Centre de Télé-Enseignement Sciences
Université de Provence

# MASTER SCIENCES
# Mention Mathématiques et Applications -- ( 2M1MAP )

| Expédition dans la semaine n° | Etape | Code UE | N° d'envoi de l'UE |
|---|---|---|---|
| | **2M1MAP** | 2TMAP14 | 5 |

***Nom de l'UE : Théorie de la mesure et Probabilités (UE 1-4)***

– **Contenu de l'envoi : Polycopié, chapitres 10 et 11 et corrigés des exercices des chapitres 10 et 11**

**- Guide du travail à effectuer**

Semaine 1 :
Etudier le chapitre 10, sections 1, 2, 3 (Transformation de Fourier)
Exercices proposés (avec corrigés) : 10.1, 10.5 (1ere question)

Semaine 2 :
Etudier le chapitre 10, sections 4 et 6 (Transf. de Fourier dans L2, Fonctions caractéristiques)
Exercices proposés (avec corrigés) : 10.10, 10.11

Semaine 3 :
Etudier le chapitre 11, section 1 (Espérance conditionnelle)
Exercices proposés (avec corrigés) : 11.1, 11.3, 11.6

Semaine 4 :
Etudier le chapitre 11, section 2 (Martingales)
Exercices proposés (avec corrigés) : 11.9, 11.18

Le deuxième devoir est à rendre à la réception de ce cinquième envoi (mi-mars)

**- Coordonnées de l'enseignant responsable de l'envoi**

Thierry Gallouet, LATP-CMI, 39 rue Joliot Curie, 13453 marseille cedex 13.
email : gallouet@cmi.univ-mrs.fr
fax : 04 91 11 35 52

Vous pouvez aussi consulter la page web: http://www.cmi.univ-mrs.fr/~gallouet/tele.d
et me poser des questions 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 $\mathbb{C}$). La notion de transformée de Fourier permet d'analyser les fonctions définies de $\mathbb{R}$ (ou $\mathbb{R}^N$) dans $\mathbb{R}$ (ou $\mathbb{C}$). 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 considèrera l'espace mesuré $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et on notera $d\lambda_N(x) = dx$. Soit $N \geq 1$, les espaces $\mathcal{L}^1_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L^1_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ ont été définis dans la section 4.10 et les espaces $\mathcal{L}^2_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L^2_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ ont été définis dans la section 6.2. On rappelle aussi que si $f$ est une fonction définie de $\mathbb{R}^N$ dans $\mathbb{C}$, la fonction $f$ est mesurable si et seulement si ses parties réelle et imaginaire sont mesurables (chaque ensemble, $\mathbb{R}^N$, $\mathbb{R}$ ou $\mathbb{C}$, étant muni de sa tribu borélienne). On peut, bien sûr, aussi définir les espaces $\mathcal{L}^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ pour tout $p \in [1, \infty]$.

**Définition 10.1 (Espaces $L^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$)** *Soit $N \geq 1$, $p \in [1,\infty]$ et $f$ une fonction mesurable de $\mathbb{R}^N$ dans $\mathbb{C}$ (c'est-à-dire $f^{-1}(A) \in \mathcal{B}(\mathbb{R}^N)$ pour tout $A \in \mathcal{B}(\mathbb{C})$). 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 $\mathcal{L}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (vue au chapitre 6). L'espace $L^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ 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 10.1** 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 $L^p_{\mathbb{C}}(\mathbb{R}^N)$ l'espace $L^p_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, pour $p \in [1, \infty]$.

**Définition 10.2 (Transformée de Fourier dans $L^1$)**

*Soit $N \geq 1$ et $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. 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 $L^1_{\mathbb{C}}(\mathbb{R}^N)$. On définit alors $\hat{f}(t)$ par :*

$$\hat{f}(t) = (2\pi)^{-\frac{N}{2}} \int f(x)e^{-ix\cdot t}dx. \tag{10.1}$$

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

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

DÉMONSTRATION :

– Le théorème de continuité sous le signe $\int$ (théorème 4.9) appliqué à la fonction $(x,t) \mapsto e^{-ix.t}f(x)$ entraîne immédiatement que $\hat{f}$ est continue.
– On montre maintenant que $\hat{f} \in C_0(\mathbb{R}, \mathbb{R})$.
  – Cas $N = 1$. On remarque que pour $t \neq 0$, on a, comme $e^{i\pi} = -1$,

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

  et donc avec un changement de variable simple,

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

  On en déduit que

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

  et donc que $|\hat{f}(t)| \leq \frac{1}{2}(2\pi)^{-\frac{1}{2}}\|f(\cdot) - f(\cdot + \frac{\pi}{t})\|_1$. Le théorème de continuité en moyenne dans $L^1$ (théorème 5.6) donne alors le fait que $\hat{f}(t) \to 0$ 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 $\mathbb{R}^N$,

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

  et donc avec un changement de variable simple,

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

  On en déduit que $|\hat{f}(t)| \leq \frac{1}{2}(2\pi)^{-\frac{N}{2}}\|f(\cdot) - f(\cdot + \frac{\pi}{t_j}e_j)\|_1$. Le théorème de continuité en moyenne dans $L^1$ (théorème 8.2) donne alors le fait que $\hat{f}(t) \to 0$ quand $|t| \to \infty$.

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

**Proposition 10.2** *Soient $f$ et $g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$, alors $\widehat{f \star g} = (2\pi)^{\frac{N}{2}} \hat{f}\hat{g}$.*

DÉMONSTRATION : Par la proposition 7.9, on $f \star g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ et pour p.p. $x \in \mathbb{R}^N$, $f \star g(x) = \int f(x-y)g(y)dy$. On a donc, pour tout $t \in \mathbb{R}^N$,

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

En appliquant le théorème de Fubini (théorème 7.3) 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 $\mathbb{R}^{2N}$ car son module est la fonction $(x,y) \mapsto |f(x-y)g(y)|$ dont l'intégrale sur $\mathbb{R}^{2N}$ est égale à $\|f\|_1\|g\|_1$), on obtient :

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

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

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

ce qui est le résultat annoncé. ■

### 10.2.2 Théorème d'inversion

Il est naturel de se poser les deux questions suivantes :

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

(ii) 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 10.1 (Inversion partielle de la transformée de Fourier)** *Soit $N \geq 1$ et $f \in L^1_{\mathbb{C}}$ t.q. $\hat{f} \in L^1_{\mathbb{C}}$. On a alors $f = \hat{\hat{f}}(-.)$ p.p., c'est-à-dire :*

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

DÉMONSTRATION : La démonstration fait l'objet de l'exercice 10.1. ■

Une conséquence de ce théorème est l'injectivité de l'application $F$, qui fournit donc une réponse positive à la question (i). En effet, soient $f$ et $g \in L^1_{\mathbb{C}}$ t.q. $\hat{f} = \hat{g}$ ; alors par linéarité, $\widehat{f-g} = 0$ et donc $\widehat{f-g} \in L^1_{\mathbb{C}}$. En appliquant le théorème d'inversion, on a donc $f = g$ p.p..
Ce théorème apporte aussi une réponse partielle à la question (ii) : on peut calculer $f$ à partir de $\hat{f}$ dès que $\hat{f} \in L^1$. Il faut remarquer à ce propos que $L^1$ n'est pas stable par transformation de Fourier (voir exercice 10.2).

### 10.2.3 Régularité et décroissance à l'infini

**Proposition 10.3 (Différentiabilité, dimension 1)**

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

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 5.8).

Le deuxième item est une conséquence immédiate du théorème 4.10 (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 $\alpha = (\alpha_1,\ldots,\alpha_N)^t \in \mathbb{N}^N$ un "multi-indice" et $f$ une fonction de $\mathbb{R}^N$ dans $\mathbb{C}$. On définit $|\alpha| = \alpha_1 + \alpha_2 + \ldots \alpha_N$ et :

$$D^\alpha f = \Big(\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}}\Big)f.$$

**Proposition 10.4 (Différentiabilité, dimension N)**

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

La proposition 10.4 montre que la dérivabilité de $f$ entraîne la décroissance de $\hat{f}$ à l'infini ("plus $f$ est dérivable, plus $\hat{f}$ décroît vite à l'infini"), et réciproquement. Cette remarque incite à définir l'espace des fonctions à décroissance rapide (souvent appelé "espace de Schwartz"), noté $\mathcal{S}(\mathbb{R}^N,\mathbb{C})$ ou $\mathcal{S}_N$ en abrégé.

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

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 $\alpha$ et $\beta$ dans la définition de $\mathcal{S}_N$, on remarque qu'il existe $C \in \mathbb{R}$ t.q. $(1+|x|^{N+1})|f(x)| \le C$. On en déduit que $f \in L^1_C(\mathbb{R}^N)$. Plus généralement, si $f \in \mathcal{S}_{\mathbb{N}}$, 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 10.5.

**Proposition 10.5** *Soit $N \ge 1$ et $f \in \mathcal{S}(\mathbb{R}^N,\mathbb{C})$. Pour tout $\alpha,\beta \in \mathbb{N}^N$ on a :*

$$D^\alpha\Big(\widehat{(-i\cdot)^\beta f}\Big) = (i\cdot)^\alpha(\widehat{-i\cdot)^\beta} f = (i\cdot)^\alpha D^\beta \hat{f}, \tag{10.2}$$

$$(\widehat{-i\cdot)^\alpha D^\beta} f = D^\alpha \widehat{D^\beta f} = D^\alpha[(i\cdot)^\beta \hat{f}]. \tag{10.3}$$

DÉMONSTRATION : La démonstration est une adaptation simple de celle de la proposition 10.4. ■

La proposition 10.5 et le théorème 10.1 nous permettent alors de remarquer que la transformée de Fourier est une bijection de $\mathcal{S}_N$ dans $\mathcal{S}_N$.

**Proposition 10.6** *Soit $N \geq 1$. L'application $F$ 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 $f = \widehat{\hat{f}}(-\cdot)$.*

DÉMONSTRATION : En utilisant la proposition 10.5, on montre facilement que $F$ envoie $\mathcal{S}_N$ dans $\mathcal{S}_N$. Le théorème d'inversion (théorème 10.1) donne alors que $f$ est injective et que $f = \widehat{\hat{f}}(-.)$ pour tout $f \in \mathcal{S}_N$. De cette dernière formule, on déduit que $F$ 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 10.3 définit la fonction $\hat{m}$ (continue et bornée de $\mathbb{R}^N$ dans $C$) 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 $\hat{m}(t) = \hat{f}(t)$ pour tout $t \in \mathbb{R}^N$.

**Définition 10.3 (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 $\hat{m}(t)$ par :*

$$\hat{m}(t) = (2\pi)^{-\frac{N}{2}} \int e^{-ix\cdot t} dm(x). \tag{10.4}$$

*La fonction $\hat{m}$ (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 10.7** *Soit $N \geq 1$. Soit $m$ une mesure signée sur les boréliens de $\mathbb{R}^N$. La fonction $\hat{m}$ appartient à $C_b(\mathbb{R}^N, \mathbb{C})$.*

DÉMONSTRATION : Le fait que $\hat{m}$ est bornée est simple. On utilise la décomposition de Hahn (proposition 2.6), 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 $|\hat{m}(t)| \leq |m|(\mathbb{R}^N) < +\infty$, où $|m| = m^+ + m^-$.

La fait que $\hat{m}$ est continue est une conséquence immédiate du théorème de continuité sous le signe $\int$ (théorème 4.9) 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 $\hat{m}$ caractérise la mesure signée $m$ et si on peut retrouver $m$ à partir de $\hat{m}$. La proposition suivante s'intéresse à ces deux questions.

**Proposition 10.8** *Soit $N \geq 1$.*

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

DÉMONSTRATION : La démonstration de cette proposition fait l'objet de l'exercice 10.5. ■

## 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 $dt = d\lambda_N(t)$) :

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

Il est clair que la définition de $\hat{f}$ 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 10.9** *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 $\hat{f}, \hat{g} \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $(f/g)_2 = (\hat{f}/\hat{g})_2$. En particulier, $\|f\|_2 = \|\hat{f}\|_2$.*

DÉMONSTRATION : Soit $f, g \in \mathcal{S}_N$. Comme $f, \hat{f} \in \mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$, on peut appliquer le théorème d'inversion (théorème 10.1). Il donne $f = \hat{\hat{f}}(-\cdot)$ et donc :

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

On utilise maintenant le théorème de Fubini (théorème 7.3). Il s'applique car $\hat{f}, \bar{g} \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. On obtient :

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

ce qui termine la démonstration. ■

La proposition 10.9 permet de définir, par un argument de densité, la transformée de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

**Théorème 10.2** *Soit $N \geq 1$. Il existe une application linéaire continue $\bar{F}$ de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ t.q. :*

*1. Si $f \in L^1_{\mathbb{C}}(\mathbb{R}^N) \cap L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a alors $\bar{F}(f) = \hat{f}$ p.p..*

*2. Pour tout $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a $(f/g)_2 = (\bar{F}(f)/\bar{F}(g))_2$.*

*3. Pour tout $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a $f = \bar{F}(\bar{F}(f))(-\cdot)$.*

*4. $\bar{F}$ est une bijection de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.*

*Pour $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, $\bar{F}(f)$ s'appelle la transformée de Fourier de $f$. Compte tenu du premier item, on notera en général, $\hat{f}$ la transformée de Fourier de $f$ si $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ (et alors $\hat{f} \in C_0(\mathbb{R}^N, \mathbb{C})$) ou si $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ (et alors $\hat{f} \in L^2_{\mathbb{C}}(\mathbb{R}^N)$).*

DÉMONSTRATION : L'application $f \mapsto \hat{f}$ est définie sur $\mathcal{S}_N$, qui est un sous espace de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, et prend ses valeurs dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (car $\mathcal{S}_N \subset L^2_{\mathbb{C}}(\mathbb{R}^N)$, en confondant, comme d'habitude, un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ avec l'un de ses représentants). Comme cette application est linéaire, continue pour la norme de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, et que $\mathcal{S}_N$ est dense dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (car $\mathcal{S}_N \supset C_c^\infty(\mathbb{R}^N, \mathbb{C})$ et $C_c^\infty(\mathbb{R}^N, \mathbb{C})$ est dense dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, voir le théorème 8.4), on en déduit que cette application se prolonge en une application, notée $\bar{F}$, de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Plus précisément, soit $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$. Il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \to f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \to \infty$. La suite $(f_n)_{n\in\mathbb{N}}$ est donc de Cauchy dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.9 donne alors que la suite $(\hat{f}_n)_{n\in\mathbb{N}}$ est donc aussi de Cauchy dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Elle converge donc dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On aimerait définir $\bar{F}(f)$ comme étant la limite (dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) de la suite $(\hat{f}_n)_{n\in\mathbb{N}}$. 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 $L^2_{\mathbb{C}}(\mathbb{R}^N)$ vers $f$. Or, ce dernier point est facile car si $(g_n)_{n\in\mathbb{N}}$ est une autre suite convergeant dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ vers $f$, on a $\|\hat{f}_n - \hat{g}_n\|_2 = \|f_n - g_n\|_2 \to 0$ quand $n \to \infty$. On a ainsi défini $\bar{F}$ de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans lui même.

La linéarité de $\bar{F}$ découle immédiatement du fait que l'application $f \mapsto \hat{f}$ est linéaire de $\mathcal{S}_N$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Enfin, soit $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \to f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.9 donne que $\|\hat{f}_n\|_2 = \|f_n\|_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quend $n \to \infty$, on en déduit $\|\bar{F}(f)\|_2 = \|f\|_2$. Ce qui prouve la continuité de $\bar{F}$ 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 8.4, 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 8.4, la suite construite pour converger vers $f$ dans $L^p$ ne dépend pas de $p$). On en déduit que $\hat{f}_n \to \hat{f}$ uniformément sur $\mathbb{R}^N$ lorsque $n \to +\infty$ (car $f_n \to f$ dans $L^1_{\mathbb{C}}(\mathbb{R}^N)$) et que $\hat{f}_n \to \bar{F}(f)$ 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 $\hat{f}_n \to \bar{F}(f)$ p.p. quand $n \to \infty$. On en déduit bien que $\hat{f} = \bar{F}(f)$ 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 10.9 donne $(\hat{f}_n/\hat{g}_n)_2 = (f_n/g_n)_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quand $n \to \infty$ on obtient bien $(\overline{F}(f)/\overline{F}(g))_2 = (f/g)_2$.
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 $\hat{f}_n \to \bar{F}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, ce qui donne aussi $\hat{f}_n(-\cdot) \to \bar{F}(f)(-\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et donc $\hat{\hat{f}}_n(-\cdot) \to \overline{F}(\overline{F}(f))(-\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \to +\infty$. La proposition 10.6 donne $f_n = \hat{\hat{f}}_n(-\cdot)$ pour tout $n \in \mathbb{N}$. On en déduit donc (par unicité de la limite dans $L^2$) que $f = \overline{F}(\overline{F}(f))(-\cdot)$ p.p..
4. L'injectivité de $\bar{F}$ (de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) découle du fait que $\|\bar{F}(f)\|_2 = \|f\|_2$ et que $\bar{F}$ est linéaire. La surjectivité est une conséquence immédiate du troisième item.

■

## 10.5 Résolution d'une EDO ou d'une EDP

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

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

$$a_N f^{(N)}(x) + \ldots + a_1 f'(x) + a_0 f(x) = g(x), \text{ pour tout } x \in \mathbb{R}. \tag{10.5}$$

Si $f \in \mathcal{S}_1$ vérifie (10.5), elle vérifie nécessairement, par transformation de Fourier :

$$a_N \widehat{f^{(N)}}(t) + \ldots + a_1 \widehat{f'}(t) + a_0 \hat{f}(t) = \hat{g}(t), \text{ pour tout } t \in \mathbb{R}. \tag{10.6}$$

c'est-à-dire :

$$a_N (it)^N \hat{f}(t) + \ldots + a_1 it \hat{f}(t) + a_0 \hat{f}(t) = \hat{g}(t), \text{ pour tout } t \in \mathbb{R}. \tag{10.7}$$

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

$$\hat{f}(t) = \frac{\hat{g}(t)}{p(t)}. \tag{10.8}$$

Comme $g \in \mathcal{S}_1$ et que $p$ ne s'annule par sur $\mathbb{R}$, on peut montrer que $\frac{\hat{g}}{p} \in \mathcal{S}_1$. En utilisant maintenant le théorème d'inversion, ou la proposition 10.6, on obtient :

$$f(t) = \hat{\hat{f}}(-t) = \widehat{\left(\frac{\hat{g}}{p}\right)}(-t), \text{ pour tout } t \in \mathbb{R}. \tag{10.9}$$

On a donc montré que $f$ est nécessairement donnée par (10.9). Réciproquement, il est facile de voir que la fonction donnée par (10.9) est solution de (10.5), c'est donc l'unique solution dans $\mathcal{S}_1$ de (10.5) (en supposant que $p$ ne s'annule par sur $\mathbb{R}$).

Soit $N \geq 1$, on cherche $u : \mathbb{R}^N \to \mathbb{R}$, de classe $C^2$ (c'est-à-dire deux fois continûment dérivable) t.q.

$$-\Delta u(x) = 0, \text{ pour tout } x \in \mathbb{R}^N. \tag{10.10}$$

Cherchons $u \in \mathcal{S}_N$ (et donc $\Delta u \in \mathcal{S}_N$) solution de (10.10). On a donc $\widehat{\Delta u} = 0$ (partout sur $\mathbb{R}^N$), c'est-à-dire $|t|^2\hat{u}(t) = 0$ et donc $\hat{u}(t) = 0$, pour tout $t \neq 0$. Comme $\hat{u}$ est continue, ceci entraîne que $\hat{u} = 0$ (partout sur $\mathbb{R}^N$), et donc $u = 0$. La fonction identiquement égale à 0 est donc la seule solution de (10.10) 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 $\Delta u$). On obtient encore que la seule solution de (10.10) 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 (10.10) (et on peut montrer que ce sont les seules fonctions, de classe $C^2$et bornées, solutions de (10.10)).

## 10.6 Fonction caractéristique d'un vecteur aléatoire

**Définition 10.4** *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 $\Omega$ (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 $\varphi_X = (2\pi)^{d/2}\hat{p}_X(-\cdot)$. La section 10.3 donne alors quelques propriétés élémentaires de la fonction caractéristique d'un v..a..

**Proposition 10.10** *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. $\varphi_X \in C_b(\mathbb{R}^d, C)$.*

*2. La loi de $X$ est entièrement déterminée par $\varphi_X$ (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 :*

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

DÉMONSTRATION : Comme $\varphi_X = (2\pi)^{d/2}\hat{p}_X(-\cdot)$, le premier item est donnée par la proposition 10.7 (qui donne $\hat{p}_X \in C_b(\mathbb{R}^d, \mathbb{C})$).

Le deuxième item est donnée par le premier item de la proposition 10.8.

Pour le troisième item, on remarque que le deuxième item de la proposition 10.8 donne $p_X = f\lambda_d$ avec $f = \hat{\hat{p_X}}(-\cdot)$, ce qui donne $\lambda_d$-p.s. en $x \in \mathbb{R}^d$ :

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

■

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 10.11** *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 9.2 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 10.8, 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, que nous donnons maintenant, sans démonstration.

**Proposition 10.12** *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $d > 1$, $(X_n)n \in \mathbb{N}$ un suite de v.a. de dimension d et $X$ 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 $u \in \mathbb{R}^d$.*

On termine cette section en donnant la fonction caractéristique d'un vecteur gaussien.

**Proposition 10.13** *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.*

*1. Soit $X$ une v.a.r. gaussienne. On note $m$ son espérance (donc, $m = E(X) \in \mathbb{R}$) et $\sigma^2$ sa variance (donc, $\sigma^2 = E((X-m)^2) \geq 0$) 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 (donc, $m = E(X) \in \mathbb{R}^d$) et $D$ sa matrice de covariance (donc, $D$ est une matrice symétrique, 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 $X$) de sorte que $X \sim \mathcal{N}(m, D)$ (proposition 9.10). On a alors :*

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

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 4.10) et donne que $\psi$ 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 $\varphi_X(u) = e^{imu}e^{\frac{-\sigma^2u^2}{2}}$, 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 169) :

$$\sigma^2 = E((X\cdot u - m\cdot u)^2) = E(u^t(X-m)(X-m)^tu) = u^t\text{Cov}(X)u = u^tDu.$$

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

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

ce qui est bien la fromule annoncée (car $u^tDu = Du\cdot u$). ■

## 10.7 Exercices

### 10.7.1 Transformation de Fourier dans $L^1$

**Exercice 10.1 (Résultat partiel d'inversion de Fourier dans $L^1$)** *Corrigé 178 page 475*
Soit $H(t) = e^{-|t|}, t \in \mathbb{R}$. On pose, pour $\lambda > 0$ :

$$h_\lambda(x) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} H(\lambda t) e^{itx} dt, x \in \mathbb{R}. \tag{10.11}$$

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

$$f \star h_\lambda(x) = \int_{\mathbb{R}} H(\lambda t) \hat{f}(t) e^{ixt} dt. \tag{10.12}$$

3. Soit $g$ une fonction mesurable bornée de $\mathbb{R}$ dans $\mathbb{C}$, continue en 0. Montrer que $g \star h_\lambda(0) \to \sqrt{2\pi} g(0)$ quand $\lambda \to 0$. [Utiliser 1. et le théorème de convergence dominée.]
4. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que :

$$\|f \star h_\lambda - \sqrt{2\pi} f\|_1 \to 0 \text{ lorsque } \lambda \to 0. \tag{10.13}$$

[On pourra utiliser la continuité en moyenne et la question précédente avec $g(y) = \int |f(x-y) - f(x)| dx$.]
5. Déduire de ce qui précède le théorème d'inversion de Fourier, théorème 10.1.

**Exercice 10.2** 1. Calculer la transformée de Fourier de $1_{[-a,a]}$, $a \in \mathbb{R}_+$. En déduire que $L^1$ n'est pas stable par transformation de Fourier.
2. On pose $g_n = 1_{[-n,n]}$. Calculer $f \star g_n$, et montrer qu'il existe $h_n \in L^1$ t.q. $\hat{h}_n = f \star g_n$. Montrer que la suite $f \star g_n$ est bornée dans $C_0(\mathbb{R}, \mathbb{R})$ alors que la suite $h_n$ n'est pas bornée dans $L^1$. En déduire que la transformée de Fourier n'est pas surjective de $L^1$ dans $C_0$.

**Exercice 10.3** Soit $f \in \mathcal{S}$, on pose $L(f)(x) = f"(x) + xf(x)$.
1. Montrer que $L(f) = 0$ entraîne $f = 0$.
2. Soit $k \in \mathcal{S}$, montrer que l'équation différentielle $h'(t) = k(t)$ a une solution dans $\mathcal{S}$ si et seulement si $\int k(t) dt = 0$.
3. Soit $g \in \mathcal{S}$, étudier l'existence et l'unicité des solutions dans $\mathcal{S}$ de l'équation $L(f) = g$. On pourra remarquer que pour $h \in \mathcal{S}$, on a :

$$i \frac{d}{dt}(h e^{-i\frac{t^3}{3}}) - t^2 (h e^{-i\frac{t^3}{3}}) = i h'(t) e^{-i\frac{t^3}{3}}.$$

**Exercice 10.4** On note $L^p$ l'espace $L^p_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.
1. Soient $f, g \in L^1$, montrer que $f\hat{g} \in L^1$, $g\hat{f} \in L^1$ et $\int f\hat{g} d\lambda = \int g\hat{f} d\lambda$.
2. Soit $B = [-\frac{1}{2}, \frac{1}{2}]$. Montrer que $1_B \star 1_B(t) = (1 - |t|)^+$.
3. On pose $\theta_n = (1 - \frac{|t|}{n})^+, n \in \mathbb{N}^\star$. Déduire de la question précédente que :

$$\hat{\theta}_n(y) = \frac{4}{\sqrt{2\pi}} \frac{\sin^2(\frac{ny}{2})}{ny^2}, \forall y \in \mathbb{R}.$$

[*Se ramener à $\theta_1$...*]

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

### 10.7.2 Transformée de Fourier d'une mesure signée

**Exercice 10.5 (Caractérisation de $m$ par $\hat{m}$)** *Corrigé 179 page 477*
Soit $d \geq 1$.

1. Soit $m$ et $\mu$ deux mesures signées sur les boréliens de $\mathbb{R}^d$. On suppose que $\hat{m} = \hat{\mu}$.
   (a) Soit $\varphi \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que $\int \hat{\varphi} dm = \int \hat{\varphi} d\mu$.
   (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 $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).
   (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 $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).
2. Soit $m$ une mesure signée sur les boréliens de $\mathbb{R}^d$. On suppose que $\hat{m} \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que $m$ est la mesure de densité $f$ par rapport à la mesure de Lebesgue avec $f = \hat{\hat{m}}(-\cdot)$.

### 10.7.3 Transformation de Fourier dans $L^2$

**Exercice 10.6** On note ici $\hat{f}$ la transformée de Fourier, pour $f \in L^1$ ou $L^2$.

1. Soient $f, g \in \mathcal{S}$, Montrer que $\widehat{fg} = \frac{1}{\sqrt{2\pi}} \hat{f} \star \hat{g}$.
2. Soient $f, g \in L^2$, Montrer que $\widehat{fg} = \frac{1}{\sqrt{2\pi}} \hat{f} \star \hat{g}$.

### 10.7.4 Fonction Caractéristique d'une v.a.r.

**Exercice 10.7 (Calcul de fonctions caractéristiques)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités et $X$ une v.a. réelle. Calculer la fonction caractéristique $\varphi_X$ 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 $p$ : $\mathcal{P}[X = 1] = p = 1 - \mathcal{P}[X = 0]$).
3. $X$ suit une loi exponentielle de paramètre $\lambda$ (avec $\lambda > 0$).

**Exercice 10.8 (Loi normale et vecteur gaussien)**
Soit $(\Omega, \mathcal{A}, p)$ un espace probabilisé, $d \geq 1$ et $X$ un v.a. de dimension $d$. Soit $m \in \mathbb{R}^d$ et $D$ une matrice s.d.p. (de taille $d \times d$). Si $X \sim \mathcal{N}(m, D)$, où $\mathcal{N}(m, D)$ est définie par la définition 9.3, montrer que $X$ est un vecteur gaussien.

**Exercice 10.9 (Vecteurs gaussiens et densité)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités, $d \geq 1$ 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. Si $D$ est inversible, montrer que la loi de $X$ est la loi $\mathcal{N}(m, D)$ donnée dans la définition 9.3.

**Exercice 10.10 (V.a. gaussiens, indépendance, covariance)** *Corrigé 180 page 478*

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités, $d \geq 1$ et $X = (X_1, \ldots, X_d)$ 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 $\mathrm{Cov}(X_1, X_2) = 0$.
   (b) On suppose que $X$ est un vecteur gaussien et que $\mathrm{Cov}(X_1, X_2) = 0$. Montrer que $X_1$ et $X_2$ sont indépendantes.
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 $(\Omega, \mathcal{A}, P)$ et $X_1$, $X_2$ de manière à avoir $\mathrm{Cov}(X_1, X_2) = 0$ sans que $X_1$, $X_2$ soient indépendantes, voir l'exercice 4.44.]
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.

**Exercice 10.11 (Suite de v.a.r.i.i.d. de Poisson)** *Corrigé 181 page 480*

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $X$ une v.a. de Poisson de paramètre $\lambda$ ($\lambda > 0$). 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 $\varphi_X$ de $X$.
2. Soit $(X_n)_{n \in \mathbb{N}^\star}$ une suite de v.a. indépendantes et de Poisson de paramètre $\lambda$.
   (a) Soit $n > 1$. Déduire de la question 1 la loi de la v.a. $Y_n = \sum_{p=1}^n X_p$.
   (b) Utiliser le théorème central limite pour démontrer que

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

**Exercice 10.12 (Sur les lois des grands nombres)** *Corrigé 182 page 481*

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 $\lambda$ désigne la mesure de Lebesgue sur les boréliens de $\mathbb{R}$).

1. Montrer que $X$ n'est pas une v.a.r. intégrable.
2. Soit $n \in \mathbb{N}^\star$, montrer que $P(\{|X| > n\}) \geq \frac{1}{n\pi}$, où $\{|X| > n\} = \{\omega \in \Omega, |X(\omega)| > n\}$. [On pourra remarquer que $\frac{2}{1+x^2} \geq \frac{1}{x^2}$ pour tout $x \geq 1$.]
3. Pour $x \in \mathbb{R}$, on pose $g(x) = e^{-|x|}$. Montrer que

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

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

   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 $n \in \mathbb{N}^\star$, on pose :

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

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

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

(b) Donner la loi de $Z_n$.

5. Pour $n \in \mathbb{N}^\star$, on pose $A_n = \{|X_n| > n\}$ et $B_n = \cup_{p\geq n} A_p$. on pose aussi $B = \cap_{n\in\mathbb{N}^\star} B_n$.

(a) Soit $n \in \mathbb{N}^\star$. 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 $P(B_n) = 1$.

(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.]

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

(d) Soit $\omega \in \Omega$. Montrer que si la suite $(Z_n(\omega))_{n\in\mathbb{N}^\star}$ converge vers une limite finie on a alors
$$\lim_{n\to\infty} \frac{X_n(\omega)}{n} = 0.$$
[Ecrire $X_n$ en fonction de $Z_n$ et $Z_{n-1}$.]

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

6. Pour tout $n \in \mathbb{N}^\star$, 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 $(Z_n)_{n\in\mathbb{N}^\star}$ ne converge pas en probabilité.

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

**Exercice 10.13 (Sur la loi d'un vecteur aléatoire)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 10.14 (Limite de Gaussiennes)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité, $X$ 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 $\sigma_n$ la variance de $X_n$.

1. Montrer tout d'abord que la suite $(\sigma_n)_{n\in\mathbb{N}}$ est bornée. Puis, montrer que la suite $(\sigma_n)_{n\in\mathbb{N}}$ est convergente dans $\mathbb{R}$. [On pourra utiliser la convergence simple de $\varphi_{X_n}$ vers $\varphi_X$.]
2. Montrer 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 9.5.]
3. Montrer que la suite $(m_n)_{n\in\mathbb{N}}$ est convergente. En déduire que $X$ est une v.a.r. gaussienne.

**Exercice 10.15 (Un exemple...)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $(X, Y)$ un couple de variables aléatoires réelles, à valeurs dans $[-1, 1]^2$. On suppose que la loi de ce couple a une densité $f$ par rapport à la mesure de Lebesgue (sur $\mathcal{B}(\mathbb{R}^2)$) avec :

$$f(x, y) = \frac{1 + xy(x^2 - y^2)}{4} 1_{[-1,1]^2}(x, y), \quad (x, y)^t \in \mathbb{R}^2.$$

1. Montrer que les lois des v.a. $X$ et $Y$ ont des densités par rapport à la mesure de Lebesgue (sur $\mathcal{B}(\mathbb{R})$). Calculer ces densités. $X$ et $Y$ sont-elles indépendantes ?
2. Calculer les espérances de $X$, $Y$ et $XY$. Que vaut $cov(X, Y)$ ?
3. Calculer les fonctions caractéristiques de $X$, $Y$ et $X + Y$.

# Chapitre 11

# Espérance conditionnelle et martingales

## 11.1 Espérance conditionnelle

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

**Définition 11.1** *Soit $(\Omega, \mathcal{A}, P)$ 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 $\mathcal{B}$" l'ensemble des applications $Z$ de $\Omega$ dans $\mathbb{R}$, $\mathcal{B}$-mesurable, intégrable et t.q. :*

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

*On note $E(X|\mathcal{B})$ cette espérance conditionnelle (c'est donc un ensemble de fonctions). (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 montrera dans la proposition 11.1 que, sous les hypothèses de la définition 11.1, l'espérance conditionnelle existe, c'est-à-dire que l'ensemble $E(X|\mathcal{B})$ est non vide, et que $E(X|\mathcal{B})$ est "unique", ceci signifiant que si $Z_1, Z_2 \in E(X|\mathcal{B})$ , on a nécessairement $Z_1 = Z_2$ p.s..
L'ensemble $E(X|\mathcal{B})$, défini dans la définition 11.1, est une ensemble de v.a.r. (car $Z$ $\mathcal{B}$-mesurable implique $Z$ $\mathcal{A}$-mesurable) mais, en pratique, on confond cet ensemble avec l'un de ces éléments (comme on confond un élément de $L^p$ 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. 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 $E(X)$.
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 aussi 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").

Dans les deux exemples précédents, l'ensemble $E(X|\mathcal{B})$ était réduit à un seul élément. Voici maintenant un exemple où $E(X|\mathcal{B})$ n'est pas réduit à un seul élément. 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 $\Omega$). 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 (exercice 11.2) que $E(X|\mathcal{B}) = \{Z_a, a \in \mathbb{R}\}$. L'ensemble $E(X|\mathcal{B})$ n'est donc pas réduit à un élément.

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

**Proposition 11.1** *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 :*

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

*2. (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 partie "existence si $X$ est de carré intégrable" est inutile. Elle 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 11.1 avec $Z = Z_1$ et $Z = Z_2$, $E(XU) = E(Z_1U)$ et $E(XU) = E(Z_2U)$. 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 $\|T\| \leq \|X\|_2$). On peut donc appliquer le théorème de Riesz dans les espaces de Hilbert (théorème 6.9), il 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 (11.1) (et donc que $Z \in E(X|B)$). Soit $U$ une application $\mathcal{B}$-mesurable bornée de $\Omega$ 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 (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 6.11, qui se démontre d'ailleurs avec le théorème de Riesz dans les espaces de Hilbert, théorème 6.9). 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 $B \in \mathcal{B}$, on pose $m(B) = \int_\Omega X1_B dP$. On définit ainsi une mesure finie, $m$, sur $\mathcal{B}$ (la $\sigma$-additivité de $m$ est immédiate). Cette mesure 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 $m(B) = 0$). Le théorème de Radon-Nikodym (théorème 6.11) donne alors l'existence de $Z$, $\mathcal{B}$-mesurable positive, t.q. $m = ZP_{\mathcal{B}}$ (c'est-à-dire que $m$ est la mesure sur $\mathcal{B}$ de densité $Z$ par rapport à $P_{\mathcal{B}}$).

La fonction $Z$ est intégrable car $\int_\Omega ZdP = \int_\Omega ZdP_B = (ZP_B)(\Omega) = m(\Omega) = E(X) < +\infty$. Il reste à montrer que $Z$ vérifie (11.1) (ce qui donnera que $Z \in E(X|B)$). Soit $U$ une application $\mathcal{B}$-mesurable bornée de $\Omega$ dans $\mathbb{R}$. On a $ZU \in \mathcal{L}^1(\Omega, \mathcal{B}, P_{\mathcal{B}})$, et donc (voir la remarque 6.22) $U \in \mathcal{L}^1(\Omega, \mathcal{B}, m)$ et :

$$E(ZU) = \int_\Omega ZUdP = \int_\Omega ZUdP_{\mathcal{B}} = \int_\Omega Udm.$$

Mais, comme $m$ est la restriction à $\mathcal{B}$ de la mesure sur $\mathcal{A}$ de densité $X$ par rapport à $P$ (notée $XP$), on a aussi (toujours par la remarque 6.22) $U \in \mathcal{L}^1(\Omega, \mathcal{A}, XP)$ et :

$$\int_\Omega Udm = \int_\Omega UXdP = E(XU).$$

On a donc, finalement, $E(ZU) = E(XU)$. L'application $Z$ vérifie donc (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 (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 11.1). 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 11.2 et la définition 11.2.

**Proposition 11.2** *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, $\geq 0$ p.s. et t.q. :*

$$E(ZU) = E(XU), \tag{11.3}$$

*pour toute application $U$ de $\Omega$ 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 (11.3). 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 (11.1). Elle vérifie donc (11.3) en ajoutant "$U$ 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 (11.3) (c'est-à-dire avec $U$ $\mathcal{B}$-mesurable positive mais non nécessairement bornée), il suffit d'utiliser le théorème de convergence monotone (théorème 4.2) 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, $\geq 0$ p.s. et vérifie (11.3), il est facile de voir que $Z$ vérifie (11.1). En effet, si $U$ est $\mathcal{B}$-mesurable bornée, on utilise (11.3) 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 11.1. On rappelle que $p_{\mathcal{B}}$ la restriction de $P$ à $\mathcal{B}$. Pour $B \in \mathcal{B}$, on pose $m(B) = \int_\Omega X 1_B dP$. La mesure $m$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$. Le théorème de Radon-Nikodym donne alors l'existence de $Z$, $\mathcal{B}$-mesurable positive, t.q. $m = ZP_{\mathcal{B}}$. Il reste à montrer que $Z$ vérifie (11.3) (ce qui donnera que $Z \in \bar{E}(X|B)$). Soit $U$ une application $\mathcal{B}$-mesurable positive de $\Omega$ dans $\mathbb{R}$. On a $E(ZU) = \int_\Omega ZUdP = \int_\Omega ZUdP_{\mathcal{B}} = \int_\Omega Udm$. Mais, comme $m$ est la restriction à $\mathcal{B}$ de la mesure sur $\mathcal{A}$ de densité $X$ par rapport à $P$ (notée $XP$), on a aussi $\int_\Omega Udm = \int_\Omega UXdP = E(XU)$. On a donc, finalement, $E(ZU) = E(XU)$. L'application $Z$ vérifie donc (11.3). Ce qui prouve que $Z \in \bar{E}(X|\mathcal{B})$.

**Unicité.** Soit $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$. prenons $U = (\mathrm{sign}(Z_1 - Z_2))^+$ (qui est bien $\mathcal{B}$-mesurable et positive). On a donc, par (11.3), $E(Z_1U) = E(Z_2U) = E(XU)$, mais on ne peut rien en déduire car il est possible que $E(XU) = +\infty$. On va donc modifier légérement $U$. Soit $n \in \mathbb{N}^\star$. On pose $B_n = \{Z_1 \leq n\} \cap \{Z_2 \leq n\}$ et $U_n = U1_{B_n}$. La fonction $U_n$ est encore $\mathcal{B}$-mesurable et positive et (11.3) donne $E(Z_1U_n) = E(Z_2U_n)$. Comme $0 \leq E(Z_1U_n) = E(Z_2U_n) \leq n$, on en déduit $E((Z_1 - Z_2)U_n) = 0$. Mais, $(Z_1 - Z_2)U_n \geq 0$. En faisant tendre $n$ vers l'infini, le théorème de convergence monotone (théorème 4.1) donne $E((Z_1 - Z_2)U) = 0$, c'est-à-dire $E((Z_1 - Z_2)^+) = 0$ et donc $Z_1 \leq Z_2$ p.s.. En changeant les rôles de $Z_1$ et $Z_2$ on a aussi $Z_2 \leq Z_1$ p.s.. D'où $Z_1 = Z_2$ p.s.. ■

**Définition 11.2** *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 $\Omega$ dans $\mathbb{R}$, $\mathcal{B}$-mesurable, positive et t.q. :*

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

*pour toute application $U$ de $\Omega$ dans $\mathbb{R}$, $\mathcal{B}$-mesurable et positive.*
*On note $\bar{E}(X|\mathcal{B})$ cette espérance conditionnelle (c'est donc un ensemble de fonctions). (Noter que dans (11.1), les applications $ZU$ et $XU$ sont bien des v.a.r. positives, leur intégrale sur $\Omega$ est donc bien définie et appartient à $\bar{R}_+$).*

La proposition 11.2 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 11.2, 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 11.2 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 $\bar{E}(X|\mathcal{B})$ et on conservera la notation $E(X|\mathcal{B})$ 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.

**Proposition 11.3** *Soit $(\Omega, \mathcal{A}, P)$ 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 $X \in \mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{A}, P)$. Soit $Z \in E(X|\mathcal{B})$. Alors, $Z \in \mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{A}, P)$ et $E(ZU) = E(XU)$ pour toute application $U$ (de $\Omega$ dans $\mathbb{R}$) $\mathcal{B}$-mesurable t.q. $|U|^q$ soit intégrable.*

DÉMONSTRATION : La démonstration fait partie de l'exercice 11.5. En fait, le cas $p = 2$ a déjà été vu dans la démonstration de la proposition 11.1. ∎

**Proposition 11.4 (Inégalité de Jensen généralisée)**
*Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{B}$ une tribu incluse dans $\mathcal{A}$ et $X$ une v.a.r. de carré intégrable. Soit $\varphi$ une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que $\varphi(X)$ est intégrable. On a alors $E(\varphi(X)|\mathcal{B}) \geq \varphi(E(X|\mathcal{B}))$ p.s..*

DÉMONSTRATION : D'après le lemme 11.1, comme $\varphi$ 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 $Z \in E(X|\mathcal{B})$. On a donc pour tout $\omega \in \Omega$ :

$$\varphi(X(\omega)) - \varphi(Z(\omega)) \geq c(Z(\omega))(X(\omega) - Z(\omega)). \tag{11.5}$$

On aimerait intégrer cette inégalité sur un élément (bien choisi) de $\mathcal{B}$ mais cela n'est pas possible car les v.a.r. $\varphi(Z)$ et $c(Z)(X - Z)$ peuvent ne pas être intégrables (bien que $Z$ et $X$ soient intégrables). Pour $p \in \mathbb{N}^\star$, 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 $A = \{E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0\}$ et $B_p = A_p \cap A$. Soit $p \in \mathbb{N}^\star$, l'inégalité (11.5) donne $1_{B_p}(\varphi(X) - \varphi(Z)) \geq 1_{B_p}c(Z)(X - Z)$ et donc, en intégrant sur $\Omega$ :

$$\int_{B_p} (\varphi(X) - \varphi(Z))dP \geq \int_{B_p} c(Z)(X - Z)dP. \tag{11.6}$$

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 :

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

et

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

Avec (11.6), 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 $P(A) = P(\cup_{p \in \mathbb{N}^\star} B_p) = 0$. Ce qui donne bien $E(\varphi(X)|\mathcal{B}) \geq \varphi(Z)$ p.s.. ∎

**Lemme 11.1** *Soit $\varphi$ 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 $\varphi$ est dérivable sur $\mathbb{R}$, la fonction $c$ existe et est unique, elle est donnée par $c = \varphi'$. L'existence de $c$ est légérement plus difficile si $\varphi$ 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 $\varphi$ 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 défnit maintenant l'espérance conditionnelle par rapport à une v.a.r. ou un v.a.

**Définition 11.3** *Soit $(\Omega, \mathcal{A}, P)$ 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 conditionnelle de $Y$ sachant $X$") l'ensemble $E(Y|\sigma(X))$, où $\sigma(X)$ est la tribu engendrée par $X$. On note $E(Y|X)$ cette espérance conditionnelle, de sorte que $E(Y|X) = E(Y|\sigma(X))$. (L'ensemble $E(Y|X)$ est donc un ensemble de v.a.r. et, comme d'habitude, on confond $E(Y|X)$ avec l'un de ces éléments.)*

Pour caractériser $E(Y|X)$ (sous les hypothèses de la définition 11.3) et pour calculer cette espérance conditionnelle, on utilise, en général, le théorème 3.1 que nous rappelons sous une forme légérement plus précise (donnée dans la démonstration du théorème 3.1).

**Théorème 11.1 ($Y$ mesurable par rapport à $X$)** *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $X$, $Y$ deux v.a.r.. On note $\sigma(X)$ la tribu engendrée par $X$. Alors :*
- *La v.a.r. $Y$ est $\sigma(X)$-mesurable si et seulement si il existe $f$, fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Y = f(X)$.*
- *La v.a.r. $Y$ est $\sigma(X)$-mesurable bornée si et seulement si il existe $f$, fonction borélienne bornée de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Y = f(X)$.*
- *La v.a.r. $Y$ est $\sigma(X)$-mesurable positive si et seulement si il existe $f$, fonction borélienne positive de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Y = f(X)$.*

DÉMONSTRATION : La démonstration de ce théorème est donnée dans la démonstration du théorème 3.1. ■

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

**Proposition 11.5 (Calcul de $E(Y|X)$)** *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $X$, $Y$ deux v.a.r.. Soit $Z$ une application de $\Omega$ dans $\mathbb{R}$.*

1. *On suppose que $Y$ est intégrable. Alors, $Z \in E(Y|X)$ si et seulement si il existe $\psi$ application borélienne de $\mathbb{R}$ dans $\mathbb{R}$ t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et*
$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \tag{11.7}$$
*pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne bornée.*
2. *On suppose que $Y$ est positive. Alors, $Z \in E(Y|X)$ si et seulement si il existe $\psi$ application borélienne positive de $\mathbb{R}$ dans $\mathbb{R}$ t.q. $Z = \psi(X)$ et*
$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \tag{11.8}$$
*pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne positive.*

DÉMONSTRATION : La démonstration est une conséquence facile du théorème 11.1. ■

La conséquence de la proposition 11.5 est que (sous les hypotèses de la proposition) l'on cherche $E(Y|X)$ sous la forme d'une fonction $\psi(X)$ (avec $\psi$ de $\mathbb{R}$ dans $\mathbb{R}$) vérifiant (11.7) (ou (11.8)). On raisonne, en général, par "condition nécessaire sur $\psi$" 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 11.5 montre également que (sous les hypotèses de la proposition 11.5) la fonction $Y$ est une fonction de $X$ (c'est-à-dire $Y = \psi(X)$ pour un certain $\psi$ 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 11.13.

## 11.2 Martingales

**Définition 11.4 (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 11.5 (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.).*

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

**Définition 11.6 (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 $\Omega$, muni de la tribu $\mathcal{A}$, dans $\mathbb{N} \cup \{+\infty\}$, muni de la tribu la plus fine). L'application $T$ s'appelle un temps d'arrêt si, pour tout $n \in \mathbb{N}$, $\{T = n\} \in \mathcal{B}_n$.*

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

**Théorème 11.2 (Convergence p.s. d'une martingale)**
*Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $(\mathcal{B}_n)_{n\in\mathbb{N}}$ une filtration et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. positives. On suppose que $(X_n)_{n\in\mathbb{N}}$ est une martingale par rapport à $(\mathcal{B}_n)_{n\in\mathbb{N}}$. Alors, il existe une v.a.r. intégrable, $X$, t.q. $X_n \to X$ p.s., quand $n \to \infty$.*

## 11.3 Exercices

### 11.3.1 Espérance conditionnelle

**Exercice 11.1 (Espérance conditionnelle selon une tribu)** *Corrigé 183 page 483*

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

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

**Exercice 11.2 (Espérance conditionnelle selon une tribu (2))**

Soit $(\Omega, \mathcal{A},\ P)$ un espace probabilisé et $X$ une v.a.r. intégrable. 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 $\Omega$). 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 $E(X|\mathcal{B}) = \{Z_a,\, a \in \mathbb{R}\}$.

**Exercice 11.3 (Espérance conditionnelle selon une v.a.r.)** *Corrigé 184 page 484*

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 $E[Y|X]$.
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 $E[Y|X]$.
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 $E[Y|X]$.

**Exercice 11.4 (Egalité d'espérances conditionnelles)**

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $\mathcal{B}$ une sous-tribu de $\mathcal{A}$, $B \in \mathcal{B}$ un événement et $X$ une v.a.r. intégrable t.q. $X1_B = Y1_B$ p.s.. Montrer que $E[X|\mathcal{B}]1_B = E[Y|\mathcal{B}]1_B$ p.s..

**Exercice 11.5 (Espérance conditionnelle d'une v.a.r. appartenant à $\mathcal{L}^p$)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités, $\mathcal{G}$ 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. $|Y| \leq M$ p.s.. Montrer que $|Z| \leq M$ p.s.. et que$E[Z\Phi] = E[Y\Phi]$ pour tout $\Phi$ $\mathcal{G}-$mesurable et intégrable.
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 $\Phi$ $\mathcal{G}-$mesurable et t.q. $|\Phi|^q$ soit intégrable.

**Exercice 11.6 (Calcul de $E(\exp(XY)|X)$ si $Y \sim \mathcal{N}(0,1)$)** *Corrigé 185 page 486*

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités 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 $E[\exp(XY)|X]$.

**Exercice 11.7 (Espérance selon une somme de v.a.r.i.i.d)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités.

1. Soit $X$ et $Y$ deux v.a.r. intégrables. Montrer que $E(X|X+Y){+}E(Y|X+Y){=}X+Y$. On suppose maintenant que $X$ et $Y$ sont indépendantes et de même loi. Montrer que

$$E(X|X+Y) = E(Y|X+Y) = \frac{X+Y}{2}.$$

2. Soit $n \in \mathbb{N}^\star$ et $X_1, \ldots, X_n$ des v.a.r. indépendantes, de même loi et intégrables. On note $S_n = \sum_{k=1}^n X_k$. Montrer que $E(X_1|S_n) = S_n/n$.

**Exercice 11.8 (Une condition nécessaire et suffisante pour avoir $X = Y$ p.s.)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités et $X$, $Y$ deux v.a.r. intégrables. On suppose que $E[X|Y] = Y$ p.s. et $E[Y|X] = X$ p.s..

1. Soit $c \in \mathbb{R}$. Montrer que

$$E\left[(X-Y)1_{X>c,Y>c}\right] = E\left[(Y-X)1_{X>c\geq Y}\right] \leq 0.$$

   En déduire que $E\left[(X-Y)1_{X>c,Y>c}\right] = 0$, puis que $P[X > c \geq Y] = 0$.
2. Montrer 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 $E((X-Y)^2)$.

   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 11.9 (Espérance du produit et produit des espérances)** *Corrigé 186 page 486*
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités et $X$, $Y$ deux v.a. intégrables t.q. $XY$ est intégrable et $E[X|Y] = E(X)$ p.s.. Montrer que $E(XY) = E(X)E(Y)$.

**Exercice 11.10 (Egalité de lois donne égalité d'espérances conditionnelles)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités et $X, Y, Z$ trois v.a.r.. On suppose que $X$ et $Y$ sont intégrables et que $(X, Z) \sim (Y, Z)$.

1. Montrer que $E[X|Z] = E[Y|Z]$ 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..

**Exercice 11.11 (Convergence faible et espérances conditionnelles)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités, $\mathcal{B}$ 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. $U$ bornée, on a $\lim_{n\to\infty} E[X_n U] = E[XU]$).

1. Montrer que pour toute v.a. $U$, $\mathcal{B}$-mesurable et bornée, on a

$$\lim_{n\to\infty} E[E[X_n|\mathcal{B}]U] = E[E[X|\mathcal{B}]U].$$

2. Montrer que $E[X_n|\mathcal{B}] \to E[X|\mathcal{B}]$ faiblement dans $L^1(\Omega, \mathcal{A}, P)$ quand $n \to \infty$.

**Exercice 11.12 (Minoration d'une espérance conditionnelle)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité, $\mathcal{B}$ 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 $E(YZ) \geq E(UZ)$. Montrer que $E(Y|B) \geq U$ p.s..

**Exercice 11.13 (Dépendance linéaire et dépendance non linéaire)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $X, Y$ deux v.a.r. non constantes.

1. (Dépendance linéaire.) On pose $\overline{X} = \frac{X-E(X)}{\sqrt{\text{Var}(X)}}$ et $\overline{Y} = \frac{Y-E(Y)}{\sqrt{\text{Var}(Y)}}$.
   (a) Montrer que $|\text{Cov}(\overline{X}, \overline{Y})| \leq 1$.
   (b) Montrer que $|\text{Cov}(\overline{X}, \overline{Y})| = 1$ si et seulement si il existe $\alpha, \beta \in \mathbb{R}$, $\alpha \neq 0$, t.q. $Y = \alpha X + \beta$.
   (c) Donner un exemple pour lequel $Y = f(X)$ (avec $f$ fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) et $\text{Cov}(\overline{X}, \overline{Y}) = 0$.
2. (Dépendance.)
   (a) On suppose que $X$ et $Y$ sont indépendantes. Montrer que $E(Y|X) = E(Y)$.
   (b) Montrer qu'il existe $f$ (fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$) t.q. $Y = f(X)$ si et seulement $E(Y|X) = Y$.

**Exercice 11.14 (Lorsque $E(Y|X) = X$ p.s.. . . )** *Corrigé 187 page 487*
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 $\varphi$ une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que la v.a.r. $\varphi(X)$ est de carré intégrable. Montrer que $\int_\Omega Y\varphi(X)dP = \int_\Omega X\varphi(X)dP$.
   (b) Montrer que $E(Y) = E(X)$ et $E(XY) = E(X^2)$.

2.(a) Montrer que $E(X^2) \leq E(Y^2)$.
   (b) Montrer que $Y = X$ p.s. si et seulement si $E(Y^2) = E(X^2)$.

**Exercice 11.15 (V.a. gaussien et espérance conditionnelle)** *Corrigé 188 page 487*
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 $D_{1,1} = \text{Var}(X)$, $D_{2,2} = \text{Var}(Y)$ et $D_{1,2} = D_{2,1} = \text{Cov}(X, Y)$). On suppose que $\text{Var}(Y) > 0$.

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 $E[XY] = E[(\alpha+\beta Y)Y]$.

   Avec $\alpha$ et $\beta$ 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 $\text{Cov}(Z, Y) = 0$. En déduire que $Z$ et $Y$ sont des v.a.r. indépendantes [On pourra utiliser la question 1(b) de l'exercice 10.10].
3. Montrer que $E[X|Y] = l(Y)$ p.s.
4. Calculer (en fonction de $D$) $\text{Var}(Z)$.

   Dans la suite, on note $\sigma = \sqrt{\text{Var}(Z)}$ et, pour $a \in \mathbb{R}$ on note $\mu_a$ 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 $\psi$ est une fonction continue (de $\mathbb{R}$ dans $\mathbb{R}$) si $\sigma > 0$.
   (b) Montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s.

**Exercice 11.16 (Convergence d'une suite d'espérance conditionnelles)**
Soient $(E, T, p)$ un espace probabilisé et $(T_n)_{n\in\mathbb{N}}$ une famille de tribus sur $E$ t.q. $T_n \subset T_{n+1}$, pour tout $n \in \mathbb{N}$, et $\cup_{n\in\mathbb{N}} T_n = T$. Soit $X \in L^2_\mathbb{R}(E, T, p)$ et $E(X|T_n)$ l'espérance conditionnelle de $X$ par rapport à la tribu $T_n$. Nous allons montrer que $E(X|T_n)$ converge vers $X$ dans $L^2_\mathbb{R}(E, T, p)$ lorsque $n \to +\infty$.

1. Montrer qu'il existe $e \in L^2_\mathbb{R}(E, T, p)$ et une sous-suite de la suite $(E(X|T_n)_{n\in\mathbb{N}}$ qui converge faiblement vers $e$ dans $L^2_\mathbb{R}(E, T, p)$.

2. Montrer que $\int XY dp = \int eY dp$, pour tout $Y \in F = \overline{\cup_{n\in\mathbb{N}} L^2_{\mathbb{R}}(E, T_n, p)}$.
3. Montrer que $F = L^2_{\mathbb{R}}(E, T, p)$ et en déduire que $e = X$ p.s..
4. Montrer que $\|E(X|T_n)\|_2 \le \|X\|_2$ et 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 11.17 (Projection sur $L^2(\Omega, \tau(X), P)$ versus projection sur ev($X$))**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités. 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 de probabilités. 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 $a \in \mathbb{R}^\star$ t.q. $X = a$ p.s.). Montrer que $P_{\text{ev}(X)} = P_{L^2(\Omega,\tau(X),P)}$ (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 $\mathcal{A}$, $P_{\text{ev}(X)} \ne P_{L^2(\Omega,\mathcal{B},P)}$ (i.e. ev$(X) \ne L^2(\Omega, \mathcal{B}, P)$).

   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.

### 11.3.2 Martingales

**Exercice 11.18 (Quelques propriétés des martingales)** *Corrigé 189 page 490*

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $(X_n)_{n\in\mathbb{N}}$ 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 $(X_n)_{n\in\mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_n$) Montrer que la suite $(E(X_n))_{n\in\mathbb{N}}$ est croissante.
2. On suppose que $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$). Montrer que $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s. pour tout $m \ge 0$.
3. Soit $\varphi$ 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 $(\mathcal{B}_n)_{n\in\mathbb{N}}$).

**Exercice 11.19 (Séries de Fourier et martingales)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $V_n = ev\{e_p, -n \le p \le n\}$ (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 $n \in \mathbb{N}^\star$, 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 $n \in \mathbb{N}^\star$. 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 11.20 (Quelques questions sur les temps d'arrêt)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $\tau$ défini de $\Omega$ 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 $\tau$ est un temps d'arrêt.
2. Soit $\nu$ et $\tau$ 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 $(\mathcal{B}_n)_{n\in\mathbb{N}}$).
3. Soit $\nu$ et $\tau$ deux temps d'arrêt, par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$, et t.q. $\nu \leq \tau$ 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 note $\mathcal{B}_T = \{A \in \mathcal{A}$ t.q., pour tout $n \in \mathbb{N}$, $A \cap \{T = n\} \in \mathcal{B}_n\}$.]

**Exercice 11.21 (Martingale construite avec les espérances d'une v.a.)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $n \in \mathbb{N}$).

**Exercice 11.22 (Il est temps de s'arrêter)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $(\mathcal{F}_n)_{n\in\mathbb{N}}$) $\tau$ borné,

$$E[|X_\tau|] < \infty.$$

   On suppose dans la suite que pour tout temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$) $\tau$ borné,

$$E[X_\tau] = E[X_0].$$

2. Soit $n \in \mathbb{N}$ et $\Lambda_n \in \mathcal{F}_n$. On définit $\sigma_n$ par

$$\sigma_n(\omega) = n \text{ si } \omega \in \Lambda_n \text{ et } \sigma_n(\omega) = n+1 \text{ si } \omega \notin \Lambda_n.$$

   Montrer que $\sigma_n$ est un temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n\in\mathbb{N}}$).
3. Soit $n \in \mathbb{N}$ et $\nu_n$ défini par $\nu_n(\omega) = n + 1$ pour tout $\omega \in \Omega$. Montrer que $\nu_n$ est aussi 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 $\tau$ et tout événement $A$), montrer que $(X_n)_{n\in\mathbb{N}}$ est une martingale.

# Chapitre 21

# Transformation de Fourier

## 21.1 Transformée de Fourier dans $L^1$

**Corrigé 178 (Résultat partiel d'inversion de Fourier dans $L^1$)**
Soit $H(t) = e^{-|t|}, t \in \mathbb{R}$. On pose, pour $\lambda > 0$ :

$$h_\lambda(x) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} H(\lambda t) e^{itx} dt, x \in \mathbb{R}. \tag{21.1}$$

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

**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 :

$$f \star h_\lambda(x) = \int_{\mathbb{R}} H(\lambda t)\hat{f}(t)e^{ixt}dt. \tag{21.2}$$

**corrigé**

Noter que $\lambda$ désigne ici à la fois le paramètre $\lambda$ 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 $h_\lambda \in L^\infty_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $f \star h_\lambda(x)$ est défini pour tout $x \in \mathbb{R}$ et on a :

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

Comme $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $H(\lambda \cdot) \in L^\infty_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, on peut utiliser le thérorème de Fubini (théorème 7.3) pour inverser l'ordre d'intégration et obtenir :

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

---

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

**corrigé**

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

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

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

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

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) :

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

---

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

$$\|f \star h_\lambda - \sqrt{2\pi} f\|_1 \to 0 \text{ lorsque } \lambda \to 0. \tag{21.3}$$

[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 :

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

En utilisant le théorème de Fubini-Tonelli (théorème 7.2), on en déduit :

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

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

Le théorème de continuité en moyenne dans $L^1$ (théorème 5.6, 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 $\mathbb{C}$. Ellle 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 $\lim_{\lambda \to 0} g \star h_\lambda(0) = 0$ et donc que $\lim_{\lambda \to 0} \|f \star h_\lambda - \sqrt{2\pi} f\|_1 = 0$.

---

5. Déduire de ce qui précède le théorème d'inversion de Fourier, théorème 10.1.

corrigé

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

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

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

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

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

---

## 21.2 Transformée de Fourier d'une mesure signée

**Corrigé 179 (Caractérisation de $m$ par $\hat{m}$)**

Soit $d \geq 1$.

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

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

**corrigé**

On remarque que $\int \hat{\varphi} dm = (2\pi)^{-\frac{d}{2}} \int \left( \int e^{-ix\cdot t} \varphi(x) dx \right) dm(t)$. (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 2.6). 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 7.2) avec les mesures $\lambda_d$ et $m^+$ et les mesures $\lambda_d$ et $m^-$. On obtient ainsi :

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

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

---

(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 $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).

**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 $\int \hat{\varphi} dm = \int \hat{\varphi} d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$. Or, l'application $\varphi \mapsto \hat{\varphi}$ est une bijection dans $\mathcal{S}(\mathbb{R}^d, \mathbb{C})$ (proposition 10.6). 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 $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).

**corrigé**

Soit $\varphi \in C_c(\mathbb{R}^d, \mathbb{R})$ et $(\varphi_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^d, \mathbb{R})$ 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 5.4 donne alors $m = \mu$.

---

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

## 21.3 Fonction caractéristique d'un v.a.

**Corrigé 180 (V.a. gaussiens, indépendance, covariance)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités, $d \geq 1$ et $X = (X_1, \ldots, X_d)$ 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 $\mathrm{Cov}(X_1, X_2) = 0$.

**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. $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$. On a donc $\varphi_{X_k}(u) = e^{ium_k} e^{-\frac{\sigma_k^2 u^2}{2}}$ 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. $a_1X_1 + a_2X_2$. Soit $u \in \mathbb{R}$, on a :

$$\varphi_{a_1X_1+a_2X_2}(u) = \int_\Omega e^{i(a_1X_1+a_2X_2)u}dP = \int_\Omega e^{ia_1X_1u}e^{ia_2X_2u}dP$$

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

$$\varphi_{a_1X_1+a_2X_2}(u) = \int_\Omega e^{ia_1X_1u}dP \int_\Omega e^{ia_2X_2u}dP = \varphi_{X_1}(a_1u)\varphi_{X_2}(a_2u).$$

Ce qui donne $\varphi_{a_1X_1+a_2X_2}(u) = e^{iu(a_1m_1+a_2m_2)}e^{-\frac{u^2(\sigma_1^2a_1^2+\sigma_2^2a_2^2)}{2}}$. Comme la loi d'une v.a.r. est entièrement déterminée par sa fonction caractéristique (voir la proposition 10.10), on en déduit que $a_1X_1 + a_2X_2 \sim \mathcal{N}(m, \sigma^2)$ avec $m = a_1m_1 + a_2m_2$ et $\sigma = \sqrt{\sigma_1^2a_1^2 + \sigma_2^2a_2^2}$. Ceci prouve bien que $X$ est un vecteur gaussien.

L'indépendance de $X_1$ et $X_2$ permet aussi de calculer la fonction caractéristique de $X$. Soit $u = (u_1, u_2)^t \in \mathbb{R}^2$ :

$$\varphi_X(u) = \int_\Omega e^{i(X_1u_1+X_2u_2)}dP = e^{i(u_1m_1+u_2m_2)}e^{-\frac{u_1^2\sigma_1^2+u_2^2\sigma_2^2}{2}}.$$

Ceci prouve que la matrice de covariance de $X$ est diagonale (proposition 10.13) et donc que

$$\mathrm{Cov}(X_1, X_2) = 0.$$

---

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

**corrigé**

D'après la proposition 10.11, pour montrer que $X_1$ et $X_2$ sont indépendantes, il suffit de montrer que $\varphi_X(u) = \prod_{j=1}^2 \varphi_{X_j}(u_j)$ pour tout $u = (u_1, u_2)^t \in \mathbb{R}^2$. Ceci est une conséquence facile du fait que $\mathrm{Cov}(X_1, X_2) = 0$. En effet, la matrice de covariance de $X$ est alors diagonale et on a bien (grâ ce à la proposition 10.13), en reprenant les notations de la question précédente (c'est-à-dire $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$),

$$\varphi_X(u) = e^{i(u_1m_1+u_2m_2)}e^{-\frac{u_1^2\sigma_1^2+u_2^2\sigma_2^2}{2}} = \varphi_{X_1}(u_1)\varphi_{X_2}(u_2),$$

c'est-à-dire $\varphi_X(u) = \prod_{j=1}^2 \varphi_{X_j}(u_j)$ 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 $(\Omega, \mathcal{A}, P)$ et $X_1$, $X_2$ de manière à avoir $\mathrm{Cov}(X_1, X_2) = 0$ sans que $X_1$, $X_2$ soient indépendantes, voir l'exercice 4.44.]

**corrigé**

On considère les deux v.a.r. $S$ et $X$ de l'exercice 4.44 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 $\mathrm{Cov}(X_1, X_2) = 0$ (voir l'exercice 4.44). 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 10.11, 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 10.13), 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 $u = (u_1, \ldots, u_d)^t \in \mathbb{R}^d$.

---

**Corrigé 181 (Suite de v.a.r.i.i.d. de Poisson)**

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $X$ une v.a. de Poisson de paramètre $\lambda$ ($\lambda > 0$). 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 $\varphi_X$ 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 $(X_n)_{n\in\mathbb{N}^\star}$ une suite de v.a. indépendantes et de Poisson de paramètre $\lambda$.

(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 10.10), 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 $(X_n)_{n\in\mathbb{N}^\star}$ 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 $n \in \mathbb{N}^\star$, 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 6.13) donne que la suite $(P_{Z_n})_{n\in\mathbb{N}^\star}$ 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 6.58 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 $n \to \infty$.

---

**Corrigé 182 (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 $\lambda$ désigne la mesure de Lebesgue sur les boréliens de $\mathbb{R}$).

1. Montrer que $X$ n'est pas une v.a.r. intégrable.
2. Soit $n \in \mathbb{N}^\star$, 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 $x \ge 1$.]
3. Pour $x \in \mathbb{R}$, on pose $g(x) = e^{-|x|}$. Montrer que

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

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

   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 $n \in \mathbb{N}^\star$, on pose :

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

4. Soit $n \in \mathbb{N}^\star$.
   (a) Montrer que
$$\varphi_{Z_n}(u) = \varphi^n_{X_1}(\frac{u}{n}) \text{ pour tout } u \in \mathbb{R}.$$
   (b) Donner la loi de $Z_n$.
5. Pour $n \in \mathbb{N}^\star$, on pose $A_n = \{|X_n| > n\}$ et $B_n = \cup_{p\ge n} A_p$. on pose aussi $B = \cap_{n\in\mathbb{N}^\star} B_n$.
   (a) Soit $n \in \mathbb{N}^\star$. En utilisant la question 2, montrer que
$$P(B_n^c) = \prod_{p\ge n} P(A_p^c) \le \prod_{p\ge n}(1 - \frac{1}{p\pi}).$$
   En déduire que $P(B_n^c) = 0$ et donc que $P(B_n) = 1$.
   (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.]

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

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

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

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

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

6. Pour tout $n \in \mathbb{N}^\star$, 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 $(Z_n)_{n\in\mathbb{N}^\star}$ ne converge pas en probabilité.

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

**corrigé**

En attente

# Chapitre 22

# Espérance conditionnelle et martingales

## 22.1 Espérance conditionnelle

**Corrigé 183 (Espérance conditionnelle selon une tribu)**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $Y$ une variable aléatoire réelle intégrable. Dans les trois cas suivants, montrer que $E[Y|\mathcal{B}]$ est réduit à un élément et déterminer $E[Y|\mathcal{B}]$ (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 $\Omega$ dans $\mathbb{R}$, $\mathcal{B}$-mesurable. Soit $a \in \mathrm{Im}(Z)$ (on suppose, bien sûr, $\Omega \neq \emptyset$). On a alors $\{Z = a\} = \{\omega \in \Omega;\, Z(\omega) = a\} \neq \emptyset$. Comme $Z$ est $\mathcal{B}$-mesurable, on a donc $\{Z = a\} = \Omega$. Une application $\mathcal{B}$-mesurable est donc une fonction constante (réciproquement, une fonction constante est bien $\mathcal{B}$-mesurable). Si $Z \in E(Y|\mathcal{B})$, il existe donc $a \in \mathbb{R}$ t.q. $Z(\omega) = a$ pour tout $\omega \in \Omega$. Le réel $a$ doit alors vérifier $E(aU) = E(UY)$ pour tout application $U$, $\mathcal{B}$-mesurable de $\Omega$ dans $\mathbb{R}$. On a donc $ab = E(ab) = E(bY) = bE(Y)$ pour tout $b \in \mathbb{R}$. La seule solution est donc $a = E(Y)$. L'ensemble $E[Y|\mathcal{B}]$ est donc réduit à un seul élément, la fonction constante et égale à $E(Y)$.

————————————————————————————————

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

———————————————corrigé———————————————

Soit $Z$ une application de $\Omega$ dans $\mathbb{R}$, $\mathcal{B}$-mesurable. Les paties $B$ et $B^c$ sont non vides (car de probabilité strictement positive). Soit $\omega_1 \in B$ et $a = Z(\omega_1)$. On a alors $\{Z = a\} = \{\omega \in \Omega;\, Z(\omega) = a\} \neq \emptyset$. Comme $Z$ est $\mathcal{B}$-mesurable, on a donc $\{Z = a\} = B$ ou $\Omega$ et donc $\{Z = a\} \supset B$. De même, soit $\omega_2 \in B^c$ et $b = Z(\omega_2)$, on a $\{Z = b\} \supset B^c$. 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 $\Omega$ 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 $(B_n)_{n\in\mathbb{N}^\star} \subset \mathcal{A}$ t.q. $B_n \cap B_m = \emptyset$ si $n \neq m$, $\Omega = \cup_{n\in\mathbb{N}^\star} B_n$ et $0 < P(B_n) < 1$ pour tout $n \in \mathbb{N}^\star$. On prend pour $\mathcal{B}$ la tribu engendrée par $(B_n)_{n\in\mathbb{N}^\star}$ (c'est-à-dire $\mathcal{B} = \{\cup_{n\in J} B_n,\ J \subset \mathbb{N}^\star\}$).

**corrigé**

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

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

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

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

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

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

---

**Corrigé 184 (Espérance conditionnelle selon 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 11.5. Soit $Z \in E[Y|X]$, il existe alors $\psi$, fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application $\varphi$ 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 $\varphi$ 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 $\psi$ 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 $A_1 = \{X = x_1\}$ 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 11.5.

Soit $Z \in E[Y|X]$, il existe alors $\psi$, fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \tag{22.1}$$

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 $\varphi$ 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 (22.1), 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

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

Ici encore, la fonction $\psi(X)$ est un élément de $E[Y|X]$ dès que $\psi$ 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 $n \in \mathbb{N}^\star$, on pose $A_n = \{X = x_n\}$. Les ensembles $A_n$ sont disjoints deux à deux, $P(A_n) > 0$ pour tout $n \in \mathbb{N}^\star$ et $\sum_{n=1}^{\infty} P(A_n) = 1$.

On utilise encore la proposition 11.5. 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 $\psi$, fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \tag{22.2}$$

On a donc $Z = \sum_{n \in \mathbb{N}^\star} \psi(x_n)1_{A_n}$ p.s.. Soit $(\alpha_n)_{n \in \mathbb{N}^\star} \subset \mathbb{R}$ une suite bornée de $\mathbb{R}$. Dans l'égalité (22.2), on prend pour $\varphi$ une fonction borélienne bornée de $\mathbb{R}$ dans $\mathbb{R}$ t.q. $\varphi(x_n) = \alpha_n$ pour tout $n \in \mathbb{N}^\star$ (un tel $\varphi$ existe, on peut prendre, par exemple, $\varphi(x) = 0$ si $x$ est différent de tous les $x_n$), on obtient :

$$\sum_{n \in \mathbb{N}^\star} \psi(x_n)\alpha_n P(A_n) = \sum_{n \in \mathbb{N}^\star} \alpha_n \int_{A_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n \in \mathbb{N}^\star} \subset \mathbb{R}$. Comme $P(A_n) > 0$ pour tout $n \in \mathbb{N}^\star$, on en déduit que $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$, pour tout $n \in \mathbb{N}^\star$, et donc que

$$Z = \sum_{n \in \mathbb{N}^\star} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n} \text{ p.s..}$$

Enfin, ici encore, la fonction $\psi(X)$ est un élément de $E[Y|X]$ dès que $\psi$ 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 $n \in \mathbb{N}^\star$. Un exemple possible est donc $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$ pour tout $n \in \mathbb{N}^\star$ et $\psi(x) = 0$ pour $x \notin \{x_n, n \in \mathbb{N}^\star\}$. Cet exemple donne la fonction $\sum_{n \in \mathbb{N}^\star} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n}$.

---

**Corrigé 185 (Calcul de $E(\exp(XY)|X)$ si $Y \sim \mathcal{N}(0,1)$)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités 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 9.2, 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 11.5 on cherche un élément de $E[e^{XY}|X]$ sous la forme $\psi(X)$ où $\psi$ est une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(e^{XY}\varphi(X)).$$

Soit $\varphi$ 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..

---

**Corrigé 186 (Espérance du produit et produit des espérances)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilités et $X$, $Y$ deux v.a. intégrables t.q. $XY$ est intégrable et $E[X|Y] = E(X)$ p.s.. Montrer que $E(XY) = E(X)E(Y)$.

**corrigé**

Grâce à la proposition 11.5 on a, pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(E[X|Y]\varphi(Y)) = E(X\varphi(Y))$$

Comme $E[X|Y] = E(X)$ p.s., on en déduit, pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$,borélienne bornée,

$$E(X)E(\varphi(Y)) = E(X\varphi(Y)). \tag{22.3}$$

Soit $n \in \mathbb{N}^\star$. Pour $s \in \mathbb{R}$, on pose $T_n(s) = \max\{-n, \min\{s, n\}\}$. La fonction $T_n$ est borélienne (car continue) bornée (par $n$) de $\mathbb{R}$ dans $\mathbb{R}$. On peut donc utiliser (22.3) avec $\varphi = T_n$. On obtient $E(X)E(T_n(Y)) = E(XT_n(Y))$.

Comme $Y$ est intégrable, on a, par convergence dominée, $\lim_{n\to\infty} E(T_n(Y)) = E(Y)$ (noter que $|T_n(Y)| \leq |Y|$).

Comme $XY$ est intégrable (et c'est uniquement ici que cette hypothèse est utilisée), on a, par convergence dominée,

$$\lim_{n\to\infty} E(XT_n(Y)) = E(XY)$$

(noter que $|XT_n(Y)| \leq |XY|$).

En passant à limite quand $n \to \infty$ sur l'égalité $E(X)E(T_n(Y)) = E(XT_n(Y))$, on a donc $E(X)E(Y) = E(XY)$.

---

**Corrigé 187 (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 $\varphi$ une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. On suppose que la v.a.r. $\varphi(X)$ est de carré intégrable. Montrer que $\int_\Omega Y\varphi(X)dP = \int_\Omega X\varphi(X)dP$.

(b) Montrer que $E(Y) = E(X)$ et $E(XY) = E(X^2)$.

2.(a) Montrer que $E(X^2) \leq E(Y^2)$.

(b) Montrer que $Y = X$ p.s. si et seulement si $E(Y^2) = E(X^2)$.

**corrigé**

En attente

---

**Corrigé 188 (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 $D_{1,1} = \text{Var}(X)$, $D_{2,2} = \text{Var}(Y)$ et $D_{1,2} = D_{2,1} = \text{Cov}(X, Y)$). On suppose que $\text{Var}(Y) > 0$.

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 $E[X] = a$, $E[Y] = b$, $E[Y^2] = \text{Var}(Y) + E[Y]^2 = D_{2,2} + b^2$ et

$$E[XY] = E[(X - E(X))(Y - E(Y))] + E[X]E[Y] = D_{1,2} + ab.$$

On cherche donc $\alpha$ et $\beta$ 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 $\alpha$ et $\beta$) 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 $\alpha$ et $\beta$ 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 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 10.10].

**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 $\alpha$ et $\beta$). Puis,

$$\mathrm{Cov}(Z,Y) = E[ZY] = E[XY] - E[(\alpha + \beta Y)Y].$$

La seconde relation satisfaite par $\alpha$ et $\beta$ 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 10.10 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 11.5, de montrer que $E(l(Y)\varphi(Y)) = E(X\varphi(Y))$ pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne bornée.

Soit donc $\varphi$ 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

$$\begin{aligned}\mathrm{Var}(Z) = E[Z^2] = E[XZ] &= E[X^2] - E[Xl(Y)] \\ &= E[X^2] - \alpha E[X] - \beta E[XY].\end{aligned}$$

Il suffit maintenant d'utiliser les valeurs de $\alpha$ et $\beta$ et le fait que

$$E[X^2] = D_{1,1} + a^2,\ E[X] = a,\ E[X,Y] = D_{1,2} + ab.$$

On obtient

$$\mathrm{Var}(Z) = D_{1,1} - \frac{D_{1,2}^2}{D_{2,2}}$$

---

Dans la suite, on note $\sigma = \sqrt{\mathrm{Var}(Z)}$ et, pour $a \in \mathbb{R}$ on note $\mu_a$ 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 $\psi$ est une fonction continue (de $\mathbb{R}$ dans $\mathbb{R}$) si $\sigma > 0$.

**corrigé**

On utilise ici le théorème 4.9. 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 $\psi$ est continue, il suffit (d'après le théorème 4.9) 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 4.9 donne alors que $\psi$ est continue sur $]-M, M[$. Comme $M$ est arbitraire, on en déduit que $\psi$ est continue sur $\mathbb{R}$.

(b) Montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s.

**corrigé**

Si $\sigma > 0$, la fonction $\psi$ est continue, elle est donc borélienne. On remarque aussi que $\psi$ 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 $\psi$ est donc aussi borélienne bornée.

Comme $\psi$ est borélienne, $\psi(l(Y))$ est donc une v.a.r.. Comme $\psi$ 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 11.5), de montrer que $E[\psi(l(Y))\varphi(Y)] = E[f(X)\varphi(Y)]$ pour toute application $\varphi$ de $\mathbb{R}$ dans $\mathbb{R}$, borélienne bornée.

Soit donc $\varphi$ 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,

$$E[f(X)\varphi(Y)] = \int_{\mathbb{R}} \psi(l(y))\varphi(y) dP_Y(y) = E[\psi(l(Y))\varphi(Y)].$$

Ce qui prouve bien $E[f(X)|Y] = \psi(l(Y))$ p.s..

## 22.2 Martingales

**Corrigé 189 (Quelques propriétés des martingales)**
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité 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 $(X_n)_{n\in\mathbb{N}}$ 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 $(X_n)_{n\in\mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). Montrer que la suite $(E(X_n))_{n\in\mathbb{N}}$ est croissante.

———————————————**corrigé**———————————————

Soit $n \in \mathbb{N}$. Comme $(X_n)_{n\in\mathbb{N}}$ est une sous-martingale, on a $E(X_{n+1}|\mathcal{B}_n) \geq X_n$ p.s. et donc

$$E(E(X_{n+1}|\mathcal{B}_n)) \geq E(X_n).$$

Or (comme les fonctions constantes sont $\mathcal{B}_n$-mesurables bornées),

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

On a donc $E(X_{n+1}) \geq E(X_n)$.

———————————————————————————————————

2. On suppose que $(X_n)_{n\in\mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n\in\mathbb{N}}$). Montrer que $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s. pour tout $m \geq 0$.

———————————————**corrigé**———————————————

Pour $m = 0$, le fait que $E(X_n|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$, découle du fait que $X_n$ est $\mathcal{B}_n$-mesurable. On montre maintenant la propriétée demandée par récurrence sur $m \in \mathbb{N}^\star$.

Pour $m = 1$ le fait que $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$, est donné dans la définition de martingale.

Soit $m \in \mathbb{N}^\star$. On suppose que $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$. On veut montrer que $E(X_{n+m+1}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$. Soit $n \in \mathbb{N}$. Comme $(X_n)_{n\in\mathbb{N}}$ est une martingale, on a $E(X_{n+m+1}|\mathcal{B}_{n+m}) = X_{n+m}$ p.s. et donc :

$$E(X_{n+m+1}U) = E(X_{n+m}U) \text{ pour tout } U \ \mathcal{B}_{n+m}\text{-mesurable bornée.}$$

Comme $\mathcal{B}_n \subset \mathcal{B}_{n+m}$, on a donc aussi

$$E(X_{n+m+1}U) = E(X_{n+m}U) \text{ pour tout } U \ \mathcal{B}_{n}\text{-mesurable bornée.}$$

L'hypothèse de récurrence donne $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s.. On a donc :

$$E(X_{n+m}U) = E(X_{n}U) \text{ pour tout } U \ \mathcal{B}_{n}\text{-mesurable bornée.}$$

On a en déduit :

$$E(X_{n+m+1}U) = E(X_{n}U) \text{ pour tout } U \ \mathcal{B}_{n}\text{-mesurable bornée.}$$

Ce qui montre que $E(X_{n+m+1}|\mathcal{B}_n) = X_n$ p.s. et termine la récurrence.

———————————————————————————————————

3. Soit $\varphi$ 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 $\varphi$ 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 11.4 sur l'inégalité de jensen. Soit $n \in \mathbb{N}$. La proposition 11.4 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.

---

That's all folks