\int_\Omega e^{XY} dP$ *en utilisant l'indépendance de* X *et* Y *(et le théorème 9.28, qui donne que* $P_{(X,Y)} = P_X \otimes P_Y$*) et le fait que* $Y \sim \mathcal{N}(0,1)$ *:*

$$\int_\Omega e^{XY} dP = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{xy} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy dP_X(x).$$

En remarquant que $xy - \frac{y^2}{2} = -\frac{1}{2}(x-y)^2 + \frac{1}{2}x^2$ *on obtient :*

$$\int_\Omega e^{XY} dP = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\frac{-(x-y)^2}{2}})dy \right) e^{\frac{x^2}{2}} dP_X(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\frac{-z^2}{2}} dz \right) e^{\frac{x^2}{2}} dP_X(x),$$

et donc

$$\int_\Omega e^{XY} dP = \int_{\mathbb{R}} e^{\frac{x^2}{2}} dP_X(x) = E(e^{\frac{X^2}{2}}) < +\infty,$$

ce qui donne que e^{XY} *est une v.a.r. intégrable.*

Selon la proposition 11.11 on cherche un élément de $E(e^{XY}|X)$ *sous la forme* $\psi(X)$ *où* ψ *est une fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$*, t.q.* $Z = \psi(X)$*,* $\psi(X)$ *est intégrable et, pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$*,borélienne bornée,*

$$E(\psi(X)\varphi(X)) = E(e^{XY}\varphi(X)).$$

Soit φ *une application borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*, on calcule* $E(e^{XY}\varphi(X))$ *en utilisant, comme précédemment, l'indépendance de* X *et* Y *et le fait que* $Y \sim \mathcal{N}(0,1)$ *:*

$$E(e^{XY}\varphi(X)) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{xy} \varphi(x) e^{-\frac{y^2}{2}} dy dP_X(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\frac{-(x-y)^2}{2}} dy \right) e^{\frac{x^2}{2}} \varphi(x) dP_X(x).$$

On a donc

$$E(e^{XY}\varphi(X)) = \int_{\mathbb{R}} e^{\frac{x^2}{2}} \varphi(x) dP_X(x) = E(e^{-\frac{X^2}{2}} \varphi(X)).$$

Ceci nous montre que $e^{-\frac{X^2}{2}}$ *est un élément de* $E(\ e^{XY}|\ X)$ *et donc (comme on confond* $E(\ e^{XY}\ |\ X)$ *avec l'un de des éléments)* $E(e^{XY}|X) = e^{-\frac{X^2}{2}}$ *p.s..*

Exercice 11.8 (Espérance selon une somme de v.a.r.i.i.d) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

1. Soit X et Y deux v.a.r. intégrables. Montrer que $E(X|X+Y) + E(Y|X+Y) = X + Y$ p.s.. On suppose maintenant que X et Y sont indépendantes et de même loi. Montrer

que

$$\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) = \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \frac{\mathrm{X}+\mathrm{Y}}{2} \text{ p.s..}$$

Corrigé – *Par linéarité de l'espérance conditionnelle (voir la remarque 11.5), on a*

$$\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) + \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \mathrm{E}(\mathrm{X}+\mathrm{Y}|\mathrm{X}+\mathrm{Y}) \textit{ p.s.,}$$

et on en déduit bien que

$$\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) + \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \mathrm{X}+\mathrm{Y} \textit{ p.s..}$$

On utilise maintenant la proposition 11.11. Soit ψ *fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$ *telle que* $\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) = \psi(\mathrm{X}+\mathrm{Y})$ *p.s.. Si* X *et* Y *sont indépendantes et de même loi, on va montrer que* $\mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \psi(\mathrm{X}+\mathrm{Y})$ *p.s..*

Soit φ *fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*. Comme* $\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) = \psi(\mathrm{X}+\mathrm{Y})$ *p.s., on a*

$$\mathrm{E}(\mathrm{X}\varphi(\mathrm{X}+\mathrm{Y})) = \mathrm{E}(\psi(\mathrm{X}+\mathrm{Y})\varphi(\mathrm{X}+\mathrm{Y})).$$

En notant m la loi commune à X *et* Y*, on a aussi, grâce à l'indépendance de* X *et* Y *(qui donne* $\mathrm{P}_{(\mathrm{X},\mathrm{Y})} = m \otimes m$*),*

$$\mathrm{E}(\mathrm{X}\varphi(\mathrm{X}+\mathrm{Y})) = \int_{\mathbb{R}^2} x\varphi(x+y)dm(x)dm(y)$$
$$= \int_{\mathbb{R}^2} y\varphi(x+y)dm(x)dm(y) = \mathrm{E}(\mathrm{Y}\varphi(\mathrm{X}+\mathrm{Y})).$$

On a donc $\mathrm{E}(\mathrm{Y}\varphi(\mathrm{X}+\mathrm{Y})) = \mathrm{E}(\psi(\mathrm{X}+\mathrm{Y})\varphi(\mathrm{X}+\mathrm{Y}))$*, ce qui prouve que* $\mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \psi(\mathrm{X}+\mathrm{Y})$ *p.s. et donc*

$$\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) = \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) \textit{ p.s..}$$

Comme $\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) + \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \mathrm{X}+\mathrm{Y}$ *p.s., on obtient, finalement,*

$$\mathrm{E}(\mathrm{X}|\mathrm{X}+\mathrm{Y}) = \mathrm{E}(\mathrm{Y}|\mathrm{X}+\mathrm{Y}) = \frac{\mathrm{X}+\mathrm{Y}}{2} \textit{ p.s..}$$

2. Soit $n \in \mathbb{N}^*$ et $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des v.a.r. indépendantes, de même loi et intégrables. On note

$$\mathrm{S}_n = \sum_{k=1}^{n} \mathrm{X}_k.$$

Montrer que $\mathrm{E}(\mathrm{X}_1|\mathrm{S}_n) = \mathrm{S}_n/n$ p.s..

Corrigé – *La méthode donnée à la question précédente se généralise facilement. On remarque tout d'abord (par linéarité de l'espérance) que*

$$\sum_{i=1}^{n} \mathrm{E}(\mathrm{X}_i|\mathrm{S}_n) = \mathrm{E}(\sum_{i=1}^{n} \mathrm{X}_i|\mathrm{S}_n) = \mathrm{E}(\mathrm{S}_n|\mathrm{S}_n) = \mathrm{S}_n \textit{ p.s..}$$

Puis, pour $i \neq j$*, en utilisant que* $\mathrm{X}_1, \ldots, \mathrm{X}_n$ *sont des v.a.r. indépendantes et de même loi (ce qui donne que* $\mathrm{P}_{(\mathrm{X}_1,\ldots,\mathrm{X}_n)} = m \otimes \ldots \otimes m$ *si m est la loi commune aux* X_i*), on montre que*

$$\mathrm{E}(\mathrm{X}_i|\mathrm{S}_n) = \mathrm{E}(\mathrm{X}_j|\mathrm{X}_n) \textit{ p.s..}$$

On en déduit alors que $\mathrm{E}(\mathrm{X}_i|\mathrm{S}_n) = \mathrm{S}_n/n$ *p.s., pour tout* $i = 1, \ldots, n$*.*

Exercice 11.9 (Une condition nécessaire et suffisante pour avoir $X = Y$ p.s.) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r. intégrables. On suppose que

$$E(X|Y) = Y \text{ p.s. et } E(Y|X) = X \text{ p.s..}$$

1. Soit $c \in \mathbb{R}$. Montrer que

$$E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) = E\Big((Y-X)1_{\{X>c\geq Y\}}\Big) \leq 0.$$

   En déduire que $E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) = 0$, puis que $P(\{X > c \geq Y\}) = 0$.

   Corrigé – *On remarque que* $(X-Y)1_{\{X>c,Y>c\}} + (X-Y)1_{\{X>c\geq Y\}} = (X-Y)1_{\{X>c\}}$. *On a donc*

$$\begin{aligned}&E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) - E\Big((Y-X)1_{\{X>c\geq Y\}}\Big)\\ &= E((X-Y)1_{\{X>c\}}) = E(X1_{\{X>c\}}) - E(Y1_{\{X>c\}}).\end{aligned}$$

   Comme $E(Y|X) = X$ *p.s. et que* $1_{\{X>c\}}$ *est une v.a.r.* $\sigma(X)$*-mesurable bornée, on a* $E(Y1_{\{X>c\}}) = E(X1_{\{X>c\}})$ *et ceci donne alors*

$$E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) - E\Big((Y-X)1_{\{X>c\geq Y\}}\Big) = 0.$$

   Comme $Y - X \leq 0$ *sur* $\{X > c \geq Y\}$, *on a bien, finalement*

$$E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) = E\Big((Y-X)1_{\{X>c\geq Y\}}\Big) \leq 0.$$

   En changeant X, Y *en* Y, X, *on montre aussi que*

$$E\Big((Y-X)1_{\{X>c,Y>c\}}\Big) \leq 0.$$

   On en déduit alors que $E\Big((X-Y)1_{\{X>c,Y>c\}}\Big) = 0$ *et donc que*

$$E\Big((Y-X)1_{\{X>c\geq Y\}}\Big) = 0.$$

   Enfin, comme $Y - X < 0$ *sur* $\{X > c \geq Y\}$, *l'égalité précédente permet de conclure que* $P(\{X > c \geq Y\}) = 0$.

2. Montrer que $X = Y$ p.s..

   Corrigé – *Pour* $c \in \mathbb{R}$, *on pose* $A_c = \{X > c \geq Y\}$. *On a donc* $P(A_c) = 0$ *pour tout* $c \in \mathbb{R}$.

   On remarque maintenant que $\{X - Y > 0\} = \cup_{c\in\mathbb{Q}} A_q$. *Par* σ*-sous additivité de* P *on en déduit que*

$$P(\{X-Y>0\}) \leq \sum_{c\in\mathbb{Q}} P(A_c) = 0.$$

   Ici encore, en changeant X, Y *en* Y, X, *on montre aussi que* $P(\{Y - X > 0\}) = 0$ *et donc que* $X = Y$ *p.s..*

3. on suppose maintenant que X et Y sont de carré intégrables. Montrer qu'une démonstration (beaucoup) plus directe de la question 2 est possible en calculant $E((X - Y)^2)$.

Corrigé – *Comme* $\mathrm{E}(\mathrm{X}|\mathrm{Y}) = \mathrm{Y}$ *p.s., on a* $\mathrm{E}(\mathrm{XY}) = \mathrm{E}(\mathrm{Y}^2)$*. De même, comme* $\mathrm{E}(\mathrm{Y}|\mathrm{X}) = \mathrm{X}$ *p.s., on a* $\mathrm{E}(\mathrm{YX}) = \mathrm{E}(\mathrm{X}^2)$*. On en déduit que*

$$\mathrm{E}((\mathrm{X}-\mathrm{Y})^2) = \mathrm{E}(\mathrm{X}^2) + \mathrm{E}(\mathrm{Y}^2) - \mathrm{E}(\mathrm{XY}) - \mathrm{E}(\mathrm{YX}) = 0,$$

et donc que $\mathrm{X} = \mathrm{Y}$ *p.s..*

N.B. Le cas où X et Y sont seulement intégrables (traité dans les questions 1 et 2) peut aussi se faire avec la question 3 en tronquant les v.a.r. X et Y.

Exercice 11.10 (Espérance du produit et produit des espérances) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et X, Y deux v.a. intégrables t.q. XY est intégrable et $\mathrm{E}(\mathrm{X}|\mathrm{Y}) = \mathrm{E}(\mathrm{X})$ p.s.. Montrer que $\mathrm{E}(\mathrm{XY}) = \mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{Y})$.

Corrigé – *Grâce à la proposition 11.11 on a, pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$*,borélienne bornée,*

$$\mathrm{E}(\mathrm{E}(\mathrm{X}|\mathrm{Y})\varphi(\mathrm{Y})) = \mathrm{E}(\mathrm{X}\varphi(\mathrm{Y}))$$

Comme $\mathrm{E}(\mathrm{X}|\mathrm{Y}) = \mathrm{E}(\mathrm{X})$ *p.s., on en déduit, pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$*,borélienne bornée,*

$$\mathrm{E}(\mathrm{X})\mathrm{E}(\varphi(\mathrm{Y})) = \mathrm{E}(\mathrm{X}\varphi(\mathrm{Y})). \tag{11.14}$$

Soit $n \in \mathbb{N}^*$*. Pour* $s \in \mathbb{R}$*, on pose* $\mathrm{T}_n(s) = \max\{-n, \min\{s, n\}\}$*. La fonction* T_n *est borélienne (car continue) bornée (par n) de* $\mathbb{R}$ *dans* $\mathbb{R}$*. On peut donc utiliser (11.14) avec* $\varphi = \mathrm{T}_n$*. On obtient* $\mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{T}_n(\mathrm{Y})) = \mathrm{E}(\mathrm{X}\mathrm{T}_n(\mathrm{Y}))$*.*

Comme Y *est intégrable, on a, par convergence dominée,* $\lim_{n\to+\infty} \mathrm{E}(\mathrm{T}_n(\mathrm{Y})) = \mathrm{E}(\mathrm{Y})$ *(noter que* $|\mathrm{T}_n(\mathrm{Y})| \le |\mathrm{Y}|$*).*

Comme XY *est intégrable (et c'est uniquement ici que cette hypothèse est utilisée), on a, par convergence dominée,*

$$\lim_{n\to+\infty} \mathrm{E}(\mathrm{X}\mathrm{T}_n(\mathrm{Y})) = \mathrm{E}(\mathrm{XY})$$

(noter que $|\mathrm{X}\mathrm{T}_n(\mathrm{Y})| \le |\mathrm{XY}|$*).*

En passant à limite quand $n \to +\infty$ *sur l'égalité* $\mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{T}_n(\mathrm{Y})) = \mathrm{E}(\mathrm{X}\mathrm{T}_n(\mathrm{Y}))$*, on a donc* $\mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{Y}) = \mathrm{E}(\mathrm{XY})$*.*

Exercice 11.11 (Égalité de lois donne égalité d'espérances conditionnelles) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et $\mathrm{X}, \mathrm{Y}, \mathrm{Z}$ trois v.a.r.. On suppose que X et Y sont intégrables et que $(\mathrm{X}, \mathrm{Z}) \sim (\mathrm{Y}, \mathrm{Z})$.

1. Montrer que $\mathrm{E}(\mathrm{X}|\mathrm{Z}) = \mathrm{E}(\mathrm{Y}|\mathrm{Z})$ p.s..

Corrigé – *On note m la loi commune à* (X, Z) *et* (Y, Z)*. Soit* ψ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*. On a*

$$\mathrm{E}(\mathrm{X}\psi(\mathrm{Z})) = \int_{\mathbb{R}^2} x\psi(z)dm(x,z),$$

et, de même

$$\mathrm{E}(\mathrm{Y}\psi(\mathrm{Z})) = \int_{\mathbb{R}^2} y\psi(z)dm(y,z) = \int_{\mathbb{R}^2} x\psi(z)dm(x,z).$$

On a donc $\mathrm{E}(\mathrm{X}\psi(\mathrm{Z})) = \mathrm{E}(\mathrm{Y}\psi(\mathrm{Z}))$*, ce qui donne bien* $\mathrm{E}(\mathrm{X}|\mathrm{Z}) = \mathrm{E}(\mathrm{Y}|\mathrm{Z})$ *p.s..*

2. Soit f une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que $f(X)$ et $f(Y)$ sont intégrables (ou que f est à valeurs dans $\mathbb{R}_+$). Montrer que $E[f(X)|Z] = E[f(Y)|Z]$ p.s..

Corrigé – *On suppose que* $f(X)$ *et* $f(Y)$ *sont intégrables. On note toujours* m *la loi commune à* (X, Z) *et* (Y, Z)*. On a alors, pour toute fonction* ψ *borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*,*

$$E(f(X)\psi(Z)) = \int_{\mathbb{R}^2} f(x)\psi(z)dm(x,z) = E(f(Y)\psi(Z)).$$

On a donc $E(f(X)\psi(Z)) = E(f(Y)\psi(Z))$*, ce qui donne bien*

$$E(f(X)|Z) = E(f(Y)|Z) \text{ p.s.}.$$

Noter que l'intégrabilité de X *et* Y *est inutile pour cette question. C'est l'intégrabilité de* $f(X)$ *et* $f(Y)$ *qui a été utilisée.*

Si on retire l'hypothèse que $f(X)$ *et* $f(Y)$ *sont intégrables mais que* f *est à valeurs dans* $\mathbb{R}_+$*, le même raisonnement permet de conclure en prenant des fonctions* ψ *boréliennes positives.*

Exercice 11.12 (Convergence faible et espérances conditionnelles) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{B}$ une sous–tribu de $\mathcal{A}$, X une v.a. intégrable et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a. intégrables. On suppose que $X_n \to X$ faiblement dans $L^1(\Omega, \mathcal{A}, P)$ quand $n \to +\infty$ (ce qui est équivalent à dire que, pour tout v.a.r. U bornée, on a $\lim_{n\to+\infty} E(X_nU) = E(XU)$).

1. Montrer que pour toute v.a.r. U, $\mathcal{B}$-mesurable et bornée, on a

$$\lim_{n\to+\infty} E(E(X_n|\mathcal{B})U) = E(E(X|\mathcal{B})U).$$

Corrigé – *Soit* U *une v.a.r.* $\mathcal{B}$*-mesurable et bornée. On a*

$$E(E(X_n|\mathcal{B})U) = E(X_n|U) \text{ et } E(E(X|\mathcal{B})U) = E(X|U).$$

Comme $\lim_{n\to+\infty} E(X_nU) = E(XU)$*, on en déduit que*

$$\lim_{n\to+\infty} E(E(X_n|\mathcal{B})U) = E(E(X|\mathcal{B})U).$$

2. Montrer que

$$E(X_n|\mathcal{B}) \to E(X|\mathcal{B}) \text{ faiblement dans } L^1(\Omega, \mathcal{A}, P) \text{ quand } n \to +\infty.$$

Corrigé – *Soit* U *une v.a.r. bornée. On pose* $V = E(U|\mathcal{B})$ *de sorte que* V *est une v.a.r.* $\mathcal{B}$*-mesurable bornée (voir l'exercice 11.6). Comme* $E(X_n|\mathcal{B})$ *(pour tout* $n \in \mathbb{N}$*) et* $E(X|\mathcal{B})$ *sont des v.a.r.* $\mathcal{B}$*-mesurable et intégrable, on a (voir aussi l'exercice 11.6)*

$$E(UE(X_n|\mathcal{B})) = E(VE(X_n|\mathcal{B})) \text{ pour tout } n \in \mathbb{N}$$

et

$$E(UE(X|\mathcal{B})) = E(VE(X|\mathcal{B})).$$

D'après la première question, on a $\lim_{n\to+\infty} \mathrm{E}(\mathrm{E}(\mathrm{X}_n|\mathcal{B})\mathrm{V}) = \mathrm{E}(\mathrm{E}(\mathrm{X}|\mathcal{B})\mathrm{V})$*, on en déduit que*

$$\lim_{n\to+\infty} \mathrm{E}(\mathrm{E}(\mathrm{X}_n|\mathcal{B})\mathrm{U}) = \mathrm{E}(\mathrm{E}(\mathrm{X}|\mathcal{B})\mathrm{U}).$$

Ce qui donne que $\mathrm{E}(\mathrm{X}_n|\mathcal{B}) \to \mathrm{E}(\mathrm{X}|\mathcal{B})$ *faiblement dans* $\mathrm{L}^1(\Omega,\mathcal{A},\mathrm{P})$ *quand* $n \to +\infty$.

Exercice 11.13 (Minoration d'une espérance conditionnelle) Soit $(\Omega,\mathcal{A},\mathrm{P})$ un espace probabilisé, $\mathcal{B}$ une sous–tribu de $\mathcal{A}$ et Y est une v.a.r. positive. Soit U une v.a.r. positive, $\mathcal{B}$-mesurable et t.q. pour toute v.a.r. positive $\mathcal{B}$-mesurable Z on ait $\mathrm{E}(\mathrm{YZ}) \ge \mathrm{E}(\mathrm{UZ})$. Montrer que $\mathrm{E}(\mathrm{Y}|\mathrm{B}) \ge \mathrm{U}$ p.s..

Corrigé – *Soit* $\mathrm{A} = \{\mathrm{E}(\mathrm{Y}|\mathrm{B}) < \mathrm{U}\}$. *Comme* U *et* $\mathrm{E}(\mathrm{Y}|\mathrm{B})$ *sont* $\mathcal{B}$*-mesurable, la v.a.r.* 1_A *est* $\mathcal{B}$*-mesurable. L'hypothèse donne alors* $\mathrm{E}(\mathrm{Y}1_\mathrm{A}) \ge \mathrm{E}(\mathrm{U}1_\mathrm{A})$. *Comme* $\mathrm{E}(\mathrm{Y}1_\mathrm{A}) = \mathrm{E}(\mathrm{E}(\mathrm{Y}|\mathcal{B})1_\mathrm{A})$*, on en déduit*

$$\mathrm{E}((\mathrm{E}(\mathrm{Y}|\mathcal{B}) - \mathrm{U})1_\mathrm{A}) = \mathrm{E}(\mathrm{E}(\mathrm{Y}|\mathcal{B})1_\mathrm{A}) - \mathrm{E}(\mathrm{U}1_\mathrm{A}) \ge 0.$$

Comme $\mathrm{E}(\mathrm{Y}|\mathcal{B}) - \mathrm{U} < 0$ *sur* A*, on a donc* $\mathrm{P}(\mathrm{A}) = 0$*, ce qui prouve que* $\mathrm{E}(\mathrm{Y}|\mathrm{B}) \ge \mathrm{U}$ *p.s..*

Exercice 11.14 (Dépendance linéaire et dépendance non linéaire) Soit $(\Omega,\mathcal{A},\mathrm{P})$ un espace probabilisé et X,Y deux v.a.r. de carré intégrable et non constantes (on rappelle que la v.a.r. X est dite constante s'il existe $a \in \mathbb{R}$ t.q. $\mathrm{X} = a$ p.s.).

1. (Dépendance linéaire.) On pose $\overline{\mathrm{X}} = \frac{\mathrm{X}-\mathrm{E}(\mathrm{X})}{\sqrt{\mathrm{Var}(\mathrm{X})}}$ et $\overline{\mathrm{Y}} = \frac{\mathrm{Y}-\mathrm{E}(\mathrm{Y})}{\sqrt{\mathrm{Var}(\mathrm{Y})}}$.

 (a) Montrer que $|\mathrm{Cov}(\overline{\mathrm{X}},\overline{\mathrm{Y}})| \le 1$.

 Corrigé – *Comme* $Var(\mathrm{X}) = \mathrm{E}((\mathrm{X}-\mathrm{E}(\mathrm{X}))^2)$*, on a, avec l'inégalité de Cauchy-Schwarz,*

$$Cov(\overline{\mathrm{X}},\overline{\mathrm{Y}}) = \frac{\mathrm{E}((\mathrm{X}-\mathrm{E}(\mathrm{X}))(\mathrm{Y}-\mathrm{E}(\mathrm{Y}))}{\sqrt{Var(\mathrm{X})Var(\mathrm{Y})}} \le \frac{\sqrt{Var(\mathrm{X})Var(\mathrm{Y})}}{\sqrt{Var(\mathrm{X})Var(\mathrm{Y})}} = 1$$

 (b) Montrer que $|\mathrm{Cov}(\overline{\mathrm{X}},\overline{\mathrm{Y}})| = 1$ si et seulement si il existe $\alpha,\beta \in \mathbb{R}$, $\alpha \ne 0$, t.q. $\mathrm{Y} = \alpha\mathrm{X} + \beta$.

 Corrigé – *Si* $Cov(\overline{\mathrm{X}},\overline{\mathrm{Y}}) = 1$*, on doit avoir une égalité dans l'inégalité de Cauchy-Schwarz utilisée à la question précédente. Les v.a.r.* $\mathrm{X}-\mathrm{E}(\mathrm{X})$ *et* $\mathrm{Y}-\mathrm{E}(\mathrm{Y})$ *sont alors colinéaires. Il existe* $a,b \in \mathbb{R}$ *t.q.* $(a,b) \ne (0,0)$ *t.q.*

$$a(\mathrm{X}-\mathrm{E}(\mathrm{X})) + b(\mathrm{Y}-\mathrm{E}(\mathrm{Y})) = 0 \textit{ p.s.}.$$

 Comme la v.a.r. X *est non constante, on a* $b \ne 0$ *et, avec* $\alpha = -a/b \ne 0$ *et* $\beta = \mathrm{E}(\mathrm{Y}) + (a/b)\mathrm{E}(\mathrm{X})$,

$$\mathrm{Y} = \alpha\mathrm{X} + \beta.$$

 (c) Donner un exemple pour lequel $\mathrm{Y} = f(\mathrm{X})$ (avec f fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) et $\mathrm{Cov}(\overline{\mathrm{X}},\overline{\mathrm{Y}}) = 0$.

Corrigé – *Voici un exemple simple. On prend* $\Omega = \{1,2,3\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ *et* P *définie par* $P(\{i\}) = 1/3$ *pour* $i = 1,2,3$. *On définit* X *et* Y *en posant*

$$X(1) = 1,\, X(2) = 0,\, X(3) = -1,$$
$$Y(1) = 1,\, Y(2) = -2,\, Y(3) = 1.$$

Les v.a.r. X *et* Y *sont bien de carré intégrable et non constantes. On a* $E(X) = E(Y) = 0$ *et* $Cov(X,Y) = 0$ *(et donc* $Cov(\overline{X},\overline{Y}) = 0$*). Enfin, on a* $Y = f(X)$ *pour tout fonction borélienne t.q.* $f(1) = f(-1) = 1$ *et* $f(0) = -2$.

2. (Dépendance.)

(a) On suppose que X et Y sont indépendantes. Montrer que $E(Y|X) = E(Y)$ p.s..

Corrigé – *Cette question est démontrée dans l'exercice 11.5.*

(b) Montrer qu'il existe f (fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) t.q. $Y = f(X)$ p.s. si et seulement si $E(Y|X) = Y$ p.s..

Corrigé – *Cette question est une conséquence de la proposition 11.11.*

Si il existe f *(fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$*) t.q.* $Y = f(X)$ *p.s., on a alors* $E(Y\psi(X)) = E(f(X)\psi(X))$ *pour toute fonction borélienne bornée* ψ *de* $\mathbb{R}$ *dans* $\mathbb{R}$. *Ceci montre que* $E(Y|X) = f(X)$ *p.s. et donc* $E(Y|X) = Y$ *p.s..*

Réciproquement, on sait qu'il existe toujours φ *fonction borélienne de* $\mathbb{R}$ *dans* $\mathbb{R}$ *t.q.* $E(Y|X) = \varphi(X)$ *p.s.. Si* $E(Y|X) = Y$ *p.s., on a donc* $Y = \varphi(X)$ *p.s..*

Exercice 11.15 (Lorsque $E(Y|X) = X$ **p.s.. . .)** Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et X, Y deux v.a.r. de carré intégrable.

On suppose que $E(Y|X) = X$ p.s..

1(a) Soit φ une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que la v.a.r. $\varphi(X)$ est de carré intégrable. Montrer que $\int_\Omega Y\varphi(X)dP = \int_\Omega X\varphi(X)dP$.

Corrigé – *Cette propriété a été vue dans la proposition 11.2. On peut la montrer à partir de la définition 11.1 en considérant* $T_n(\varphi(X))$ *avec*

$$T_n(s) = \min\{\max\{-n,s\},n\}.$$

En effet, pour tout $n \in \mathbb{N}$, *la v.a.r.* $T_n(\varphi(X))$ *est* $\sigma(X)$*-mesurable bornée. Comme* $E(Y|X) = X$ *p.s., on a donc*

$$\int_\Omega YT_n(\varphi(X))dP = \int_\Omega XT_n(\varphi(X))dP.$$

Comme $\varphi(X) \in L^2_{\mathbb{R}}(\Omega,\mathcal{A},P)$, *le théorème de convergence dominée permet de passer à la limite dans cette égalité quand* $n \to +\infty$. *On obtient bien* $\int_\Omega Y\varphi(X)dP = \int_\Omega X\varphi(X)dP$, *c'est-à-dire*

$$E(Y\varphi(X)) = E(X\varphi(X)).$$

(b) Montrer que $E(Y) = E(X)$ et $E(XY) = E(X^2)$.

Corrigé – *En prenant* $\varphi(s) = 1$ *pour tout* $s \in \mathbb{R}$*, la question précédente donne* $E(Y) = E(X)$*. Puis, en prenant* $\varphi(s) = s$ *(ce qui est possible car* X^2 *est intégrable), on a* $E(YX) = E(X^2)$*.*

2(a) Montrer que $E(X^2) \le E(Y^2)$.

Corrigé – *Comme* $E(XY) = E(X^2)$*, on a*

$$0 \le E((Y-X)^2) = E(Y^2) + E(X^2) - 2E(XY) = E(Y^2) - E(X^2).$$

Ce qui donne bien $E(X^2) \le E(Y^2)$*.*

On peut aussi faire cette question en remarquant que l'inégalité de Jensen donne $E(Y^2|X) \ge E(Y|X)^2$ *p.s.. On a donc* $E(Y^2|X) \ge X^2$ *p.s.. On en déduit, en particulier, que* $E(Y^2) \ge E(X^2)$*.*

(b) Montrer que $Y = X$ p.s. si et seulement si $E(Y^2) = E(X^2)$.

Corrigé – *Si* $Y = X$ *p.s., on a, bien sûr,* $E(Y^2) = E(X^2)$*. Réciproquement, si* $E(Y^2) = E(X^2)$*, la question précédente donne* $E((Y-X)^2) = 0$ *et donc* $Y = X$ *p.s..*

Exercice 11.16 (V.a. gaussien et espérance conditionnelle) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(X, Y)^t$ un v.a. gaussien de dimension 2, On note a l'espérance de X, b l'espérance de Y et D la matrice de covariance du v.a. $(X, Y)^t$ (on a donc $D_{1,1} = \mathrm{Var}(X)$, $D_{2,2} = \mathrm{Var}(Y)$ et $D_{1,2} = D_{2,1} = \mathrm{Cov}(X, Y)$). On suppose que $\mathrm{Var}(Y) > 0$.

1. Calculer $\alpha, \beta \in \mathbb{R}$ (en fonction de a, b et D) de manière à avoir $E[X] = E[\alpha + \beta Y]$ et $E[XY] = E[(\alpha + \beta Y)Y]$.

Corrigé – *On a* $E[X] = a$*,* $E[Y] = b$*,* $E[Y^2] = Var(Y) + E[Y]^2 = D_{2,2} + b^2$ *et*

$$E[XY] = E[(X - E(X))(Y - E(Y))] + E[X]E[Y] = D_{1,2} + ab.$$

On cherche donc α *et* β *t.q.*

$$\alpha + b\beta = a,$$
$$b\alpha + (D_{2,2} + b^2)\beta = D_{1,2} + ab.$$

Comme $D_{2,2} \neq 0$*, ce système de 2 équations à 2 inconnues (qui sont* α *et* β*) a bien une unique solution. Cette solution est*

$$\beta = \frac{D_{1,2}}{D_{2,2}}, \quad \alpha = a - b\frac{D_{1,2}}{D_{2,2}}.$$

Avec α et β ainsi déterminés, on définit la fonction affine l de $\mathbb{R}$ dans $\mathbb{R}$ par $l(s) = \alpha + \beta s$ pour $s \in \mathbb{R}$ et on définit la v.a.r. Z par $Z = X - l(Y)$.

2. Montrer que $(Z, Y)^t$ est un v.a. gaussien. Montrer que $\mathrm{Cov}(Z, Y) = 0$. En déduire que Z et Y sont des v.a.r. indépendantes [On pourra utiliser la question 1(b) de l'exercice 10.15].

Corrigé – *Soit $a_1, a_2 \in \mathbb{R}$. On va montrer que la v.a.r. $a_1 Z + a_2 Y$ suit une loi gaussienne (ceci montre bien que le vecteur aléatoire $(Z, Y)^t$ est un vecteur gaussien). Comme $Z = X - l(Y) = X - \alpha - \beta Y$, on a*

$$a_1 Z + a_2 Y = a_1 X + (a_2 - a_1 \beta) Y - a_1 \alpha.$$

Comme $(X, Y)^t$ est un v.a. gaussien, la v.a.r. $a_1 X + (a_2 - a_1 \beta) Y$ suit une loi gaussienne. Il existe donc $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+$ t.q. $a_1 X + (a_2 - a_1 \beta) Y \sim \mathcal{N}(m, \sigma^2)$. On a alors

$$a_1 X + (a_2 - a_1 \beta) Y - a_1 \alpha \sim \mathcal{N}(m - a_1 \alpha, \sigma^2).$$

ceci prouve que $a_1 X + (a_2 - a_1 \beta) Y - a_1 \alpha$ suit une loi gaussienne. La v.a.r. $a_1 Z + a_2 Y$ suit donc une loi gaussienne pour tout $a_1, a_2 \in \mathbb{R}$. Ceci prouve bien que le vecteur aléatoire $(Z, Y)^t$ est un vecteur gaussien.

On calcule maintenant $Cov(Z, Y)$. On rappelle que $Z = X - (\alpha + \beta Y)$. On remarque d'abord que $E[Z] = E[X] - E[\alpha + \beta Y] = 0$ (grâce à la première relation satisfaite par α et β). Puis,

$$Cov(Z, Y) = E[ZY] = E[XY] - E[(\alpha + \beta Y) Y].$$

La seconde relation satisfaite par α et β donne alors $Cov(Z, Y) = 0$.

Comme $(Z, Y)^t$ est un vecteur gaussien et que $Cov(Z, Y) = 0$, la question 1(b) de l'exercice 10.15 donne que Z et Y sont indépendantes.

3. Montrer que $E(X|Y) = l(Y)$ p.s.

Corrigé – *La v.a.r. $l(Y)$ est intégrable (car sa loi est gaussienne). Pour montrer que $E(X|Y) = l(Y)$ p.s., il suffit, d'après la proposition 11.11, de montrer que $E(l(Y)\varphi(Y)) = E(X\varphi(Y))$ pour toute application φ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne bornée.*

Soit donc φ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne bornée. Comme Z et Y sont indépendantes et que $E[Z] = 0$, on a $E[Z\varphi(Y)] = E[Z]E[\varphi(Y)] = 0$. Comme $Z = X - l(Y)$ on a donc

$$E[X\varphi(Y)] = E[l(Y)\varphi(Y)].$$

Ce qui prouve bien que $E(X|Y) = l(Y)$ p.s.

4. Calculer (en fonction de D) Var(Z).

Corrigé – *Comme $E[Z] = 0$ et $E[Z\,l(Y)] = E[Z]E[l(Y)] = 0$, on a*

$$\begin{aligned} Var(Z) = E[Z^2] = E[XZ] &= E[X^2] - E[Xl(Y)] \\ &= E[X^2] - \alpha E[X] - \beta E[XY]. \end{aligned}$$

Il suffit maintenant d'utiliser les valeurs de α et β et le fait que

$$E[X^2] = D_{1,1} + a^2,\ E[X] = a,\ E[X, Y] = D_{1,2} + ab.$$

On obtient

$$Var(Z) = D_{1,1} - \frac{D_{1,2}^2}{D_{2,2}}$$

Dans la suite, on note $\sigma = \sqrt{\mathrm{Var}(Z)}$ et, pour $a \in \mathbb{R}$ on note μ_a la loi normale $\mathcal{N}(a, \sigma^2)$.

5. Soit f une fonction borélienne bornée de $\mathbb{R}$ dans $\mathbb{R}$. Pour $a \in \mathbb{R}$, on pose $\psi(a) = \int_{\mathbb{R}} f d\mu_a$.

(a) Montrer que ψ est une fonction continue (de $\mathbb{R}$ dans $\mathbb{R}$) si $\sigma > 0$.

Corrigé – *On utilise ici le théorème 4.52. On remarque que*

$$\psi(a) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} f(x) e^{\frac{-(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} F(x,a) dx,$$

avec $F(x,a) = f(x)e^{\frac{-(x-a)^2}{2\sigma^2}}$. *La fonction* $a \mapsto F(x,a)$ *est continue, pour tout* $x \in \mathbb{R}$. *Pour montrer que* ψ *est continue, il suffit (d'après le théorème 4.52) de montrer que la fonction* $x \mapsto F(x,a)$ *est dominée, localement uniformément par rapport à a, par une fonction intégrable. Nous montrons maintenant cette domination.*

Soit $M > 0$. *Pour tout* $a \in]-M, M[$ *et tout* $x \in \mathbb{R}$, *on a* $|F(x,a)| \le g_M(x)$ *avec*

$$g_M(x) = |f(x)| e^{\frac{-((|x|-M)^+)^2}{2\sigma^2}}.$$

(La vérification de $|F(x,a)| \le g_M(x)$ *peut se faire en distinguant les cas* $x \in [-M, M]$, $x > M$ *et* $x < -M$.*) La fonction* g_M *est bien intégrable (pour la mesure de Lebesgue). Le théorème 4.52 donne alors que* ψ *est continue sur* $]-M, M[$. *Comme* M *est arbitraire, on en déduit que* ψ *est continue sur* $\mathbb{R}$.

(b) Montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s.

Corrigé – *Si* $\sigma > 0$, *la fonction* ψ *est continue, elle est donc borélienne. On remarque aussi que* ψ *est bornée (en effet, soit* $M \in \mathbb{R}_+$ *t.q.* $|f(x)| \le M$ *pour tout* $x \in \mathbb{R}$. *On a alors aussi* $|\psi(a)| \le M$ *pour tout* $a \in \mathbb{R}$*).*

Si $\sigma = 0$, *on a* $\psi(a) = f(a)$ *pour tout* $a \in \mathbb{R}$. *La fonction* ψ *est donc aussi borélienne bornée.*

Comme ψ *est borélienne,* $\psi(l(Y))$ *est donc une v.a.r.. Comme* ψ *est bornée, la v.a.r.* $\psi(l(Y))$ *est bornée et donc intégrable. Pour montrer que* $E[f(X)|\,Y] = \psi(l(Y))$ *p.s., il suffit, comme à la question précédente (cf proposition 11.11), de montrer que* $E[\psi(l(Y))\varphi(Y)] = E[f(X)\varphi(Y)]$ *pour toute application* φ *de* $\mathbb{R}$ *dans* $\mathbb{R}$, *borélienne bornée.*

Soit donc φ *borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$. *On a, comme* Z *et* Y *sont indépendantes,*

$$\begin{aligned} E[f(X)\varphi(Y)] &= E[f(Z + l(Y))\varphi(Y)] \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(z + l(y)) dP_Z(z) \right) \varphi(y) dP_Y(y). \end{aligned}$$

On utilise maintenant le fait que $Z \sim \mathcal{N}(0, \sigma^2)$, *on en déduit*

$$\int_{\mathbb{R}} f(z + l(y)) dP_Z(z) = \int_{\mathbb{R}} f(z + l(y)) d\mu_0(z).$$

le changement de variable $\overline{z} = z + l(y)$ *donne alors*

$$\int_{\mathbb{R}} f(z + l(y)) d\mu_0(z) = \int_{\mathbb{R}} f(\overline{z}) d\mu_{l(y)}(\overline{z}) = \psi(l(y)).$$

On a donc, finalement,

$$\mathrm{E}[f(\mathrm{X})\varphi(\mathrm{Y})] = \int_{\mathbb{R}} \psi(l(y))\varphi(y)d\mathrm{P}_{\mathrm{Y}}(y) = \mathrm{E}[\psi(l(\mathrm{Y}))\varphi(\mathrm{Y})].$$

Ce qui prouve bien $\mathrm{E}[f(\mathrm{X})|\mathrm{Y}] = \psi(l(\mathrm{Y}))$ *p.s.*.

Exercice 11.17 (Indépendance et espérance) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et $\mathcal{B}_1, \mathcal{B}_2$ deux sous–tribus de $\mathcal{A}$. On pose $\mathcal{B} = \sigma(\mathcal{B}_1 \cup \mathcal{B}_2)$ (c'est-à-dire que $\mathcal{B}$ est la tribu engendrée par $\mathcal{B}_1$ et $\mathcal{B}_2$). On pose aussi $\mathcal{C} = \{\mathrm{B}_1 \cap \mathrm{B}_2; \mathrm{B}_1 \in \mathcal{B}_1, \mathrm{B}_2 \in \mathcal{B}_2\}$.

1. Montrer que $\mathcal{B} = \sigma(\mathcal{C})$ (c'est-à-dire que $\mathcal{B}$ est la tribu engendrée par $\mathcal{C}$).

   Soit Y une v.a.r. intégrable. On note $\sigma(\mathrm{Y})$ la tribu engendrée par Y. On suppose que $\sigma(\mathrm{Y})$ et $\mathcal{B}_1$ sont indépendantes de $\mathcal{B}_2$.

2. Soit U_1 une v.a.r. $\mathcal{B}_1$-mesurable et bornée et U_2 une v.a.r. $\mathcal{B}_2$-mesurable et intégrable. Montrer que YU_1 et U_2 sont des v.a.r. indépendantes. En déduire que $\mathrm{YU}_1\mathrm{U}_2$ est intégrable et que $\mathrm{E}(\mathrm{YU}_1\mathrm{U}_2) = \mathrm{E}(\mathrm{YU}_1)\mathrm{E}(\mathrm{U}_2)$.

3. On pose $\mathrm{Z}_1 = \mathrm{E}[\mathrm{Y}|\mathcal{B}_1]$ (la v.a.r. Z_1 est $\mathcal{B}_1$-mesurable intégrable et $\mathrm{E}(\mathrm{YU}) = \mathrm{E}(\mathrm{Z}_1\mathrm{U})$ pour toute v.a.r. $\mathcal{B}_1$-mesurable bornée).

   (a) Montrer que $\mathrm{E}(\mathrm{Y}1_{\mathrm{C}}) = \mathrm{E}(\mathrm{Z}_1 1_{\mathrm{C}})$ pour tout $\mathrm{C} \in \mathcal{C}$.

   (b) Montrer que $\mathrm{E}(\mathrm{Y}1_{\mathrm{B}}) = \mathrm{E}(\mathrm{Z}_1 1_{\mathrm{B}})$ pour tout $\mathrm{B} \in \mathcal{B}$.

4. Montrer que $\mathrm{E}[\mathrm{Y}|\mathcal{B}] = \mathrm{E}[\mathrm{Y}|\mathcal{B}_1]$ p.s..

Exercice 11.18 (Espérance conditionnelle d'une somme) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé et $(\mathrm{X}_i)_{i\in\mathbb{N}^*}$ une suite de v.a.r.i.i.d.. On suppose que X_1 est intégrable (on a donc aussi X_i intégrable pour tout $i \in \mathbb{N}^*$). Pour $n \in \mathbb{N}^*$, on pose $\mathrm{S}_n = \mathrm{X}_1 + \cdots + \mathrm{X}_n$.

1. (Cette question reprend l'exercice 11.8.) Soit $n > 1$. Montrer que $\mathrm{E}(\mathrm{X}_1|\mathrm{S}_n) = \dfrac{\mathrm{S}_n}{n}$ p.s..

2. Soit $n \in \mathbb{N}^*$. On note $\mathcal{D}_n$ la tribu engendrée par l'ensemble des v.a.r. S_k, $k \geq n$. Montrer que $\mathrm{E}(\mathrm{X}_1|\mathcal{D}_n) = \mathrm{E}(\mathrm{X}_1|\mathrm{S}_n) = \frac{\mathrm{S}_n}{n}$ p.s.. [On pourra utiliser l'exercice 11.17 en choisissant convenablement $\mathcal{B}_1$ et $\mathcal{B}_2$.]

Exercice 11.19 (Exercice liminaire aux martingales) Soient (E, T, p) un espace probabilisé et $(\mathrm{T}_n)_{n\in\mathbb{N}}$ une famille de tribus sur E t.q. $\mathrm{T}_n \subset \mathrm{T}_{n+1}$, pour tout $n \in \mathbb{N}$, et t.q. T est la tribu engendrée par $\bigcup_{n\in\mathbb{N}} \mathrm{T}_n$. Soit $\mathrm{X} \in \mathrm{L}^2_{\mathbb{R}}(\mathrm{E}, \mathrm{T}, p)$ et $\mathrm{E}(\mathrm{X}|\mathrm{T}_n)$ l'espérance conditionnelle de X par rapport à la tribu T_n. Nous allons montrer que $\mathrm{E}(\mathrm{X}|\mathrm{T}_n)$ converge vers X dans $\mathrm{L}^2_{\mathbb{R}}(\mathrm{E}, \mathrm{T}, p)$ lorsque $n \to +\infty$.

1. Montrer qu'il existe $e \in \mathrm{L}^2_{\mathbb{R}}(\mathrm{E}, \mathrm{T}, p)$ et une sous-suite de la suite $(\mathrm{E}(\mathrm{X}|\mathrm{T}_n))_{n\in\mathbb{N}}$ qui converge faiblement vers e dans $\mathrm{L}^2_{\mathbb{R}}(\mathrm{E}, \mathrm{T}, p)$.

2. Montrer que $\int \mathrm{XY}dp = \int e\mathrm{Y}dp$, pour tout v.a.r. Y T_n-mesurable et de carré intégrable.

3. Montrer que $\int XYdp = \int eYdp$, pour tout v.a.r. Y de carré intégrable. En déduire que $e = X$ p.s..
4. Montrer que $\|E(X|T_n)\|_2 \le \|X\|_2$. En déduire que la suite $(E(X, T_n))_{n\in\mathbb{N}}$ converge vers X dans $L^2_{\mathbb{R}}(E, T, p)$.

Exercice 11.20 (Espérance conditionnelle pour une suite décroissante de tribus) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(\mathcal{B}_n)_{n\in\mathbb{N}}$ une suite de tribus incluses dans $\mathcal{A}$. On suppose $\mathcal{B}_{n+1} \subset \mathcal{B}_n$, pour tout $n \in \mathbb{N}$, et on pose $\mathcal{B} = \bigcap_{n\in\mathbb{N}} \mathcal{B}_n$ (de sorte que $\mathcal{B}$ est aussi une tribu incluse dans $\mathcal{A}$).

1. Soit X une v.a.r.. On suppose que pour tout $n \in \mathbb{N}$ il existe Y_n v.a.r. $\mathcal{B}_n$-mesurable t.q. $X = Y_n$ p.s.. Montrer qu'il existe Y v.a.r. $\mathcal{B}$-mesurable t.q. $X = Y$ p.s..

   N.B. Cette première question montre en quel sens on peut écrire $L^p(\Omega, \mathcal{B}, P) = \bigcap_{n\in\mathbb{N}} L^p(\Omega, \mathcal{B}_n, P)$ pour tout $p \in [1, +\infty]$.

   Corrigé – *Pour tout* $n \in \mathbb{N}$, *il existe* $A_n \in \mathcal{A}$ *t.q.* $P(A_n) = 0$ *et* $X = Y_n$ *sur* A_n^c. *On pose* $A = \bigcup_{n\in\mathbb{N}} A_n$. *On a donc* $P(A) = 0$ *(par σ-sous–additivité de* P*) et, pour tout* $n \in \mathbb{N}$, $X = Y_n$ *sur* A^c.

   On définit maintenant $\bar{Y}$ *de* Ω *dans* $\overline{\mathbb{R}}$ *en posant* $\bar{Y}(\omega) = \liminf_{n\to+\infty} Y_n(\omega)$. *Comme* $Y_n = X$ *sur* A^c *(pour tout* $n \in \mathbb{N}$*), on a donc* $\bar{Y} = X$ *sur* A^c *et donc* $\bar{Y} = X$ *p.s..*

   Soit $p \in \mathbb{N}$. *Comme* Y_n *est* $\mathcal{B}_p$*-mesurable pour* $n \ge p$ *(car* $\mathcal{B}_n \subset \mathcal{B}_p$*), la stabilité des fonctions mesurables donne que* $\bar{Y}$ *est* $\mathcal{B}_p$*-mesurable. Soit* C *un borélien de* $\overline{\mathbb{R}}$, *on a donc* $\bar{Y}^{-1}(C) \in \mathcal{B}_p$ *pour* $p \in \mathbb{N}$. *Par la définition de* $\mathcal{B}$, *on en déduit que* $\bar{Y}^{-1}(C) \in \mathcal{B}$. *On a donc* $\bar{Y}$ $\mathcal{B}$*-mesurable de* Ω *dans* $\overline{\mathbb{R}}$ *et* $\bar{Y} = X$ *p.s.. Il nous reste à modifier légèrement* $\bar{Y}$ *pour obtenir une v.a.r.. Pour cela, on pose* $E = \{\omega \in \Omega, \bar{Y}(\omega) = \pm\infty\}$. *On a* $E \in \mathcal{B}$ *(car* $\bar{Y}$ *est* $\mathcal{B}$*-mesurable). On peut définir* Y *par*

   $$Y(\omega) = \bar{Y}(\omega) \text{ si } \omega \notin E,$$
   $$Y(\omega) = 0 \text{ si } \omega \in E.$$

   La fonction Y *est ainsi une v.a.r.* $\mathcal{B}$*-mesurable et* $Y = X$ *p.s. (noter que* $E \subset A$*).*

2. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r.. On suppose que, pour tout $n \in \mathbb{N}$, X_n est $\mathcal{B}_n$-mesurable et de carré intégrable.

   (a) On suppose que $X_n \to X$ dans $L^2(\Omega, \mathcal{A}, P)$ quand $n \to +\infty$. Montrer que X est $\mathcal{B}$-mesurable (au sens qu'il existe Y $\mathcal{B}$-mesurable t.q. $X = Y$ p.s.).

   Corrigé – *On peut supposer, après extraction éventuelle d'une sous-suite (encore notée* $(X_n)_{n\in\mathbb{N}}$*), que* $X_n \to X$ *p.s.. (Après cette extraction, on a toujours* $\bigcap_{n\in\mathbb{N}} \mathcal{B}_n = \mathcal{B}$.*)*

   Comme $X_n \to X$ *p.s., il existe donc* $A \in \mathcal{A}$ *t.q.* $P(A) = 0$ *et* $X_n(\omega) \to X(\omega)$ *(quand* $n \to +\infty$*) pour tout* $\omega \in A^c$. *On procède alors comme à la question précédente en posant* $\bar{Y}(\omega) = \liminf_{n\to+\infty} X_n(\omega)$. *La fonction* $\bar{Y}$ *va de* Ω *dans* $\overline{\mathbb{R}}$ *et est* $\mathcal{B}_p$*-mesurable pour tout* $p \in \mathbb{N}$ *(car* X_n *est* $\mathcal{B}_p$*-mesurable pour* $n \ge p$*). La fonction* $\bar{Y}$

est donc $\mathcal{B}$-mesurable et $\bar{\mathrm{Y}} = \mathrm{X}$ p.s. (car $\mathrm{Y} = \mathrm{X}$ sur A^c). Enfin, on pose $\mathrm{E} = \{\omega \in \Omega,$ $\bar{\mathrm{Y}}(\omega) = \pm\infty\}$. On a $\mathrm{E} \in \mathcal{B}$ (car $\bar{\mathrm{Y}}$ est $\mathcal{B}$-mesurable) et on définit Y par

$$\mathrm{Y}(\omega) = \bar{\mathrm{Y}}(\omega) \text{ si } \omega \notin \mathrm{E},$$
$$\mathrm{Y}(\omega) = 0 \text{ si } \omega \in \mathrm{E}.$$

La fonction Y est ainsi une v.a.r. $\mathcal{B}$-mesurable et $\mathrm{Y} = \mathrm{X}$ p.s..

(b) On suppose que $\mathrm{X}_n \to \mathrm{X}$ faiblement dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ quand $n \to +\infty$. Montrer que X est $\mathcal{B}$-mesurable.

Corrigé – *On peut montrer cette question en utilisant une petite remarque d'analyse fonctionnelle (voir la remarque 6.84). Puisque $\mathrm{X}_n \to \mathrm{X}$ faiblement dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ quand $n \to +\infty$, il existe une suite $(\mathrm{Y}_n)_{n\in\mathbb{N}}$ t.q.*

i. $\mathrm{Y}_n \to \mathrm{X}$ *dans* $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ *quand* $n \to +\infty$,

ii. pour tout $n \in \mathbb{N}$, Y_n est une combinaison convexe de l'ensemble des X_p, $p \geq n$.

Comme X_p est $\mathcal{B}_n$-mesurable pour $p \geq n$, la v.a.r. Y_n est aussi B_n-mesurable. On est ainsi ramené à la question précédente et on obtient qu'il existe Y v.a.r. $\mathcal{B}$-mesurable t.q. $\mathrm{X} = \mathrm{Y}$ p.s..

3. Soit X une v.a.r. de carré intégrable.
Montrer que $\mathrm{E}(\mathrm{X}|\mathcal{B}_n) \to \mathrm{E}(\mathrm{X}|\mathcal{B})$ dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ quand $n \to +\infty$.

Corrigé – *On pose $\mathrm{Z}_n = \mathrm{E}(\mathrm{X}|\mathcal{B}_n)$ (en étant quelque peu pointilleux, on devrait plutôt dire qu'on choisit un élément de l'ensemble $\mathrm{E}(\mathrm{X}|\mathcal{B}_n)$). On pose aussi $\mathrm{Z} = \mathrm{E}(\mathrm{X}|\mathcal{B})$ et on raisonne par l'absurde.*

Si $\mathrm{Z}_n \not\to \mathrm{Z}$ dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ quand $n \to +\infty$, il existe $\varepsilon > 0$ et une sous-suite, encore notée $(\mathrm{Z}_n)_{n\in\mathbb{N}}$, t.q.

$$\|\mathrm{Z}_n - \mathrm{Z}\|_2 \geq \varepsilon \text{ pour tout } n \in \mathbb{N}. \tag{11.15}$$

On remarque maintenant que la suite $(\mathrm{Z}_n)_{n\in\mathbb{N}}$ est bornée dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ (plus précisément, on a $\mathrm{E}(\mathrm{Z}_n^2) \leq \mathrm{E}(\mathrm{X}^2)$ car $\mathrm{E}(\mathrm{Z}_n^2) = \mathrm{E}(\mathrm{Z}_n\mathrm{X}) \leq \sqrt{\mathrm{E}(\mathrm{Z}_n^2)}\sqrt{\mathrm{E}(\mathrm{X}^2)}$). Comme $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ est un espace de Hilbert, on peut donc supposer, toujours après extraction d'une sous-suite, qu'il existe $\bar{\mathrm{Z}}$ t.q.

$$\mathrm{Z}_n \to \bar{\mathrm{Z}} \text{ faiblement dans } \mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P}) \text{ quand } n \to +\infty.$$

On montre maintenant que $\bar{\mathrm{Z}} = \mathrm{Z}$ p.s..

La question précédente montre que $\bar{\mathrm{Z}}$ est $\mathcal{B}$-mesurable. Puis pour tout U v.a.r. $\mathcal{B}$-mesurable et de carré intégrable on a

$$\mathrm{E}(\mathrm{Z}_n\mathrm{U}) = \mathrm{E}(\mathrm{X}\mathrm{U}) \text{ pour tout } n \in \mathbb{N}.$$

Comme $\mathrm{Z}_n \to \bar{\mathrm{Z}}$ faiblement dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$, on peut passer à la limite (quand $n \to +\infty$) dans cette égalité. On obtient

$$\mathrm{E}(\bar{\mathrm{Z}}\mathrm{U}) = \mathrm{E}(\mathrm{X}\mathrm{U}).$$

Ce qui prouve que $\bar{\mathrm{Z}} = \mathrm{E}(\mathrm{X}|\mathcal{B})$ p.s. et donc que $\bar{\mathrm{Z}} = \mathrm{Z}$ p.s.. Pour conclure, il reste à montrer que $\mathrm{Z}_n \to \mathrm{Z}$ dans $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ (en contradiction avec (11.15)). Pour cela, il suffit de montrer que $\mathrm{E}(\mathrm{Z}_n^2) \to \mathrm{E}(\mathrm{Z}^2)$ (car $\mathrm{E}((\mathrm{Z}_n - \mathrm{Z})^2) = \mathrm{E}(\mathrm{Z}_n^2) - 2\mathrm{E}(\mathrm{Z}_n\mathrm{Z}) + \mathrm{E}(\mathrm{Z}^2)$

et $E(Z_n Z) \to E(Z^2)$ *grâce à la convergence faible de* Z_n *vers* Z*). On utilise une nouvelle fois la convergence faible de* Z_n *vers* Z *(et le fait que* $Z_n = E(X|\mathcal{B}_n)$ *et* $Z = E(X|\mathcal{B})$ *pour écrire que*

$$E(Z_n^2) = E(Z_n X) \to E(ZX) = E(Z^2) \text{ quand } n \to +\infty.$$

Finalement, on obtient bien que $Z_n \to Z$ *dans* $L^2(\Omega, \mathcal{A}, P)$*, en contradiction avec (11.15).*

Exercice 11.21 (Projection sur $L^2(\Omega, \tau(X), P)$ versus projection sur ev(X)) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Si $\mathcal{B}$ est une sous–tribu de $\mathcal{A}$, on note encore P la restriction de P à $\mathcal{B}$, de sorte que $(\Omega, \mathcal{B}, P)$ est encore un espace probabilisé. On rappelle que $\mathcal{L}^2(\Omega, \mathcal{B}, P) \subset \mathcal{L}^2(\Omega, \mathcal{A}, P)$ et que l'on peut considérer $L^2(\Omega, \mathcal{B}, P)$ comme un s.e.v. de $L^2(\Omega, \mathcal{A}, P)$. L'espace $L^2(\Omega, \mathcal{B}, P)$ est donc un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$.

Soit $X \in L^2(\Omega, \mathcal{A}, P)$ (en choisissant un représentant de X, X est donc une v.a.). On note ev(X) le s.e.v. engendré par X (noter que ev(X) est un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$).

Si V est un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$, on note P_V l'opérateur de projection orthogonale sur V.

1. On suppose que X est constante et non nulle (c'est-à-dire qu'il existe $a \in \mathbb{R}^*$ t.q. $X = a$ p.s.). Montrer que $P_{ev(X)} = P_{L^2(\Omega, \tau(X), P)}$ (*i.e.* $ev(X) = L^2(\Omega, \tau(X), P)$).
2. On suppose que X n'est pas constante. Montrer que pour toute sous–tribu $\mathcal{B}$ de $\mathcal{A}$, $P_{ev(X)} \neq P_{L^2(\Omega, \mathcal{B}, P)}$ (*i.e.* $ev(X) \neq L^2(\Omega, \mathcal{B}, P)$).

   Remarque : Si Y est une v.a. de carré intégrable, on a $E(Y|X) = E[Y|\tau(X)] = P_{L^2(\Omega, \tau(X), P)} Y$.

Exercice 11.22 (Loi de X et Y et loi de (X, Y)) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r.. On suppose que le couple (X, Y) a pour loi une loi de densité par rapport à la mesure de Lebesgue (sur les boréliens de $\mathbb{R}^2$) et que cette densité est donnée par la fonction g de $\mathbb{R}^2$ dans $\mathbb{R}_+$ définie par

$$g(x, y) = \frac{1 + x^2}{4\sqrt{2\pi}} e^{-\frac{(1+x^2)|y|}{2} - \frac{x^2}{2}} \text{ pour } x, y \in \mathbb{R}.$$

1. Montrer que X a pour loi la loi normale réduite.

   Corrigé – *Soit* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*. En utilisant la loi de* (X, Y)*, on obtient*

   $$E(\varphi(X)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi(x) g(x, y) dy dx.$$

Or,

$$\int_{\mathbb{R}} g(x,y)dy = \frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_{\mathbb{R}} e^{-\frac{(1+x^2)|y|}{2}}dy = \frac{1+x^2}{2\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_0^{+\infty} e^{-\frac{(1+x^2)y}{2}}dy$$
$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

On a donc $\mathrm{E}(\varphi(\mathrm{X})) = \int_{\mathbb{R}} \varphi(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$*. Ce qui prouve que* X *a pour loi la loi normale réduite.*

2. Montrer que Y a pour loi une loi de densité par rapport à la mesure de Lebesgue (sur les boréliens de $\mathbb{R}$) et donner une expression de cette densité.

Corrigé – *Soit* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$*. En utilisant la loi de* (X,Y)*, on obtient*

$$\mathrm{E}(\varphi(\mathrm{Y})) = \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(y)g(x,y)dxdy.$$

Pour $y\in\mathbb{R}$*, on pose* $h(y) = \int_{\mathbb{R}} g(x,y)dx$*, c'est-à-dire*

$$h(y) = \int_{\mathbb{R}} \frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{x^2}{2}}e^{-\frac{(1+x^2)|y|}{2}}dx.$$

On a alors $\mathrm{E}(\varphi(\mathrm{Y})) = \int_{\mathbb{R}} \varphi(y)h(y)dy$*. Ce qui prouve que* Y *a pour loi une loi de densité par rapport à la mesure de Lebesgue et cette densité est donnée par la fonction* h*.*

3. Soit $y \le 0$. Montrer que

$$\mathrm{P}(\{\mathrm{Y} < y\}) = \frac{1}{2\sqrt{1-y}}e^{\frac{y}{2}}.$$

Soit $y > 0$, montrer que

$$\mathrm{P}(\{\mathrm{Y} < y\}) = 1 - \frac{1}{2\sqrt{1+y}}e^{-\frac{y}{2}}.$$

Corrigé – *Soit* $y\in\mathbb{R}$*. En utilisant la densité* h *de la loi de* Y*, on a*

$$\mathrm{P}(\{\mathrm{Y} < y\}) = \int_{-\infty}^{y} h(z)dz = \int_{-\infty}^{y}\Big(\int_{\mathbb{R}} \frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{x^2}{2}}e^{-\frac{(1+x^2)|z|}{2}}dx\Big)dz.$$

Avec le théorème de Fubini-Tonelli, on en déduit

$$\mathrm{P}(\{\mathrm{Y} < y\}) = \int_{\mathbb{R}}\Big(\int_{-\infty}^{y} e^{-\frac{(1+x^2)|z|}{2}}dz\Big)\frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

Si $y \le 0$*, on a*

$$\int_{-\infty}^{y} e^{-\frac{(1+x^2)|z|}{2}}dz = \int_{-\infty}^{y} e^{\frac{(1+x^2)z}{2}}dz = \frac{2}{(1+x^2)}e^{\frac{(1+x^2)y}{2}}.$$

Ce qui donne

$$\mathrm{P}(\{\mathrm{Y}<y\}) = \frac{1}{2\sqrt{2\pi}}\int_{\mathbb{R}} e^{\frac{(1+x^2)y}{2}} e^{-\frac{x^2}{2}}dx = e^{\frac{y}{2}}\frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{\frac{x^2(y-1)}{2}}dx.$$

Dans cette dernière intégrale, on effectue le changement de variable $x\sqrt{1-y}=u$ *et on obtient*

$$\mathrm{P}(\{\mathrm{Y}<y\}) = \frac{e^{\frac{y}{2}}}{\sqrt{1-y}}\frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{u^2}{2}}du = \frac{e^{\frac{y}{2}}}{2\sqrt{1-y}}.$$

On déduit en particulier de cette formule que $\mathrm{P}(\{\mathrm{Y}<0\}) = 1/2$ *et comme* $h(-z)=h(z)$*, on trouve bien que*

$$\mathrm{P}(\{\mathrm{Y}\in\mathbb{R}\}) = 2\int_{-\infty}^0 h(z)dz = 2\mathrm{P}(\{\mathrm{Y}<0\}) = 1.$$

On calcule maintenant $\mathrm{P}(\{\mathrm{Y}<y\})$ *si* $y>0$*. On a*

$$\mathrm{P}(\{\mathrm{Y}<y\}) = \mathrm{P}(\{\mathrm{Y}\in\mathbb{R}\}) - \mathrm{P}(\{\mathrm{Y}\geq y\}) = 1 - \mathrm{P}(\{\mathrm{Y}\geq y\}).$$

et, avec le théorème de Fubini-Tonelli,

$$\mathrm{P}(\{\mathrm{Y}\geq y\}) = \int_y^{+\infty} h(z)dz = \int_{\mathbb{R}}\Big(\int_y^{+\infty} e^{-\frac{(1+x^2)|z|}{2}}dz\Big)\frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

Comme $y>0$*, on a*

$$\int_y^{+\infty} e^{-\frac{(1+x^2)|z|}{2}}dz = \int_y^{+\infty} e^{-\frac{(1+x^2)z}{2}}dz = \frac{2}{(1+x^2)}e^{\frac{-(1+x^2)y}{2}}.$$

Ce qui donne

$$\mathrm{P}(\{\mathrm{Y}\geq y\}) = \frac{1}{2\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{(1+x^2)y}{2}} e^{-\frac{x^2}{2}}dx = e^{-\frac{y}{2}}\frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{x^2(y+1)}{2}}dx.$$

Dans cette dernière intégrale, on effectue le changement de variable $x\sqrt{1+y}=u$ *et on obtient*

$$\mathrm{P}(\{\mathrm{Y}\geq y\}) = \frac{e^{-\frac{y}{2}}}{\sqrt{1+y}}\frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{u^2}{2}}du = \frac{e^{-\frac{y}{2}}}{2\sqrt{1+y}}.$$

On a bien, finalement,

$$\mathrm{P}(\{\mathrm{Y}<y\}) = 1 - \frac{e^{-\frac{y}{2}}}{2\sqrt{1+y}}.$$

4. (Espérances conditionnelles) Montrer que $\mathrm{E}(\mathrm{Y}|\mathrm{X}) = 0$ p.s.. Donner $\mathrm{E}(\mathrm{Y}^2|\mathrm{X})$ en fonction de X. En déduire la variance de Y.

Corrigé – *On montre tout d'abord que* Y *est une v.a.r. de carré intégrable. En utilisant le téorème de Fubini-Tonelli et la question précédente, on a*

$$\begin{aligned}\mathrm{E}(\mathrm{Y}^2) &= \int_\Omega \mathrm{Y}^2 d\mathrm{P} = \int_\Omega\Big(\int_0^{+\infty} 1_{[0,\mathrm{Y}^2(\omega)[}(t)dt\Big)d\mathrm{P}(\omega)\\ &= \int_0^{+\infty}\Big(\int_\Omega 1_{]t,+\infty[}(\mathrm{Y}^2(\omega))d\mathrm{P}(\omega)\Big)dt = \int_0^{+\infty}\mathrm{P}(\{\mathrm{Y}^2>t\})dt\\ &= \int_0^{+\infty} 2u\mathrm{P}(\{|\mathrm{Y}|>u\})du = \int_0^{+\infty} 2u\frac{e^{-\frac{u}{2}}}{\sqrt{1+u}}du < +\infty.\end{aligned}$$

Comme Y *est de carré intégrable les espérances conditionnelles* $\mathrm{E(Y|X)}$ *et* $\mathrm{E(Y^2|X)}$ *sont bien définies.*

Pour calculer $\mathrm{E(Y|X)}$ *et* $\mathrm{E(Y^2|X)}$, *on utilise la proposition 11.11. Soit* φ *une fonction borélienne bornée de* $\mathbb{R}$ *dans* $\mathbb{R}$, *on a*

$$\mathrm{E(Y\varphi(X))} = \int_{\mathbb{R}}\int_{\mathbb{R}} y\varphi(x)g(x,y)dydx = \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} yg(x,y)dy\Big)\varphi(x)dx.$$

Comme $g(x,-y)=g(x,y)$, *on a* $\int_{\mathbb{R}} yg(x,y)dy = 0$ *pour tout* x *et donc* $\mathrm{E(Y\varphi(X))}=0$.

On en déduit que $\mathrm{E(Y|X)}=0$ *p.s.*.

On procède de manière analogue pour Y^2. *On a*

$$\mathrm{E(Y^2\varphi(X))} = \int_{\mathbb{R}}\int_{\mathbb{R}} y^2\varphi(x)g(x,y)dydx = \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} y^2g(x,y)dy\Big)\varphi(x)dx.$$

Pour $x\in\mathbb{R}$, *on a*

$$\int_{\mathbb{R}} y^2g(x,y)dy = \int_{\mathbb{R}} y^2\frac{1+x^2}{4\sqrt{2\pi}}e^{-\frac{(1+x^2)|y|}{2}-\frac{x^2}{2}}dy = \frac{1+x^2}{2\sqrt{2\pi}}e^{-\frac{x^2}{2}}\int_0^{+\infty} y^2e^{-\frac{(1+x^2)y}{2}}dy.$$

On calcule cette dernière intégrale en intégrant deux fois par parties. On obtient

$$\int_{\mathbb{R}} y^2g(x,y)dy = \frac{8}{\sqrt{2\pi}(1+x^2)^2}e^{-\frac{x^2}{2}}.$$

On a donc

$$\mathrm{E(Y^2\varphi(X))} = \int_{\mathbb{R}}\frac{8}{\sqrt{2\pi}(1+x^2)^2}e^{-\frac{x^2}{2}}\varphi(x)dx.$$

Ce qui donne, en posant $\psi(x)=\frac{8}{(1+x^2)^2}$,

$$\mathrm{E(Y^2\varphi(X))} = \int_{\mathbb{R}}\psi(x)\varphi(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \mathrm{E(\psi(X)\varphi(X))}.$$

On en déduit que

$$\mathrm{E(Y^2|X)} = \psi(\mathrm{X}) = \frac{8}{(1+\mathrm{X}^2)^2}\ \textit{p.s.}\ .$$

On calcule maintenant la variance de Y, *notée* $Var(\mathrm{Y})$. *Comme* $\mathrm{E(Y)}=\mathrm{E(E(Y|X))}=0$ *et* $\mathrm{E(Y^2)}=\mathrm{E(E(Y^2|X))}$, *on a* $Var(\mathrm{Y})=\mathrm{E(E(Y^2|X))}=\mathrm{E(\psi(X))}$ *et donc*

$$Var(\mathrm{Y}) = 8\mathrm{E}\Big(\frac{1}{(1+\mathrm{X}^2)^2}\Big).$$

On remarque maintenant que

$$\mathrm{E}\Big(\frac{1}{(1+\mathrm{X}^2)^2}\Big) = \mathrm{E}\Big(\frac{1}{1+\mathrm{X}^2}\Big) - \mathrm{E}\Big(\frac{\mathrm{X}^2}{(1+\mathrm{X}^2)^2}\Big). \tag{11.16}$$

. Mais, on a

$$\sqrt{2\pi}\mathrm{E}\Big(\frac{\mathrm{X}^2}{(1+\mathrm{X}^2)^2}\Big) = 2\int_0^{+\infty}\frac{x}{(1+x^2)^2}xe^{-\frac{x^2}{2}}dx.$$

En intégrant par parties, on obtient

$$\begin{aligned}\sqrt{2\pi}\mathrm{E}\Big(\frac{\mathrm{X}^2}{(1+\mathrm{X}^2)^2}\Big) &= \int_0^{+\infty}\frac{1}{1+x^2}e^{-\frac{x^2}{2}}(1-x^2)dx\\ &= 2\int_0^{+\infty}\frac{1}{1+x^2}e^{-\frac{x^2}{2}}dx - \int_0^{+\infty}e^{-\frac{x^2}{2}}dx.\end{aligned}$$

Ce qui donne

$$\mathrm{E}(\frac{\mathrm{X}^2}{(1+\mathrm{X}^2)^2}) = \mathrm{E}(\frac{1}{(1+\mathrm{X}^2)}) - \frac{1}{2}.$$

En revenant à (11.16), on a donc $\mathrm{E}(\frac{1}{(1+\mathrm{X}^2)^2}) = \frac{1}{2}$*, ce qui donne* $Var(\mathrm{Y}) = 4$.

11.3.2 Martingales

Exercice 11.23 (Quelques propriétés des martingales) Soit $(\Omega, \mathcal{A}, \mathrm{P})$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$ (c'est-à-dire d'une suite croissante de sous tribus de $\mathcal{A}$) et $(\mathrm{X}_n)_{n\in\mathbb{N}}$ une suite de de v.a.r. (c'est-à-dire un processus réel). On suppose que X_n est intégrable pour tout $n \in \mathbb{N}$.

1. On suppose que $(\mathrm{X}_n)_{n\in\mathbb{N}}$ est une sous–martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). Montrer que la suite $(\mathrm{E}(\mathrm{X}_n))_{n\in\mathbb{N}}$ est croissante.

Corrigé – *Soit* $n \in \mathbb{N}$*. Comme* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *est une sous–martingale, on a*

$$\mathrm{E}(\mathrm{X}_{n+1}|\mathcal{B}_n) \geq \mathrm{X}_n \textit{ p.s. et donc } \mathrm{E}(\mathrm{E}(\mathrm{X}_{n+1}|\mathcal{B}_n)) \geq \mathrm{E}(\mathrm{X}_n).$$

Or (comme les fonctions constantes sont $\mathcal{B}_n$*-mesurables bornées),*

$$\mathrm{E}(\mathrm{E}(\mathrm{X}_{n+1}|\mathcal{B}_n)) = \mathrm{E}(\mathrm{X}_{n+1}).$$

On a donc $\mathrm{E}(\mathrm{X}_{n+1}) \geq \mathrm{E}(\mathrm{X}_n)$.

2. On suppose que $(\mathrm{X}_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$). Montrer que $\mathrm{E}(\mathrm{X}_{n+m}|\mathcal{B}_n) = \mathrm{X}_n$ p.s. pour tout $m \geq 0$.

Corrigé – *Pour* $m = 0$*, le fait que* $\mathrm{E}(\mathrm{X}_n|\mathcal{B}_n) = \mathrm{X}_n$ *p.s., pour tout* $n \in \mathbb{N}$*, découle du fait que* X_n *est* $\mathcal{B}_n$*-mesurable. On montre maintenant la propriété demandée par récurrence sur* $m \in \mathbb{N}^*$.

Pour $m = 1$ *le fait que* $\mathrm{E}(\mathrm{X}_{n+1}|\mathcal{B}_n) = \mathrm{X}_n$ *p.s., pour tout* $n \in \mathbb{N}$*, est donné dans la définition de martingale.*

Soit $m \in \mathbb{N}^*$*. On suppose que* $\mathrm{E}(\mathrm{X}_{n+m}|\mathcal{B}_n) = \mathrm{X}_n$ *p.s., pour tout* $n \in \mathbb{N}$*. On veut montrer que* $\mathrm{E}(\mathrm{X}_{n+m+1}|\mathcal{B}_n) = \mathrm{X}_n$ *p.s., pour tout* $n \in \mathbb{N}$*. Soit* $n \in \mathbb{N}$*. Comme* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *est une martingale, on a* $\mathrm{E}(\mathrm{X}_{n+m+1}|\mathcal{B}_{n+m}) = \mathrm{X}_{n+m}$ *p.s. et donc :*

$$\mathrm{E}(\mathrm{X}_{n+m+1}\mathrm{U}) = \mathrm{E}(\mathrm{X}_{n+m}\mathrm{U}) \textit{ pour tout } \mathrm{U}\ \mathcal{B}_{n+m}\textit{-mesurable bornée.}$$

Comme $\mathcal{B}_n \subset \mathcal{B}_{n+m}$*, on a donc aussi*

$$\mathrm{E}(\mathrm{X}_{n+m+1}\mathrm{U}) = \mathrm{E}(\mathrm{X}_{n+m}\mathrm{U}) \textit{ pour tout } \mathrm{U}\ \mathcal{B}_n\textit{-mesurable bornée.}$$

L'hypothèse de récurrence donne $\mathrm{E}(\mathrm{X}_{n+m}|\mathcal{B}_n) = \mathrm{X}_n$ *p.s.. On a donc :*

$$\mathrm{E}(\mathrm{X}_{n+m}\mathrm{U}) = \mathrm{E}(\mathrm{X}_n\mathrm{U}) \textit{ pour tout } \mathrm{U}\ \mathcal{B}_n\textit{-mesurable bornée.}$$

On a en déduit :

$$\mathrm{E}(\mathrm{X}_{n+m+1}\mathrm{U}) = \mathrm{E}(\mathrm{X}_n\mathrm{U}) \textit{ pour tout } \mathrm{U}\ \mathcal{B}_n\textit{-mesurable bornée.}$$

Ce qui montre que $\mathrm{E}(\mathrm{X}_{n+m+1}|\mathcal{B}_n) = \mathrm{X}_n$ *p.s. et termine la récurrence.*

3. Soit φ une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$) et que $\varphi(X_n)$ est intégrable pour tout $n \in \mathbb{N}$ (on rappelle que $\varphi(X_n)$ est une notation pour désigner $\varphi \circ X_n$). Montrer que $(\varphi(X_n))_{n\in\mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$).

Corrigé – *On remarque tout d'abord que $\varphi(X_n)$ est bien $\mathcal{B}_n$-mesurable (car X_n est $\mathcal{B}_n$-mesurable et φ est borélienne), pour tout $n \in \mathbb{N}$. Pour montrer que $(\varphi(X_n))_{n\in\mathbb{N}}$ est une sous-martingale, il suffit alors d'utiliser le proposition 11.7 sur l'inégalité de jensen. Soit $n \in \mathbb{N}$. La proposition 11.7 donne*

$$E(\varphi(X_{n+1})|\mathcal{B}_n) \geq \varphi(E(X_{n+1}|\mathcal{B}_n)) \text{ p.s..}$$

Comme $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s., on en déduit $E(\varphi(X_{n+1})|\mathcal{B}_n) \geq \varphi(X_n)$ p.s., ce qui montre bien que $(\varphi(X_n))_{n\in\mathbb{N}}$ est une sous-martingale.

Exercice 11.24 (Martingale construite avec les espérances d'une v.a.)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$. Soit X une v.a.r. intégrable. Montrer que la suite $(X_n)_{n\in\mathbb{N}}$ définie par $X_n = E(X|\mathcal{B}_n)$ (pour tout $n \in \mathbb{N}$) est une martingale par rapport à la filtration $(\mathcal{B}_n)$.

Montrer que $(X_n)_{n\in\mathbb{N}}$ est aussi une martingale pour la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ définie par $\mathcal{F}_n = \tau(X_1, \ldots, X_n)$ (pour tout $n \in \mathbb{N}$).

Corrigé – *Soit $n \in \mathbb{N}$. On remarque tout d'abord que X_n est bien $\mathcal{B}_n$-mesurable et intégrable. Il reste à montrer que $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s..*

Soit U uner v.a.r. $\mathcal{B}_n$-mesurable bornée. Comme $X_n = E(X|\mathcal{B}_n)$ p.s., on a $E(X_n U) = E(XU)$. Mail, comme $\mathcal{B}_n \subset \mathcal{B}_{n+1}$, la v.a.r. U est aussi $\mathcal{B}_{n+1}$-mesurable et le fait que $X_{n+1} = E(X|\mathcal{B}_{n+1})$ p.s. donne alors $E(X_{n+1}U) = E(XU)$. On a donc

$$E(X_n U) = E(X_{n+1} U).$$

On en déduit bien que $X_n = E(X_{n+1}|\mathcal{B}_n)$ p.s.. La suite $(X_n)_{n\in\mathbb{N}}$ est donc une martingale par rapport à la filtration $(\mathcal{B}_n)$.

Pour montrer que la suite $(X_n)_{n\in\mathbb{N}}$ est aussi une martingale par rapport à la filtration $(\mathcal{F}_n)$, il suffit de montrer que $X_n = E(X|\mathcal{F}_n)$ pour tout $n \in \mathbb{N}$ (le raisonnement précédent permet alors de conclure en remplaçant $\mathcal{B}_n$ par $\mathcal{F}_n$).

Soit donc $n \in \mathbb{N}$. Il est clair que X_n est $\mathcal{F}_n$-mesurable. Comme X_p est $\mathcal{B}_n$-mesurable pour tout $p \leq n$, il est clair aussi que $\mathcal{F}_n \subset \mathcal{B}_n$. Soit U une v.a.r. $\mathcal{F}_n$ mesurable bornée. La v.a.r. U est donc aussi $\mathcal{B}_n$-mesurable. Puisque $X_n = E(X|\mathcal{B}_n)$, on a

$$E(X_n U) = E(XU).$$

On en déduit bien que $X_n = E(X|\mathcal{F}_n)$ p.s., ce qui termine cette démonstration.

Exercice 11.25 (Martingale d'un jeu équilibré) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, X_0 une v.a.r. intégrable et $(J_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. intégrable et de moyenne nulle. On suppose que la suite formée de X_0 et $(J_n)_{n\in\mathbb{N}^*}$ est une suite de v.a.r. indépendantes. On pose alors, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$ et $\mathcal{B}_n$ la tribu engendrée par $X_0, \ldots, X_n$. Montrer que la suite $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$).

Corrigé – *Par récurrence sur n, on remarque tout d'abord que* X_n *est bien intégrable pour tout* $n \in \mathbb{N}$. *Pour tout* $n \in \mathbb{N}$, *la v.a.r.* X_n *est donc* $\mathcal{B}_n$*-mesurable et intégrable. Il reste à montrer que* $E(X_{n+1}|\mathcal{B}_n) = X_n$ *p.s..*

Soit U *une v.a.r.* $\mathcal{B}_n$*-mesurable bornée. On a* $E(X_{n+1}U) = E(X_nU) + E(J_{n+1}U)$. *En utilisant la proposition 3.30, on remarque que les v.a.r.* $X_0, \ldots, X_n$, J_{n+1} *sont indépendantes. Puis, avec la proposition 2.59, on remarque que la tribu engendrée par* J_{n+1} *est indépendante de la tribu engendrée par* $X_0, \ldots, X_n$ *(qui est* $\mathcal{B}_n$*). On en déduit que* J_{n+1} *et* U *sont des v.a.r. indépendantes. Ceci donne* $E(J_{n+1}U) = E(J_{n+1})E(U) = 0$ *(car* J_{n+1} *est de moyenne nulle). Finalement, on a donc* $E(X_{n+1}U) = E(X_nU)$, *ce qui donne bien* $E(X_{n+1}|\mathcal{B}_n) = X_n$ *p.s..*

Exercice 11.26 (Séries de Fourier et martingales) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $\mathcal{B}$ une sous tribu de $\mathcal{A}$. Montrer que

$$X \in L^2(\Omega, \mathcal{B}, P) \Rightarrow X^+ \in L^2(\Omega, \mathcal{B}, P).$$

$(L^2(\Omega, \mathcal{B}, P) = L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P).)$

On prend maintenant $\Omega =]0,1[$, $\mathcal{A} = \mathcal{B}(]0,1[)$ et $P = \lambda$ (la mesure de Lebesgue sur $\mathcal{B}(]0,1[)$). On pose $H = L^2_{\mathbb{C}}(\Omega, \mathcal{A}, P)$. Pour $p \in \mathbb{Z}$, on définit $e_p \in H$ par $e_p(x) = \exp(2i\pi px)$ pour $x \in]0,1[$. On rappelle que $\{e_p, p \in \mathbb{Z}\}$ est une base hilbertienne de H. pour $n \in \mathbb{N}$, on pose $V_n = ev\{e_p, -n \leq p \leq n\}$ (c'est un s.e.v. fermé de H).

Soit $X \in H$. On sait que $X_n \to X$ dans H, quand $n \to +\infty$, avec $X_n = P_{V_n}X$ où P_{V_n} désigne l'opérateur de projection orthogonale sur V_n (X_n est donc une somme partielle de la série de Fourier de X).

1. Montrer que $P_{V_n}(X_{n+1}) = X_n$ pour tout $n \in \mathbb{N}$.
2. Soit $n \in \mathbb{N}^*$, monter qu'il n'existe pas de sous–tribu $\mathcal{B}$ de $\mathcal{A}$ t.q. $V_n = L^2_{\mathbb{C}}(\Omega, \mathcal{B}, P)$. [On pourra, par exemple, commencer par remarquer que V_n est formé de fonctions analytiques.]
3. Soit $n \in \mathbb{N}^*$. On pose $\mathcal{B}_n = \tau(e_p, -n \leq p \leq n)$. A-t-on $V_n \subset L^2_{\mathbb{C}}(\Omega, \mathcal{B}_n, P)$ ou $L^2_{\mathbb{C}}(\Omega, \mathcal{B}_n, P) \subset V_n$?

Exercice 11.27 (Quelques questions sur les temps d'arrêt) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$.

1. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r., adaptée à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$. Soit $E \in \mathcal{B}(\mathbb{R})$ et τ défini de Ω dans $\mathbb{R}$ par
$$\tau(\omega) = 1 \text{ si } X_1(\omega) \in E \text{ et } \tau = 5 \text{ si } X_1(\omega) \notin E.$$
Montrer que τ est un temps d'arrêt.

   Corrigé – *En posant* $C = \{w \in \Omega$*, t.q.* $X_1(\omega) \in E\}$*, on a* $\tau = 1_C + 5\,1_{C^c}$*. Comme* X_1 *est une v.a.r., on a* $C \in \mathcal{A}$*. La fonction* τ *est donc bien mesurable (de* Ω *muni de* $\mathcal{A}$ *dans* $\mathbb{N}$ *muni de* $\mathcal{P}(\mathbb{N})$*).*

   Soit $n \in \mathbb{N}$*, on va montrer que* $\{\tau = n\} \in \mathcal{B}_n$*.*

   Si $n \notin \{1,5\}$*, on a* $\{\tau = n\} = \emptyset \in \mathcal{B}_n$*.*
   Si $n = 1$*, on a* $\{\tau = 1\} = C = X_1^{-1}(E) \in \mathcal{B}_1$ *car* X_1 *est* $\mathcal{B}_1$*-mesurable.*
   Si $n = 5$*, on a* $\{\tau = 5\} = C^c \in \mathcal{B}_1 \subset \mathcal{B}_5$ *car* X_1 *est* $\mathcal{B}_1$*-mesurable et que* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *est une filtration.*

   On en déduit bien que τ *est un temps d'arrêt.*

2. Soit ν et τ deux temps d'arrêt par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$. Montrer que $\nu + \tau$ est encore un temps d'arrêt (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$).

   Corrigé – *La mesurabilité de* $\nu + \tau$ *(de* Ω *muni de* $\mathcal{A}$ *dans* $\mathbb{N} \cup \{+\infty\}$ *muni de* $\mathcal{P}(\mathbb{N} \cup \{+\infty\})$*) découle de la mesurabilité de* ν *et* τ*.*

   Soit $n \in \mathbb{N}$*. On a* $\{\nu + \tau = n\} = \cup_{p=0}^{n}(\{\nu = p\} \cap \{\tau = n-p\})$*. Pour tout* $0 \le p \le n$*, on a* $\{\nu = p\} \in \mathcal{B}_p \subset \mathcal{B}_n$ *et* $\{\tau = n-p\} \in \mathcal{B}_{n-p} \subset \mathcal{B}_n$ *et donc* $\{\nu = p\} \cap \{\tau = n-p\} \subset \mathcal{B}_n$*. On en déduit que* $\{\nu + \tau = n\} \in \mathcal{B}_n$*, ce qui montre bien que* $\nu + \tau$ *est un temps d'arrêt.*

3. Soit ν et τ deux temps d'arrêt, par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$, et t.q. $\nu \le \tau$ p.s.. Soit $\mathcal{B}_\nu$ et $\mathcal{B}_\tau$ les deux tribus associées. Montrer que $\mathcal{B}_\nu \subset \mathcal{B}_\tau$. [Si T est un temps d'arrêt, on rappelle que $\mathcal{B}_T = \{A \in \mathcal{B}_\infty$ t.q., pour tout $n \in \mathbb{N}$, $A \cap \{T = n\} \in \mathcal{B}_n\}$.]

   Corrigé – *Soit* $A \in B_\nu$*. Soit* $n \in \mathbb{N}$*. Comme* $\nu \le \tau$*, on a* $\{\tau = n\} = \cup_{p=0}^{n}\{\nu = p\}$ *et donc*
$$A \cap \{\tau = n\} = \cup_{p=0}^{n}(A \cap \{\nu = p\}).$$
   Comme $A \in B_\nu$*, on a* $A \cap \{\nu = p\} \in \mathcal{B}_p \subset \mathcal{B}_n$ *pour tout* $p \le n$*. On a donc* $A \cap \{\tau = n\} \in \mathcal{B}_n$*, ce qui prouve que* $A \in B_\tau$*.*

Exercice 11.28 (Martingale arrêtée à un temps d'arrêt) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$ et $(X_n)_{n\in\mathbb{N}}$, $(\phi_n)_{n\in\mathbb{N}^*}$ deux suites de v.a.r.. On suppose que, pour tout $n \in \mathbb{N}$, X_n est intégrable et que, pour tout $n \in \mathbb{N}^*$, ϕ_n est bornée. Pour $n \in \mathbb{N}^*$, on pose $(\Delta X)_n = X_n - X_{n-1}$ et $(\phi \cdot X)_n = \sum_{k=1}^{n} \phi_k (\Delta X)_k$ (ceci est une intégrale stochastique discrète).

1. On suppose que $(X_n)_{n\in\mathbb{N}}$ est une martingale et que $(\phi_n)_{n\in\mathbb{N}^*}$ est prévisible (c'est-à-dire que ϕ_{n+1} est $\mathcal{B}_n$-mesurable pour tout $n \in \mathbb{N}$). Montrer que $((\phi \cdot X)_n)_{n\in\mathbb{N}^*}$ est une martingale.

Corrigé – *Soit $n \in \mathbb{N}^*$. Pour tout $1 \le k \le n$, la v.a.r. $(\Delta X)_k$ est $\mathcal{B}_k$-mesurable et donc $\mathcal{B}_n$-mesurable. La v.a.r. ϕ_k est $\mathcal{B}_{k-1}$-mesurable et donc aussi $\mathcal{B}_n$-mesurable. La stabilité des fonctions mesurables montre alors que $(\phi \cdot X)_n$ est $\mathcal{B}_n$-mesurable. De plus, comme les ϕ_k sont bornées et les X_k sont intégrables, la v.a.r. $(\phi \cdot X)_n$ est aussi intégrable. (elle donc $\mathcal{B}_n$-mesurable et intégrable).*

Soit $n \in \mathbb{N}^*$. On montre maintenant que $E((\phi \cdot X)_{n+1}|\mathcal{B}_n) = (\phi \cdot X)_n$ p.s.. Pour cela, on utilise la linéarité de l'éspérance conditionnelle, on obtient

$$E((\phi \cdot X)_{n+1}|\mathcal{B}_n) = \sum_{k=1}^{n+1} E(\phi_k (\Delta X)_k|\mathcal{B}_n).$$

Pour $k \le n$, la v.a.r. $\phi_k(\Delta X)_k$ est $\mathcal{B}_n$-mesurable et donc

$$E(\phi_k(\Delta X)_k|\mathcal{B}_n) = \phi_k(\Delta X)_k \text{ p.s..}$$

Pour $k = n+1$, ϕ_{n+1} est $\mathcal{B}_n$-mesurable et $E(X_{n+1}|\mathcal{B}_n) = X_n$, on a donc (avec l'exercice 11.4)

$$E(\phi_{n+1}(\Delta X)_{n+1}|\mathcal{B}_n) = \phi_{n+1} E(X_{n+1} - X_n|\mathcal{B}_n) = \phi_{n+1}(X_n - X_n) = 0 \text{ p.s..}$$

On en déduit que

$$E((\phi \cdot X)_{n+1}|\mathcal{B}_n) = \sum_{k=1}^{n} \phi_k(\Delta X)_k \text{ p.s.} = (\phi \cdot X)_n \text{ p.s..}$$

Ce qui prouve bien que $((\phi \cdot X)_n)_{n \in \mathbb{N}^*}$ est une martingale.

2. (Exemple) Soit ν un temps d'arrêt. On prend ici $\phi_n = 1_{\{\nu \ge n\}}$ pour tout $n \in \mathbb{N}^*$. Montrer que $(\phi_n)_{n \in \mathbb{N}^*}$ est prévisible et $(\phi \cdot X)_n = X_{\nu \wedge n} - X_0$ pour tout $n \in \mathbb{N}^*$. (On rappelle que $\nu \wedge n(\omega) = \min\{\nu(\omega), n\}$ et donc que $X_{\nu \wedge n}(\omega) = X_{\min\{\nu(\omega), n\}}(\omega)$, pour tout $\omega \in \Omega$.)

   Remarque : La question précédente permet de montrer qu'une martingale arrêtée à un temps d'arrêt est encore une martingale et donne donc une démonstration du théorème 11.18.

   Corrigé – *Soit $n \in \mathbb{N}^*$. On a $1 - \phi_n = 1_{\{\nu < n\}} = \sum_{k=0}^{n-1} 1_{\{\nu = k\}}$. Pour $k \le n-1$, on a $\{\nu = k\} \in \mathcal{B}_k \subset \mathcal{B}_{n-1}$, la v.a.r. $1_{\{\nu = k\}}$ est donc $\mathcal{B}_{n-1}$-mesurable. On en déduit que $1 - \phi_n$ (et donc aussi ϕ_n) est $\mathcal{B}_{n-1}$-mesurable. Ceci prouve que $(\phi_n)_{n \in \mathbb{N}^*}$ est prévisible.*

   Pour $1 \le k \le n$, on a $\phi_k(\Delta X)_k = 1_{\nu \ge k}(X_k - X_{k-1})$. On a donc

$$(\phi \cdot X)_n = \sum_{k=1}^{n} 1_{k \le \nu} (\Delta X)_k = \sum_{k=1}^{\min\{\nu, n\}} (\Delta X)_k = X_{\nu \wedge n} - X_0.$$

Exercice 11.29 (Caractérisation des martingales régulières) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ et $(X_n)_{n \in \mathbb{N}}$ une martingale bornée dans L^1 (c'est-à-dire t.q. $\sup\{E(|X_n|), n \in \mathbb{N}\} < \infty$).

1. Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ est équi-intégrable si et seulement si $(X_n)_{n \in \mathbb{N}}$ est une martingale régulière (c'est-à-dire que la suite $(X_n)_{n \in \mathbb{N}}$ converge dans $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$). [On pourra utiliser le théorème de Vitali, théorème 4.51 et le théorème 11.19.]

Corrigé – *On remarque tout d'abord que, d'après le théorème 11.19, il existe une v.a.r.* X *t.q.* $X_n \to X$ *p.s.. Le théorème de Vitali (théorème 4.51) montre alors que* $X_n \to X$ *dans* $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$ *si et seulement si la suite* $(X_n)_{n\in\mathbb{N}}$ *est équi-intégrable. Ceci donne bien le résultat demandé.*

2. On suppose qu'il existe $q > 1$ t.q. $\sup\{E(|X_n|^q), n \in \mathbb{N}\} < \infty$. Montrer que $(X_n)_{n\in\mathbb{N}}$ est une martingale régulière.

Corrigé – *Comme cela a été dit à la question précédente il existe une v.a.r.* X *t.q.* $X_n \to X$ *p.s.. La suite* $(X_n)_{n\in\mathbb{N}}$ *étant bornée dans* $L^q_{\mathbb{R}}(\Omega, \mathcal{A}, P)$ *avec* $q > 1$, *le théorème 6.10 donne la convergence de* $(X_n)_{n\in\mathbb{N}}$ *vers* X *dans* $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$. *La suite* $(X_n)_{n\in\mathbb{N}}$ *est donc une martingale régulière.*

Exercice 11.30 (Une condition pour qu'un temps d'arrêt soit intégrable) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$ et T un temps d'arrêt par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$.

1. On suppose qu'il existe $\varepsilon_0 \in]0,1[$ et $n_0 \in \mathbb{N}^*$ t.q.

$$P(\{T > kn_0\}) \leq (1-\varepsilon_0)^{k-1} \text{ pour tout } k \in \mathbb{N}^*. \tag{11.17}$$

Montrer que $E(T) < \infty$.

Corrigé – *Comme* T *est mesurable et prend ses valeurs dans* $\mathbb{N} \cup \{+\infty\}$, *son espérance est bien définie. grâce au théorème de convergence monotone, on peut écrire*

$$E(T) = \int_{\{T \leq n_0\}} T dP + \sum_{k=1}^{+\infty} \int_{\{kn_0 < T \leq (k+1)n_0\}} T dP.$$

On en déduit, en posant $u_k = (k+1)(1-\varepsilon_0)^{k-1}$, *que*

$$E(T) \leq n_0 + \sum_{k=1}^{+\infty} (k+1) n_0 P(\{T > kn_0\}) \leq n_0 (1 + \sum_{k=1}^{+\infty} u_k).$$

La série de terme général u_k *est convergente car* $\lim_{k\to+\infty} \frac{u_{k+1}}{u_k} = (1-\varepsilon_0) < 1$. *On en déduit que* $E(T) < +\infty$.

2. On suppose qu'il existe $\varepsilon_0 \in]0,1[$ et $n_0 \in \mathbb{N}^*$ t.q. $P(\{n+n_0 \geq T > n\}) \geq \varepsilon_0 P(\{T > n\})$ pour tout $n \in \mathbb{N}$. Montrer que l'inégalité (11.17) est vraie.

Corrigé – *On montre (11.17) par récurrence sur k. Pour* $k = 1$, *l'inégalité (11.17) est vraie puisque* $P(\Omega) = 1$. *Soit* $k \in \mathbb{N}^*$, *on suppose que (11.17) est vraie (pour cette valeur de k) et on démontre (11.17) pour* $k+1$. *D'après l'hypothèse de cette question, on a*

$$P(\{(k+1)n_0 \geq T > kn_0\}) \geq \varepsilon_0 P(\{T > kn_0\}),$$

ce qui donne

$$P(\{T > kn_0\}) - P(\{T > (k+1)n_0\}) \geq \varepsilon_0 P(\{T > kn_0\}),$$

et donc

$$P(\{T > (k+1)n_0\}) \le (1-\varepsilon_0)P(\{T > kn_0\}) \le (1-\varepsilon_0)^k.$$

Ce qui termine la récurrence.

3. On suppose qu'il existe $\varepsilon_0 \in]0,1[$ et $n_0 \in \mathbb{N}^*$ t.q. $E(1_{\{T\le n+n_0\}}|\mathcal{B}_n) \ge \varepsilon_0$ p.s., pour tout $n \in \mathbb{N}$. Montrer que $P(\{n+n_0 \ge T > n\}) \ge \varepsilon_0 P(\{T > n\})$ pour tout $n \in \mathbb{N}$. En déduire que $E(T) < \infty$.

   Corrigé – *Soit* $n \in \mathbb{N}$. *On a* $P(\{n+n_0 \ge T > n\}) = E(1_{\{T\le n+n_0\}}1_{\{T>n\}})$. *Comme* $\{T \le n\} \in \mathcal{B}_n$, *on a* $\{T > n\} = \{T \le n\}^c \in \mathcal{B}_n$. *La v.a.r.* $1_{\{T>n\}}$ *est donc* $\mathcal{B}_n$*-mesurable (et, bien sûr, bornée). On a donc (en utilisant* $E(1_{\{T\le n+n_0\}}|\mathcal{B}_n) > \varepsilon_0$ *p.s.)*

   $$E(1_{\{T\le n+n_0\}}1_{\{T>n\}}) = E(E(1_{\{T\le n+n_0\}}|\mathcal{B}_n)1_{\{T>n\}}) \ge \varepsilon_0 E(1_{\{T>n\}}) = \varepsilon_0 P(\{T > n\}).$$

   Ce qui donne bien $P(\{n+n_0 \ge T > n\}) \ge \varepsilon_0 P(\{T > n\})$.

   Par la question 2, (11.17) est vraie et donc, par la question 1, $E(T) < +\infty$.

Exercice 11.31 (On joue à pile ou face...)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(J_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r.i.i.d. ne prenant que les valeurs 1 et −1 et t.q. $P(\{J_n = 1\}) = P(\{J_n = -1\}) = 1/2$. On pose $X_0 = 0$ et, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$. Soit a et b deux entiers strictement positifs et, pour tout $\omega \in \Omega$,

$$\begin{aligned} T(\omega) &= \inf\{n \ge 0 \text{ t.q. } X_n(\omega) = -a \text{ ou } X_n(\omega) = b\} \\ &\qquad \text{s'il existe } n \in \mathbb{N} \text{ t.q. } X_n(\omega) \in \{-a, b\}, \\ T(\omega) &= +\infty \text{ sinon.} \end{aligned}$$

On note $p_a = P(\{\omega \in \Omega, \text{ t.q. } X_{T(\omega)}(\omega) = -a\})$ (p_a est donc la probabilité que X_n atteigne $-a$ avant b). Pour $n \in \mathbb{N}$, on désigne par $\mathcal{B}_n$ la tribu engendrée par $X_0, \ldots, X_n$ et on pose, pour $\omega \in \Omega$, $Y_n(\omega) = X_{\min(n,T(\omega))}(\omega)$.

1. Montrer que $(\mathcal{B}_n)_{n\in\mathbb{N}}$ est une filtration et que T est un temps d'arrêt pour cette filtration.

   Corrigé – *On a bien* $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{A}$ *pour tout* $n \in \mathbb{N}$. *La suite* $(\mathcal{B}_n)_{n\in\mathbb{N}}$ *est donc une filtration.*

   Soit $n \in \mathbb{N}$. *La définition de* T *donne*

   $$\{T = n\} = X_n^{-1}(\{-a, b\}) \cap \Big(\cap_{k=0}^{n-1} X_k^{-1}(\{-a, b\}^c) \Big).$$

   Comme $X_n^{-1}(\{-a, b\}) \in \mathcal{B}_n$ *et* $X_k^{-1}(\{-a, b\})^c \in \mathcal{B}_n$ *pour tout* $k \le n$, *on a bien* $\{T = n\} \in \mathcal{B}_n$. *Ce qui montre que* T *est un temps d'arrêt.*

2. Montrer que $E(T) < \infty$.

   Corrigé – *On utilise ici l'exercice 11.30. On pose* $n_0 = a + b$ *et* $\varepsilon_0 = (1/2)^{n_0}$. *On va montrer que, pour tout* $n \in \mathbb{N}$, $E(1_{\{T\le n+n_0\}}|\mathcal{B}_n) \ge \varepsilon_0$ *(l'exercice 11.30 donne alors* $E(T) < +\infty$*).*

Soit $n \in \mathbb{N}$. Pour tout $0 \le k \le n$, la v.a.r. X_k ne prend qu'un nombre fini de valeurs (X_k ne prend que des valeurs entières entre $-k$ et k). La tribu $\mathcal{B}_n$, qui est la tribu engendrée par $X_0, \ldots, X_n$, est donc engendrée par une partition de Ω. Un élément de cette partition est déterminé par les valeurs prises par les X_k, $k = 0, \ldots, n$. On note $A_1, \ldots, A_m$ la partition de Ω engendrant $\mathcal{B}_n$. Il est possible de montrer que $P(A_i) > 0$ pour tout $i = 1, \ldots, m$, mais ceci n'est pas nécessaire pour la suite. En utilisant l'exercice 11.2, on a alors

$$E(1_{\{T \le n+n_0\}} | \mathcal{B}_n) = \sum_{i=1}^{m} \frac{\int_{A_i} 1_{\{T \le n+n_0\}} dP}{P(A_i)} 1_{A_i} = \sum_{i=1}^{m} \frac{P(\{T \le n+n_0\} \cap A_i\}}{P(A_i)} 1_{A_i} \text{ p.s.}. \tag{11.18}$$

Si pour certaines valeurs de i on a $P(A_i) = 0$, l'égalité (11.18) est encore vraie en supprimant ces valeurs de i dans la somme. C'est pour cela qu'il est inutile de démontrer que $P(A_i) > 0$ pour tout i.

Soit $1 \le i \le m$, on montre maintenant que $P(\{T \le n+n_0\} \cap A_i\} \ge \varepsilon_0 P(A_i)$. Comme T est un temps d'arrêt, on a $\{T \le n\} \in \mathcal{B}_n\}$, il existe donc $I \subset \{1, \ldots, m\}$ t.q.

$$\{T \le n\} = \cup_{i \in I} A_i.$$

On distingue maintenant les cas $i \in I$ et $i \notin I$.

Premier cas, $i \in I$. *Ce cas est facile car on alors $T \le n \le n + n_0$ sur A_i et donc $P(\{T \le n+n_0\} \cap A_i\} = P(A_i) \ge \varepsilon_0 P(A_i)$.*

Second cas, $i \notin I$. *Dans ce cas on a $T > n$ sur A_i. Ceci n'est possible que si, pour tout $k \in \{0, \ldots, n\}$, X_k est strictement entre $-a$ et b. Il existe donc $c \in \mathbb{Z}$ t.q. $X_n = c$ sur A_i, avec $-a < c < b$. Comme $c + n_0 = c + a + b > b$, on remarque alors que*

$$A_i \cap \Big(\cap_{j=1}^{n_0} \{J_{n+j} = 1\} \Big) \subset \{T \le n + n_0\}.$$

On a donc

$$P(\{T \le n+n_0\}) \ge P(A_i \cap \Big(\cap_{j=1}^{n_0} \{J_{n+j} = 1\} \Big)).$$

On utilise maintenant l'indépendance de la tribu $\mathcal{B}_n$ (engendrée par $J_1, \ldots, J_n$) et des tribus engendrées par $J_{n+1}, \ldots, J_{n+n_0}$, ce qui est donné par l'hypothèse d'indépendance des J_k et la proposition 2.59, on obtient

$$P(\{T \le n+n_0\}) \ge p(A_i) \prod_{j=1}^{n_0} P(\{J_{n+j} = 1\}) = p(A_i) (\frac{1}{2})^{n_0} = \varepsilon_0 P(A_i).$$

On peut maintenant revenir à (11.18). On obtient

$$E(1_{\{T \le n+n_0\}} | \mathcal{B}_n) \ge \sum_{i=1}^{m} \varepsilon_0 1_{A_i} \text{ p.s.},$$

et on en déduit bien, comme $A_1, \ldots, A_m$ est une partition de Ω, $E(1_{\{T \le n+n_0\}} | \mathcal{B}_n) \ge \varepsilon_0$ p.s.. L'exercice 11.30 donne alors $E(T) < +\infty$.

3. Montrer que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).

Corrigé – *Le deuxième item de l'exemple 11.16 (démontré dans l'exercice 11.25) donne que la suite $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).*

4. Montrer que $(Y_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$) et que $p_a = b/(a+b)$.

   Corrigé – *Comme* T *est un temps d'arrêt, le théorème 11.18 (sur les martingales arrêtées) et la question précédente donnent que la suite* $(Y_n)_{n\in\mathbb{N}}$ *est aussi une martingale (par rapport à la filtration* $(\mathcal{B}_n)_{n\in\mathbb{N}}$*).*

   Comme $(Y_n)_{n\in\mathbb{N}}$ *est une martingale, on a* $E(Y_n) = E(Y_0)$ *pour tout* $n \in \mathbb{N}$*. Mais,* $Y_0 = X_0 = 0$*, on a donc* $E(Y_n) = 0$ *pour tout* $n \in \mathbb{N}$*.*

   On utilise maintenant le fait $E(T) < +\infty$*. Ceci donne que* $T < +\infty$ *p.s.. On en déduit que* Y_n *tend p.s. vers* X_T *(c'est-à-dire vers la v.a.r.* $\omega \mapsto X_{T(\omega)}$ *en posant, par exemple,* $X_\infty = 0$*). Comme* $|Y_n| \le \max\{a,b\}$ *p.s. (et pour tout* $n \in \mathbb{N}$*), le théorème de convergence dominée donne*

   $$E(Y_n) \to E(X_T) \text{ quand } n \to +\infty.$$

   Comme $E(Y_n) = 0$ *on a donc* $E(X_T) = 0$*. On conclut en remarquant que*

   $$E(X_T) = -aP(\{X_T = -a\}) + bP(\{X_T = b\}) = -ap_a + b(1-p_a) = b - (a+b)p_a,$$

   et donc $p_a = \dfrac{b}{a+b}$.

N.B. : On peut aussi montrer que $E(T) = ab$.

Exercice 11.32 (Il est temps de s'arrêter) Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé muni d'une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r., adaptée à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. On suppose aussi que $E[|X_n|] < \infty$ pour tout $n \in \mathbb{N}$.

1. Montrer que pour tout temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$) τ borné,

   $$E[|X_\tau|] < \infty.$$

   Corrigé – τ *est borné, il existe donc* $n \in \mathbb{N}$ *t.q.* $\tau \le n$*, on a alors* $|X_\tau| \le \sum_{k=0}^{n} |X_k|$*, et donc*

   $$E(|X_\tau|) \le \sum_{k=0}^{n} E(|X_k|) < +\infty.$$

   On suppose dans la suite que pour tout temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$) τ borné,

   $$E(X_\tau) = E(X_0).$$

2. Soit $p \in \mathbb{N}$ et $\Lambda \in \mathcal{F}_p$. On définit σ_p par

   $$\sigma_p(\omega) = p \text{ si } \omega \in \Lambda \text{ et } \sigma_p(\omega) = p+1 \text{ si } \omega \notin \Lambda.$$

   Montrer que σ_p est un temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$).

Corrigé – *Soit* $n \in \mathbb{N}$.
Si $n = p$, *on a* $\{\sigma_p = n\} = \Lambda \in \mathcal{F}_p = \mathcal{F}_n$.
Si $n = p+1$, *on a* $\{\sigma_p = n\} = \Lambda^c \in \mathcal{F}_p = \mathcal{F}_{n-1} \subset \mathcal{F}_n$.
Si $n \notin \{p, p+1\}$, *on a* $\{\sigma_p = n\} = \emptyset \in \mathcal{F}_n$.

On a donc $\{\sigma_p = n\} \in \mathcal{F}_n$ *pour tout* $n \in \mathbb{N}$ *et ceci montre que* σ *est un temps d'arrêt par rapport à* $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

3. Soit $p \in \mathbb{N}$ et ν_p défini par $\nu_p(\omega) = p+1$ pour tout $\omega \in \Omega$.
Montrer que ν_p est aussi un temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$).

Corrigé – *Soit* $n \in \mathbb{N}$. *On a* $\{\nu_p = n\} = \Omega \in \mathcal{F}_n$ *si* $n = p+1$ *et* $\{\nu_p = n\} = \emptyset \in \mathcal{F}_n$ *si* $n \neq p+1$. *On a donc* $\{\nu_p = n\} \in \mathcal{F}_n$ *pour tout* $n \in \mathbb{N}$ *et ceci montre que* ν *est un temps d'arrêt par rapport à* $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

4. En remarquant que $X_\tau = X_\tau 1_A + X_\tau 1_{A^c}$ (pour tout temps d'arrêt τ et tout événement A), montrer que $(X_n)_{n\in\mathbb{N}}$ est une martingale.

Corrigé – *Soit* $n \in \mathbb{N}$. *On veut montrer que* $E(X_{n+1}|\mathcal{F}_n) = X_n$ *p.s., c'est-à-dire que*

$$E(X_{n+1}U) = E(X_n U), \tag{11.19}$$

pour toute v.a.r. U $\mathcal{F}_n$*-mesurable bornée.*

On commence par montrer (11.19) si $U = 1_\Lambda$ *avec* $\Lambda \in \mathcal{F}_n$. *Soit donc* $\lambda \in \mathcal{F}_n$. *On utilise alors les temps d'arrêt* σ_n *et* ν_n, *définis dans les deux questions précédentes, avec* $p = n$. *Comme ces deux temps d'arrêt sont bornés on a*

$$E(X_{\nu_n}) = E(X_{\sigma_n}) = E(X_0).$$

Mais, $X_{\sigma_n} = X_{\sigma_n} 1_\Lambda + X_{\sigma_n} 1_{\Lambda^c} = X_n 1_\Lambda + X_{n+1} 1_{\Lambda^c}$ *et* $X_{\nu_n} = X_{n+1}$. *On a donc*

$$E(X_{n+1}) = E(X_n 1_\Lambda + X_{n+1} 1_{\Lambda^c}) = E(X_n 1_\Lambda) + E(X_{n+1} 1_{\Lambda^c}),$$

ce qui donne $E(X_{n+1}(1 - 1_{\Lambda^c})) = E(X_n 1_\Lambda)$. *Comme* $1 - 1_{\Lambda^c} = 1_\Lambda$, *on a donc montré (11.19) pour* $U = 1_\Lambda$.

Par linéarité de l'espérance, on remarque alors que (11.19) est encore pour U *v.a.r. étagée formée avec des éléments de* $\mathcal{F}_n$, *c'est-à-dire toute v.a.r.* U *de la forme* $\sum_{i=1}^k \alpha_i 1_{A_i}$ *avec* $A_i \in \mathcal{F}_n$ *et* $\alpha_i \in \mathbb{R}$.

Soit maintenant U $\mathcal{F}_n$*-mesurable bornée. Il existe alors une suite* $(U_k)_{k\in\mathbb{N}}$ *de v.a.r. étagées formées avec des éléments de* $\mathcal{F}_n$ *t.q.* $U_k \to U$ *p.s. quand* $k \to +\infty$ *et* $|U_k| \le |U|$ *p.s. (et pour tout* k*). Grâce au théorème de convergence dominée, on peut alors passer à la limite quand* $k \to +\infty$ *dans l'égalité* $E(X_{n+1}U_k) = E(X_n U_k)$ *et on obtient bien (11.19). Ceci prouve que* $E(X_{n+1}|\mathcal{F}_n) = X_n$ *p.s. et donc que* $(X_n)_{n\in\mathbb{N}}$ *est une martingale.*

That's all folks !

Mesure-Intégration-Probabilités, DM2, janvier 2021

Exercice 1 (Une condition suffisante pour que deux v.a.r. suivent une loi normale)**.** Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé de X une v.a.r. telle $E(X) = 0$ et $E(X^2) < +\infty$. On note ϕ la fonction caractéristique de X.

1. Quelle régularité minimale a-t-on sur ϕ ? (plus précisément peut t-on assurer que ϕ est continue ? de classe C^1 ? de classe C^2 ? ... de classe C^∞ ?). Quelle est la valeur de $\phi(0)$? Quelle est la valeur de $\phi'(0)$?

On se donne maintenant une deuxième v.a.r., notée Y, de même loi que X. On suppose que X et Y sont indépendantes et que les v.a.r. $X + Y$ et $X - Y$ sont aussi indépendantes. On pose $\sigma^2 = E(X^2)$.

2. Montrer que $E((X - Y)^2 e^{it(X+Y)}) = 2\sigma^2 \phi(t)^2$ pour tout $t \in \mathbb{R}$.
3. Montrer que $-\phi''(t)\phi(t) + \phi'(t)^2 = \sigma^2 \phi(t)^2$ pour tout $t \in \mathbb{R}$.
4. Montrer que la loi de X est la loi $\mathcal{N}(0, \sigma^2)$.

Exercice 2 (Constante d'Euler)**.**
Soit $(X_k)_{k\geq 1}$ une famille i.i.d. de v.a. de loi exponentielle $\mathcal{E}(1)$. Pour tout $n \geq 1$, on introduit les v.a. M_n et Z_n définies par :

$$M_n = \max(X_1, \ldots, X_n)\ ,\ Z_n = M_n - \ln(n).$$

On introduit également les fonctions $f : \mathbb{R} \to \mathbb{R}_+$ et $F : \mathbb{R} \to]0, 1[$ définies par :

$$\forall x \in \mathbb{R}\ ,\ f(x) = e^{-x} e^{-e^{-x}}\ ,\ F(x) = e^{-e^{-x}}.$$

1. Calculer les limites $\lim_{x\to-\infty} F(x)$ et $\lim_{x\to+\infty} F(x)$, ainsi que la fonction dérivée $F'(x)$. En déduire que f est une densité de probabilité sur $\mathbb{R}$.
2. Montrer que les fonctions de répartition des v.a. M_n et Z_n sont respectivement données par :

$$\forall x \in \mathbb{R}\ ,\ F_{M_n}(x) = (1 - e^{-x})^n 1_{x\geq 0}\ ,\ F_{Z_n}(x) = \left(1 - \frac{e^{-x}}{n}\right)^n 1_{x\geq -\ln(n)}.$$

   En déduire la densité f_n de la loi de Z_n.
3. Montrer que la suite de fonctions $(f_n)_{n\geq 1}$ converge simplement vers la fonction f.
4. Montrer que pour tout $x \geq 0$ et tout $n \geq 1$, on a $0 \leq f_n(x) \leq e^{-x}$.
5. Soit $x \leq 0$ et $n \geq 1$. On veut montrer que $0 \leq f_n(x) \leq e^x$ pour n assez grand, indépendamment de x.
   (a) Montrer que cette inégalité est vraie si $x < -\ln(n)$.
   (b) Montrer que si $x \in [-\ln(n), 0]$ et si $-\frac{n-1}{n} e^{-x} \leq 2x$, alors $0 \leq f_n(x) \leq e^x$.
   (c) Conclure en remarquant que $e^z \geq ez$ pour tout $z \geq 0$.
6. Déduire des questions précédentes qu'il existe $N \in \mathbb{N}$ tel que $|f_n(x)| \leq e^{-|x|}$ pour tout $x \in \mathbb{R}$ et tout $n \geq N$.
7. Montrer que la suite $(\mathbb{E}[Z_n])_{n\geq 1}$ est convergente, de limite $\mathbb{E}[Z]$, où Z suit la loi de densité f.
8. Montrer que $\mathbb{E}[M_n] = \int_0^1 \frac{1-t^n}{1-t} dt$ (on pourra utiliser la fonction de répartition F_{M_n}), puis que $\mathbb{E}[M_n] = \sum_{k=1}^n \frac{1}{k}$.
9. Déduire des résultats montrés dans l'exercice que la suite $\left(\sum_{k=1}^n \frac{1}{k} - \ln(n)\right)_{n\geq 1}$ est convergente, et écrire sa limite sous la forme d'une intégrale.