ESPACES DE HILBERT. PROBABILITE

J. Mellac

Juillet 1998

Table des matières

Première partie

ESPACES DE HILBERT

Chapitre 1

ESPACES METRIQUES

1.1 Définitions

Un espace métrique est un couple (E, d) où E est un ensemble et d une application de $E \times E$ dans $I\!R^+$ vérifiant :
- $d(x,y) = 0 \Leftrightarrow x = y$.
- $\forall x, y \in E,\ d(x,y) = d(y,x)$.
- $\forall x, y, z \in E,\ d(x,z) \leq d(x,y) + d(y,z)$. (Inégalité triangulaire).

Si $x_0 \in E$ et $r \in I\!R^{+*}$ on appelle boule ouverte de centre x_0 et de rayon r, l'ensemble noté $B(x_0, r) = \{x \in E | d(x_0, x) < r\}$.

On appelle ouvert de E toute réunion de boules ouvertes. Un fermé de E est une partie de E dont le complémentaire est un ouvert.
Remarque : Une partie A de E est ouverte si et seulement si tout $x_0 \in A$ est centre d'une boule ouverte incluse dans A.

1.2 Exemples d'espaces métriques

Exemple 1 :
$E = I\!R,\quad d(x,y) = |x - y|$. Une boule ouverte est dans ce cas un intervalle ouvert. Les trois propriétés d'une distance son aisément vérifiées.

Exemple 2 :
$E = I\!R^2,\quad d((x_1,y_1),(x_2,y_2)) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$ Pour démontrer l'inégalité triangulaire il suffit d'utiliser l'inégalité
$a_1b_1 + a_2b_2 \leq \sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}$ avec
$a_i = x_i - y_i,\ \ b_i = y_i - z_i,\ \ i = 1, 2$.
Les boules ouvertes de cet espace sont les disques ouverts. On peut définir sur cet ensemble d'autres métriques.
$d_1(x,y) = \max(|x_1 - y_1|, |x_2 - y_2|)$ dont les boules ouvertes sont formés de l'intérieur des carrés de cotés parallèles aux axes.
$d_2(x,y) = |x_1 - y_1| + |x_2 - y_2|$ dont les boules ouvertes sont formés de l'intérieur des carrés dont les diagonales sont parallèles aux axes.

1.3 Propriétés.

Théoreme 1.1 *Un espace métrique E est séparé, ce qui signifie que si x et y sont des points distincts de E, il existe des boules ouvertes contenant x et y respectivement et d'intersection vide.*

Démonstration
La distance $r = d(x, y) > 0$ car $x \neq y$. Les boules $B_x = B(x, r/4)$ et $B_y = B(y, r/4)$ sont telles que $B_x \cap B_y = \emptyset$ et elles contiennent respectivement x et y.

Définition 1.1 *Un point $\lambda \in E$ est un point d'accumulation d'une partie $A \subset E$ si toute boule ouverte de centre λ contient au moins un point de A différent de λ.*

Exemple

Si $E = I\!R$ et si $d(x, y) = |x - y|$, 0 est point d'accumulation de $A = \{1, 1/2, 1/3, \cdots\}$ mais aussi de $A' = A \cup \{0\}$. Ce qui montre qu'un point d'accumulation peut appartenir ou non à l'ensemble.

Théoreme 1.2 *Un ensemble $A \subset E$ est fermé si et seulement si il contient tous ses points d'accumulation.*

Démonstration
Dans le cas où $x \in E - A$qui est un ouvert, il existe une boule $B(x, \epsilon)$ incluse dans $E - A$, x ne peut donc être un point d'accumulation de A.
Réciproquement, si A contient tous ses points d'accumulation et si $x \in E - A$, il existe une boule ouverte de centre x ne contenant ancun point de A. $E - A$ est donc ouvert et A est fermé.

Définition 1.2 *On appelle fermeture ou adhérence de $A \subset E$ l'ensemble $\overline{A} = A \cup A_a$ où A_a est l'ensemble des points d'accumulation de A. C'est donc le plus petit fermé contenant A.*

Remarque : A est fermé $\Leftrightarrow A = \overline{A}$

1.4 Espaces métriques complets

Définition 1.3
- *Une suite (x_n) de points de l'espace métrique converge vers $x \in E$ si*
$$\lim_{n\to\infty} d(x_n, x) = 0.$$
- *Une suite (x_n) est de Cauchy si*
$$lim_{n,m\to\infty} d(x_n, x_m) = 0.$$
- *un espace métrique (E, d) est complet si toute suite de Cauchy de E converge.*

Remarques
1) Une partie $A \subset E$ est fermée si et seulement si elle contient les limites de ses suites convergentes.
2) $(I\!R, |.|)$ est complet mais $Q\!\!\!\!Q$ muni de la même distance ne l'est pas. La première affirmation provient de la construction même de $I\!R$. Pour montrer la deuxième notons x_n la représentation décimale par défaut de $\sqrt{2}$ contenant n décimales, c'est une suite de $Q\!\!\!\!Q$, mais la limite $\sqrt{2} \notin Q\!\!\!\!Q$ (c'est une suite de Cauchy convergeant dans $I\!R$ mais pas dans $Q\!\!\!\!Q$).

1.5 Espaces vectoriels normés

Dans la suite le corps de base des espaces vectoriel considérés est noté K. Il s'agira de $I\!R$ ou $C\!\!\!\!C$.

Définition 1.4 *Un espace vectoriel E est normé s'il est muni d'une application*

$$||.|| : E \to I\!R^+$$

vérifiant

— $||x|| = 0 \Leftrightarrow x = 0$.
— $\forall \lambda \in K$, $\forall x \in E$ $||\lambda x|| = |\lambda| \, ||x||$.
— $\forall x, y \ \in E$ $||x + y|| \leq ||x|| + ||y||$.

Il en découle immédiatement que $d(x, y) = ||x - y||$ définit une distance sur E.

Exemples
a) $E = I\!R^n$, $||x|| = \sqrt{\sum_{i=1}^n x_i^2}$ est un $I\!R$-espace vectoriel.
b) $E = C([a, b])$ (fonctions réelles ou complexes continues sur l'intervalle $[a, b]$) $||f|| = \max_{a \leq x \leq b} |f(x)|$.
Une suite (f_n) converge alors vers f pour la distance d associée à $||.||$ si et seulement si (f_n) converge uniformément vers f.
c) $E = C\!\!\!\!C^n$, $||x|| = \sqrt{\sum_{i=1}^n |x_i|^2}$ est un $C\!\!\!\!C$-espace vectoriel normé.
d) $E = \left\{ (x_n)_{n \in I\!N} \mid \sum_{n \in I\!N} |x_n|^2 < +\infty \right\}$ où $x_n \in K$ pour tout n est un K-espace vectoriel normé pour la norme

$$N_2(x) = (\sum_{n \in I\!N} |x_n|^2)^{1/2}.$$

Remarque.
En notant μ la mesure de décompte sur $I\!N$ et $x = (x_n)$ on a

$$N_2(x) = (\int_{I\!N} |x|^2 d\mu)^{1/2}$$

cet espace est donc un espace de Riesz particulier étudié dans le cours d'intégration, il est noté $l^2(K)$.

1.6 Espaces vectoriels normés complets

Définition 1.5 *On appelle espace de Banach tout espace vectoriel normé complet.*

Exemples.
a) $\mathbb{R}^n$ et $\mathbb{C}^n$ sont complets pour la norme précisée précédemment. Ceci provient du fait que $\mathbb{R}$ et $\mathbb{C}$ sont complets et qu'une suite de terme général $x_k = (x_k^1, x_k^2, \cdots, x_k^n)$ converge vers $x = (x_1, x_2, \cdots, x_n)$ dans K^n si et seulement si pour tout i x_k^i converge vers x_i.
b) L'espace $l^2(K)$ est complet. Ceci a été démontré dans le cours d'intégration (tous les espaces L^p sont complets). Cet exemple est très important car cet espace est isomorphe à tous les espaces de Hilbert que l'on traitera dans ce cours.
c) L'espace $(X, \mathcal{B}, \mu)$ étant un espace mesuré ($\mathcal{B}$ est une tribu sur X, μ une mesure sur cet espace) on définit

$$L^p_\mu(X) = \left\{ f : X \to K \mid \int_X |f|^p d\mu < +\infty \right\}.$$

Tous ces espaces correspondants à $p \in \mathbb{R}$, $p \geq 1$ sont des espaces de Banach (voir cours d'intégration). L'exemple précédent est un cas particulier où $X = \mathbb{N}$.

1.7 Applications linéaires continues

Rappel.

— Les espaces (E, d) et (F, δ) étant métriques et f une application de E vers F, on dit que f est continue en $a \in E$ si

$$\forall \epsilon > 0 \ , \ \exists \eta > 0 \ , \ \forall x \in E \ , \ d(x, a) < \eta \Rightarrow \delta(f(x), f(a)) < \epsilon.$$

f est continue sur E si elle est continue en tout point de E. Il revient au même de dire que l'image réciproque de tout ouvert est un ouvert (ou l'image réciproque de tout fermé est un fermé).

— On dit que f est lipschitzienne s'il existe $k \in \mathbb{R}^+$ tel que

$$\forall x, y \in E \ , \ \delta(f(x), f(y)) \leq k d(x, y).$$

Une application lipschitzienne est continue sur E.

— Une application f est continue en $a \in A$ si et seulement si pour toute suite (x_n) convergeant vers a dans (E, d), la suite $(f(x_n))$ converge vers $f(a)$ dans (F, δ).
Dans la suite E et F seron deux espaces vectoriels normés, munis des distances associées à ces normes, qui seront toutes deux notées $||.||$.

Théoreme 1.3 *Soit u une application linéaire de E dans F, les trois propriétés suivantes sont équivalentes*

1) u est continue en o.
2) u est bornée sur toute partie bornée de E.

3) u est lipschitzienne.

Démonstration
(1) $\Rightarrow$ (2) : Si u est continue en 0, il existe $r > 0$ tel que $||x|| \leq r \Rightarrow ||u(x)|| \leq 1$. u est donc bornée sur la boule $B(0, r)$ donc par homothétie sur toute boule de centre 0 et donc sur toute partie bornée de E.
(2) $\Rightarrow$ (3) : Si (2) est vérifiée on a $||x|| = 1 \Rightarrow ||u(x)|| \leq M$.
Si $x \neq 0$, $||u(x/||x||)|| \leq M \Rightarrow ||u(x)|| \leq M||x||$ et ceci est vérifié si $x = 0$. On a alors $||u(x) - u(y)|| = ||u(x - y)|| \leq M||x - y||$ et u est bien lipschitzienne.
(3) $\Rightarrow$ (1) est immédiat.

Remarque. Soit u une application linéaire de E vers F, la plus petite constante de Lipschtiz associée à E est le nombre

$$||u|| = \sup_{x=0} \frac{||u(x)||}{||x||} = \sup_{||x||\leq 1} ||u(x)|| = \sup_{||x||=1} ||u(x)||.$$

Ce nombre est appelé norme de u. On a donc
$||u(x)|| \leq ||u||\,||x||$, le signe $||.||$ ayant trois significations différentes dans cette inégalité.

Théoreme 1.4 *L'espace vectoriel $\mathcal{L}(E, F)$ des applications linéaires continues de E dans F est normé par $u \to ||u||$. Si F est complet $\mathcal{L}(E, F)$ est complet pour la norme qui vient d'être définie.*

Démonstration Soient $u, v \in \mathcal{L}(E, F)$

$$||u + v|| = \sup_{||x||=1} ||u(x) + v(x)|| \leq \sup_{||x||=1} ||u(x)|| + \sup_{||x||=1} ||v(x)|| = ||u|| + ||v||$$

Pour $\lambda \in K$, on a immédiatement $||\lambda u|| = |\lambda|\,||u||$
$||u|| = 0 \Rightarrow \forall x \in E \;\; u(x) = 0$ c'est à dire $u = 0$.
L'espace vectoriel $\mathcal{L}(E, F)$ est donc bien normé à l'aide de l'application définie plus haut.
Supposons F complet et considérons une suite de Cauchy (u_n) dans $\mathcal{L}(E, F)$.
L'inégalité (1) $||u_p(x) - u_q(x)|| \leq ||u_p - u_q||\,||x||$ entraine que $(u_n(x))$ est une suite de cauchy dans F pour tout $x \in E$, elle converge donc vers un élément de F que l'on notera $u(x)$.

$$u(\lambda x + \mu y) = \lim_{n\to\infty} u_n(\lambda x + \mu y) = \lim_{n\to\infty} (\lambda u_n(x) + \mu u_n(y)) = \lambda u(x) + \mu u(y)$$

u est donc bien linéaire, montrons qu'elle est continue.
Soit $\epsilon > 0$ et choisissons $n \in \mathbb{N}$ tel que $p, q \geq n \Rightarrow ||u_p - u_q|| \leq \epsilon$
en faisant tendre q vers l'infini (1) donne $||u_p(x) - u(x)|| \leq \epsilon||x||$, $u_p - u$ est donc Lipschitzienne, donc continue et $u = (u - u_p) + u_p$ est continue.
$||u_p - u_q|| \leq \epsilon \Rightarrow ||u_p - u|| \leq \epsilon$ $(q \to \infty)$, ce qui exprime que $\lim_{p\to\infty} u_p = u$ dans $\mathcal{L}(E, F)$ qui est donc complet.

Remarque.
a) Si $E = F$, $\mathcal{L}(E, F)$ est noté $\mathcal{L}(E)$ et on a $||u \circ v|| \leq ||u||\,||v||$.
b) Si $F = \mathbb{R}$ ou $\mathbb{C}$, $\mathcal{L}(E, F)$ est appelé dual topologique de E et noté E^*. C'est un espace vectoriel normé complet.

Définition 1.6 *Les espaces vectoriels E et F étant normés, l'application linéaire $u : E \to F$ est un isomorphisme d'espaces normés si elle est bijective et bicontinue, c'est à dire que u et u^{-1} sont continues.*

Théoreme 1.5 *L'application linéaire bijective u est un isomorphime si et seulement il existe deux réels strictement positifs A et B tels que*

$$\forall x \in E \; A||x|| \leq ||u(x)|| \leq b||x||.$$

La première inégalité est équivalente à la continuité de u^{-1} et la deuxième à la continuité de u.

Remarque.
L'espace vectoriel E étant muni des deux normes $||.||_1$ et $||.||_2$, elles donnent les mêmes ouverts (on dit que les topologies sont identiques) si et seulement si l'application identité $Id : (E, ||.||_1) \to (E, ||.||_2)$ est un isomorphisme l'image réciproque d'un ouvert (par Id ou par $Id^{-1} = Id$) étant alors un ouvert.

Corollaire 1.5.1 $Id : (E, ||.||_1 \to (E, ||.||_2)$ *est un isomorphisme d'espaces vectoriels normés si et seulement si les normes sont équivalentes c'est à dire si elles vérifient*

$$\forall x \in E \; , \; A\,||x||_1 \leq ||x||_2 \leq B\,||x||_1$$

Ce résultat est une conséquence immédiate du résultat précédent.

Théoreme 1.6 *Sur un espace vectoriel de dimension finie toutes les normes sont équivalentes, ce qui entraine qu'un tel espace n'a qu'une seule topologie.*

Démontrons d'abord un lemme préliminaire.

lemme 1.6.1 *L'espace vectoriel K^n étant muni de la norme définie par $p(x_1, x_2, \cdots, x_n) = \sqrt{|x_1|^2 + |x_2|^2 + \cdots |x_n|^2}$, toute autre norme sur cet espace est continue pour p.*

Démonstration
Soit $(e_i)_{1 \leq i \leq n}$ la base canonique de K^n et $x \in K^n$, $x = \sum_{i=1}^{n} x_i e_i$ Si $||.||$ est une norme quelconque sur K^n

$$||x|| = ||\sum_{i=1}^{n} x_i e_i|| \leq \sum_{i=1}^{n} |x_i|\,||e_i|| \leq A\,p(x) \text{ où } A = \left(\sum_{i=1}^{n} ||e_i||^2)\right)^{1/2}$$

$||.||$ est donc continue en 0. L'inégalité $|\,||x|| - ||a||\,| \leq ||x - a||$ entraine la continuité en tout point a.

Démontrons maintenant le théorème. Appelons q et r deux normes sur K^n, elles sont continues par rapport à p. Sur $S = \{x \in K^n | p(x) = 1\}$, sphère unité de K^n, q et r ne s'annulent pas, S étant compacte (fermée et bornée) r/q a un maximum et un minimum sur S

$$\forall x \in S \; , \; \exists (A, B) \in I\!R^2 \;\; A \leq \frac{r(x)}{q(x)} \leq B \Leftrightarrow Aq(x) \leq r(x) \leq Bq(x)$$

Ces inégalités se généralisent à K^n, ce qui montre que sur K^n toutes les normes sont équivalentes. Considérons un espace vectoriel E quelconque de dimension n sur K et appelons f un isomorphisme algébrique de K^n sur E (toute application linéaire envoyant la base canonique de K^n sur une base de E est un tel isomorphisme algébrique). Dans le cas où q et r sont deux normes sur E, $q \circ f$ et $r \circ f$ sont des normes sur K^n et à l'aide de ce qui précède

$$\forall x \in K^n \ , \ Aq\left(f(x)\right) \leq r\left(f(x)\right) \leq Bq\left(f(x)\right)$$

or $f(x)$ est un vecteur quelconque de E, ce qui montre que q et r sont équivalentes.
Remarque. Ceci est faux en dimension infinie.
Soit $E = C[0,1], I\!R)$, $||f||_u = \sup_{t\in[0,1]} |f(t)|$, $||f||_1 = \int_0^1 |f(t)|dt$. La suite de fonctions définie par $f_n(t) = t^n$ vérifie $||f_n||_u = 1$ et $||f_n||_1 = 1/(n+1)$. Il est façile de montrer que $||f||_1 \leq ||f||_u$ pour tout $f \in E$, mais il n'est pas possible de trouver une constante une contante k telle que $\forall f \in E$, $||f||_u \leq k||f||_1$ car $\frac{||f_n||_u}{||f_n||_1} = n+1$ et ce quotient n'est pas borné.

Théoreme 1.7 *Tout espace normé de dimension n est isomorphe à K^n, il est donc complet.*

Démonstration
Soit $f : K^n \to E$ un isomorphisme algébrique. Si q est la norme sur E, $r = q \circ f$ est une norme sur K^n, elle est donc équivalente à la norme euclidienne p

$$\forall x \in K^n \ , \ Ap(x) \leq q(f(x)) \leq Bp(x)$$

f est donc non seulement un isomrphisme algébrique mais aussi un isomorphisme d'espace normé.

Théoreme 1.8 *Toute application linéaire f d'un espace vectoriel E de dimension finie n vers un espace vectoriel quelconque F est continue.*

Démonstration
Soit $f : K^n \to E$ un isomorphisme et $g : E \to F$une application linéaire. On pose $h = g \circ f$, $h(x_1, x_2, \cdots, x_n) = \sum_{i=1}^n x_i h(e_i)$

$$||h(x_1, x_2, \cdots, x_n)|| \leq \sum_{i=1}^n |x_i| \, ||h(e_i)|| \leq \left(\sum_{i=1}^n ||h(e_i)||^2\right)^{1/2} p(x_1, x_2, \cdots, x_n)$$

ce qui entraine la continuité de h et donc également la continuité de $g = h \circ f^{-1}$.

Chapitre 2

Espaces de Hilbert

2.1 Formes Hermitiennes

Dans toute la suite on prendra $K = \mathbb{C}$.

Définition 2.1 — *E et F étant deux espaces vectoriels complexes, une application $u : E \to F$ est dite semi-linéaire si*

$$\forall x, y \in E \;\; u(x + y) = u(x) + u(y) \;,\; \forall \lambda \in \mathbb{C} \;,\; \forall x \in E \;\; u(\lambda x) = \overline{\lambda} u(x).$$

— *E, F, G étant trois espaces vectoriels complexes, une application $u : E \times F \to G$ est dite sesquilinéaire si $x \to u(x, y)$ est linéaire et $y \to u(x, y)$ est semi-linéaire. Si $G = \mathbb{C}$ u est une forme sesquilinéaire.*

— *Une forme sesquilinéaire sur l'espace vectoriel complexe $E \times E$ est hermitienne si $\forall x, y \in E \;\; f(y, x) = \overline{f(x, y)}$. Dans le cas $K = \mathbb{R}$ l'analogue des formes hermitiennes sont les formes bilinéaires symétriques.*

Exemples

1) Sur $\mathbb{C}^n \times \mathbb{C}^n$ la forme hermitienne la plus générale s'exprime à l'aide des coordonnées x_i et y_i par

$$f(x, y) = \sum_{i,j=1}^{n} \alpha_{ij} x_i \overline{y_j} \;,\; \alpha_{ij} = f(e_i, e_j) \in \mathbb{C}.$$

f est hermitienne si et seulement si $\alpha_{ji} = \overline{\alpha}_{ij} \;\; i, j = 1, 2, \cdots, n$. Dans ce cas on dit que la matrice $A = (\alpha_{ij})_{i,j}$ est hermitienne.

2) On peut munir l'espace

$$l^2 = \{(x_n)_{n \in \mathbb{N}} | \forall n x_n \in \mathbb{C} \;,\; \sum_{n=0}^{\infty} |x_n|^2 < +\infty\}$$

de la forme hermitienne définie par $f((x_n), (y_n)) = \sum_{i=0}^{\infty} x_i \overline{y_i}$. Cette série est bien définie sur l^2 car $|x_i \overline{y_i}| \leq 1/2(|x_i^2| + |y_i^2|)$. Il est alors façile de vérifier que f est une forme hermitienne.

3) Sur $E = C([0, 1], \mathbb{C})$ la forme définie par

$$f(x, y) = \int_0^1 x(t) \overline{y(t)} dt$$

est hermitienne.
4) $E = L^2_\mu(X)$ où $(X, \mathcal{A}, \mu)$ est un espace mesuré et E l'espace vectoriel constitué des classes d'équivalence des fonctions mesurables de carré intégrable, à valeurs dans $\mathbb{C}$. Soit $f, g \in E$, l'application $\phi(f, g) = \int_X f.\overline{g} d\mu$ est une forme hermitienne sur E, l'intégrale étant bien définie d'après l'inégalité de Hölder.

Théoreme 2.1 *Soit f une forme sesquilinéaire sur E, on a*

$$(1) \;\; f(x,y) = \frac{1}{4}\left(f(x+y,x+y) - f(x-y,x-y) + if(x+iy,x+iy) - f(x-iy,x-iy)\right)$$

Démonstration
$f(x+y,x+y) = f(x,x) + f(y,y) + f(x,y) + f(y,x)$
$f(x+iy,x+iy) = f(x,x) + f(y,y) - if(x,y) + if(y,x)$
En remplaçant y par $-y$ dans ces expressions et en effectuant les opérations précisées on obtient le résultat.
f est donc déterminée par les valeurs prises sur la diagonale.

Corollaire 2.1.1 *Une forme sesquilinéaire est hermitenne si et seulement si elle prend des valeurs réelles sur la diagonale de $E \times E$.*

Si f est hermitienne on a $f(x,x) = \overline{f(x,x)} \in \mathbb{R}$.
Réciproquement la formule (1) s'écrit $f(x,y) = 1/4(A - B + iC - iD)$ où $A, B, C, D \in \mathbb{R}$, donc $f(y,x) = \overline{f(x,y)}$.

Définition 2.2 *Une forme hermitienne est définie positive si $\forall x \in E, x \neq 0 \;\; f(x,x) > 0$.*

Théoreme 2.2 *Si f est une forme hermitienne définie positive sur E on a*

$$\forall (x,y) \in E \times E \; , \; |f(x,y)|^2 \leq f(x,x)\, f(y,y)$$

,c'est l'inégalité de Cauchy-Schwarz.

Démonstration
Si $x = y = 0$, l'inégalité est immédiate. Supposons par exemple $f(x,y) = a > \; 0$ $(x \neq 0)$. Posons $f(x,y) = b$, $f(y,y) = c$.

$$\forall \lambda \in \mathbb{C} \; , \; f(x+\lambda y, x+\lambda y) = a\lambda\overline{\lambda} + b\lambda + \overline{b\lambda} + c = a\left[|\lambda + \overline{b}/a|^2 + \frac{ac - |b|^2}{a^2}\right] \geq 0$$

La valeur $\lambda = -\overline{b}/a$ entraîne $ac - |b|^2 \geq 0$, ce qui est l'inégalité à démontrer.

Théoreme 2.3 *L'inégalité ci-dessus devient une égalité si et seulement si x et y sont linéairement indépendants.*

Démonstration
Si $\lambda x + y = 0$, $f(\lambda x + y, \lambda x + y) = 0 \Rightarrow |b|^2 = ac$.
Réciproquement si $|b|^2 = ac$ dans le cas $a = 0$ on a $x = 0$ et si $a > 0$
$f(\lambda x + y, \lambda x + y) = a|\lambda + \overline{b}/a|^2$ ce qui entraine $\lambda x + y = 0$ si $\lambda = -\overline{b}/a$.

Corollaire 2.3.1 *Si f est une forme hermitienne définie positive sur $E \times E$, l'application $p : x \to \sqrt{f(x,x)}$ est une norme sur E.*

Démonstration

- $p(x) = 0 \Leftrightarrow x = 0$.
- $p(\lambda x) = |\lambda| p(x)$.
- $p^2(x+y) = f(x+y, x+y) = f(x,x) + f(y,y) + 2\Re(f(x,y))$
  $\geq p^2(x) + p^2(y) + 2p(x)p(y) = (p(x)+p(y))^2$
  d'où $p(x+y) \leq p(x) + p(y)$.

2.2 Espaces de Hilbert

Définition 2.3 *Un espace de Hilbert (ou espace Hilbertien) est un espace vectoriel complexe muni d'une forme hermitienne définie positive et complet pour la norme associée à cette forme.*
Cette forme définie sur E (on dit aussi produit scalaire) sera notée $(x|y)$ et la norme associée $||x||$.

Avec ces notations on obtient

$$|(x|y)| \leq ||x|| \, ||y|| \text{ (Inégalité de Cauchy-Schwarz) et}$$

$$(x|y) = \frac{1}{4}\left(||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2\right)$$

Exemples. 1) l^2 est un espace de Hilbert associé au produit scalaire

$$(x|y) = \sum_{i=0}^{\infty} x_i \overline{y_i}$$

où $x = (x_i)_{i \in \mathbb{N}}$, $y = (y_i)_{i \in \mathbb{N}} \in l^2$.
2) L'espace $C([0,1], \mathbb{C})$ peut être muni de la forme hermitienne défine positive

$$(x|y) = \int_0^1 x(t)\overline{y(t)}dt$$

mais il n'est pas complet pour la norme associée.
Posons $x_n(t) = \sin(\pi/t)$ pour $1/n \leq t \leq 1$, $x_n(t) = 0$ pour $0 \leq t \leq 1/n$

$$||x_n - x_m||^2 = \int_{1/n}^{1/m} \sin(\pi/t)dt \ (n > m) \ , \ \lim_{n,m\to\infty} ||x_n - x_m||^2 = 0$$

Ceci montre que (x_n) est une suite de Cauchy dans cet espace. Si elle admettait une limite $x \in E$ on aurait pour $a > 0$

$$\lim_{n\to\infty} \int_a^1 |x(t) - x_n(t)|^2 dt = \int_a^1 |x(t) - \sin(\pi/t)|^2 dt = 0$$

ce qui entrainerait $x(t) = \sin(\pi/t)$ pour $0 < a \leq t \leq 1$ car $x(t) - \sin(\pi/t)$ est continue, d'où $x(t) = \sin(\pi/t)$ pour $0 < t \leq 1$ or une telle fonction ne peut être continue en 0 quelle que soit la valeur prise par $x(0)$.
3) $E = L^2_\mu(X)$ muni du produit scalaire $(f|g) = \int_X f\overline{g}d\mu$ est un espace de Hilbert. La norme $||.||_2$ en fait un espace vectoriel normé complet (cours d'intégration).

Théoreme 2.4 *Dans un espace de Hilbert on a les relations suivantes*

- $||x+y||^2 = ||x||^2 + ||y||^2 + 2\Re(x|y)$
- $||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2$. *Identité du parallélogramme.*
- $||x||^2 + ||y||^2 = 2||\frac{x+y}{2}||^2 + \frac{1}{2}||x-y||^2$. *Théorème de la médiane.*

La vérification est immédiate.
Notons que si $(x|y) = 0$ la première relation est le théorème de Pythagore, mais la réciproque n'est vraie que si l'espace vectoriel est réel.

2.3 Théorème de projection

Définition 2.4

- *Soit F une partie d'un espace de Hilbert E et $a \in E$. On appelle projection de a sur F le point $\alpha \in F$ (s'il existe) tel que*

$$||a-\alpha|| = \inf_{x\in F} ||a-x||.$$

C'est le point réalisant le minimum de la distance de a à F (si ce point existe et dans ce cas il peut y en avoir plusieurs).

- *Une partie F d'un espace vectoriel E est convexe si lorsqu'elle contient deux points elle contient le segment joignant ces deux points, où le segment $[a,b]$ est l'ensemble des points $\{x \in E | x = \alpha a + (1-\alpha)b \ , \ 0 \leq \alpha \leq 1\}$*

Théoreme 2.5 *Soit E un espace Hilbertien et F une partie fermée convexe non vide de E. Chaque point de E a une projection unique sur F.*

Démonstration
On pose $d = \inf_{x\in F} ||a-x||$. Les propriétés de la bornes inférieure entraine l'existence d'une suite (x_n) de points de F telle que $\lim_{n\to\infty} ||a-x_n|| = d$. En appliquant le théorème de la médiane à $x = a - x_n$ et $y = a - x_m$ on obtient

$$(1) \ ||a-x_n||^2 + ||a-x_m||^2 = 2||a - \frac{x_n + x_m}{2}||^2 + 1/2||x_n - x_m||^2$$

$\lim_{n\to\infty} ||a-x_n||^2 = \lim_{m\to\infty} ||a-x_m||^2 = d^2$, F étant convexe $\frac{x_n+x_m}{2} \in F$. La définition de d entraine $||a - \frac{x_n+x_m}{2}|| \geq d$. On a $\limsup_{n,m\to\infty} ||x_n - x_m|| \leq 0$ ce qui donne $\lim_{n,m\to\infty} ||x_n - x_m|| = 0$. La suite (x_n) est donc une suite de Cauchy de F, E étant complet, elle converge vers un élément $\alpha \in E$ et F étant fermé $\alpha \in F$. Ceci montre l'existence d'un point $\alpha \in F$ vérifiant $d = ||a-\alpha||$, α est donc bien une projection de a sur F.
S'il existait une autre projection β on pourrait écrire

$$||\beta-\alpha||^2 = 2||\beta-a||^2 + 2||\alpha-a||^2 - 4||a - \frac{\alpha+\beta}{2}||^2$$

d'où $||\beta-\alpha||^2 \leq 2d^2 + 2d^2 - 4d^2 \leq 0 \Rightarrow ||\beta-\alpha|| = 0 \Rightarrow \alpha = \beta$.

Théoreme 2.6 (Caractérisation de la projection) *Soit F partie fermée convexe non vide de l'espace de Hilbert E, et $a \in E$ on a*

$$\alpha \textit{ est la projection de } a \textit{ sur } F \Leftrightarrow \forall x \in F \ , \ \Re(a-\alpha | x-\alpha) \leq 0$$

Démonstration
Si $\Re(a-\alpha|x-\alpha) \leq 0$ pour $x \in F$ on a

$$||a-x||^2 = ||a-\alpha||^2 + ||x-\alpha||^2 - 2\Re(a-\alpha|x-\alpha) \geq ||a-\alpha||^2$$

et donc α est bien la projection de a sur F.
Réciproquement, si α est la projection de a sur F on a pour $x \in F$
$(1-t)\alpha + tx \in F$, $t \in [0,1]$ d'après la convexité de F.

$\forall x \in F$, $\forall t \in [0,1]$, $||a-((1-t)\alpha+tx)||^2 = ||(a-\alpha)-t(x-\alpha)||^2 \geq ||a-\alpha||^2$

Après developpement et simplification on obtient
$-2t\Re(a-\alpha-x-\alpha) + t^2||x-\alpha||^2 \geq 0$ en divisant par $t \in]0,1]$ et en faisant tendre t vers 0 on obtient $\Re(a-\alpha|x-\alpha) \leq 0$ pour tout $x \in F$.
Remarque.
Si le corps de base était $I\!R$, en posant $\cos\theta = \frac{(x|y)}{||x||\,||y||}$ ceci exprime que l'angle θ est obtu ou droit.

2.4 Projection sur un sous-espace vectoriel fermé

Un sous-espace vectoriel étant convexe, ce qui précède s'applique aux sous-espaces fermés d'un espace de Hilbert E.

Définition 2.5 — *Les vecteurs x et y de E sont orthogonaux si $(x|y) = 0$.*
- *F étant une partie de E, x est orthogonal à F s'il est orthogonal à tout élément de F.*
  On note $F^\perp$ l'ensemble des éléments orthogonaux à F.Cet ensemble, appelé orthogonal de F est un sous-espace vectoriel fermé de E car $F^\perp = \cap_{a\in F}\{a\}^\perp$ et l'orthogonal d'un élément $a \in E$ est le noyau de la forme linéaire continue $\phi_a : x \to (a|x)$, c'est donc $\phi_a^{-1}(\{0\})$ qui est fermé.
- *Deux parties F et G sont orthoganales si tout vecteur de l'une est orthogonal à tout vecteur de l'autre.*

Remarque.
En appelant G l'espace vectoriel engendré par F on a $F^\perp = G^\perp = \overline{G}^\perp$. La première égalité est vraie car un vecteur orthogonal à d'autres vecteurs est orhogonal à leurs combinaisons linéaires, la deuxième s'obtient par passage à la limite en utilisant la continuité des ϕ_a.

Théoreme 2.7 (de projection sur un sous espace fermé) *La projection d'un point a d'un espace Hilbertien E sur un sous-espace fermé F est l'unique point α de F tel que $a-\alpha$ soit orthogonal à F.*

Démonstration Soit $x \in F$, le vecteur $y = x+\alpha$ est alors élément de F qui est espace vectoriel.
$\Re(a-\alpha|y-\alpha) = \Re(a-\alpha|x) \leq 0$, en remplaçant x par $-x$ on obtient $\Re(a-\alpha|x) = 0$, puis en remplaçant x par ix on obtient $\Im(a-\alpha|x) = 0$ ce qui entraine $(a-\alpha|x) = 0$ pour tout $x \in F$, ce qui exprime que $a-\alpha \perp F$.

Si $\beta \in F$ est tel que $a - \beta \perp F$ on a $(a - \alpha|x) = (a - \beta|x) = 0$ pour tout $x \in F$, en faisant la différence on obtient $(\alpha - \beta|x) = 0$ et si $x = \alpha - \beta$

$$(x|x) = 0 \Rightarrow x = \alpha - \beta = 0 \Rightarrow \alpha = \beta$$

il existe donc un seul point α vérifiant $a - \alpha \perp F$.
Dans la suite on notera P_F la projection de E sur F.

Théoreme 2.8 *Si F n'est pas réduit au sous-espace $\{0\}$ P_F est linéaire, continue, de norme 1. $\ker P_F = F^\perp$ et $E = F \oplus F^\perp$ (les deux sous-espaces sont supplémentaires).*

Démonstration
Soient $x, y \in E$, $\forall z \in F$ on a $(P_F x - x|z) = (P_F y - y|z) = 0$
d'où $(P_F x + P_F y - (x + y)|z) = 0$ ce qui entraine $P_F(x + y) = P_F(x) + P_F(y)$.
De même si $\lambda \in \mathbb{C}$ $(\lambda P_F(x) - \lambda x|z) = 0 \Rightarrow P_F(\lambda x) = \lambda P_F(x)$. P_F est donc linéaire.
L'égalité $(P_F(x) - x|P_F(x)) = 0$ entraine $||P_F x||^2 = (x|P_F x) \leq ||x||\,||P_F x||$ donc $||P_F x|| \leq ||x||$ et $||P_F|| \leq 1$.
Si $x \in F$ on a $||P_F x|| = ||x||$ d'où $||P_F|| \geq 1$ et donc $||P_F|| = 1$.

$$P_F x = 0 \Leftrightarrow \forall z \in F \ (x|z) = 0 \Leftrightarrow x \in F^\perp$$

ce qui démontre que $\ker P_F = F^\perp$.
Un vecteur $x \in E$ s'écrit $x = (x - P_F x) + P_F x$, le premier vecteur est un élément de $F^\perp$ et le second un élément de F, on a de plus $F \cap F^\perp = \{0\}$ ces deux résultats établissent que $E = F \oplus F^\perp$.

lemme 2.8.1 *Soit v un vecteur d'un espace de Hilbert E. $f_v : x \to (x|v)$ est une forme linéaire continue sur E de norme $||v||$. L'application $v \to f_v$ est une isométrie semi linéaire de E vers E^*.*

Démonstration
En appliquant l'inégalité de Cauchy-Schwarz on a
$|(x|v)| \leq ||x||\,||v|| \Rightarrow ||f_v|| \leq ||v||$. De plus $f_v(v) = ||v||^2$, on déduit donc
$||f_v|| \geq ||v||$ d'où $||f_v|| = ||v||$.
$v \to f_v$ est bien une isométrie semi-linéaire de E dans E^*, le théorème suivant montre qu'elle est surjective.

Théoreme 2.9 *Si f est une forme linéaire continue sur l'espace de Hilbert E, il existe un vecteur unique v tel que $\forall x \in E$, $f(x) = (x|v)$.*

Démonstration Si $f = 0$ $v = 0$ convient. Sinon $F = f^{-1}(\{0\})$ est un hyperplan fermé, f étant continue. Le théorème de projection entraine l'existence d'un vecteur $u \perp F$. Posons $g(x) = (x|u)$, cette forme linéaire a le même noyau que f, on a donc $g = \lambda f$ pour un $\lambda \in \mathbb{C}$, soit $\forall x \in E$ $f(x) = \lambda(x|u)$, on prend alors $v = \overline{\lambda} u$
Application.
Pour toute forme linéaire continue ϕ sur $E = L^2_\mu(X)$ on a :

$$\exists g \in E \ , \ \forall f \in E \ , \ \phi(f) = \int_X f\overline{g}d\mu \text{ et } ||\phi|| = ||g||_2.$$

On peut donc identifier E et son dual topologique.

2.5 Espaces de Hilbert et systèmes orthogonaux

Définition 2.6 — *Un espace de Hilbert E est séparable s'il existe un sous-ensemble dénombrable $A = \{x_1, x_2, \cdots\}$ dense dans E, c'est à dire tel que $\overline{A} = E$. Il est équivalent de dire que toute boule ouverte non vide de E contient au moins un élément de A.*

— *Une famille $(e_i)_{i\in I}$ de vecteurs de E est orthonormée si $(e_i|e_j) = 0$ pour $i \neq j$ et $||e_i|| = 1$pour tout $i \in E$.*

Théoreme 2.10 — *Toute famille orthonormée est libre, ce qui signifie que toute combinaison linéaire composée d'un nombre fini de vecteurs donne le vecteur nul seulement si tous les coefficients scalaires sont nuls.*

— *Dans un espace de Hilbert séparable toute famille orthonormée est finie ou dénombrable.*

Démonstration

Soit J une partie finie quelconque de I et $(e_i)_{i\in I}$ une famille orthonormée.

$$\sum_{i\in J} \alpha_i e_i = 0 \Rightarrow \forall j \in J\ (e_j | \sum_{i\in J} \alpha_i e_i) = 0 = \sum_{i\in J} \overline{\alpha_i}(e_j|e_i) = \overline{\alpha_j}$$

ce qui démontre qu'une famille orthonormée est une famille libre.
Soit $(e_i)_{i\in I}$ une famille orthonormée dans un espace de Hilbert séparable E.Si $e_j \neq e_i$ $||e_j - e_i||^2 = ||e_j||^2 + ||e_i||^2 - 2\Re(e_j|e_i) = 2$ donc $||e_j - e_i|| = \sqrt{2}$.
Soit $A = \{x_1, x_2, \cdots\}$ un ensemble dénombrable dense dans E. Pour chaque $i \in I$ il existe un entier $n(i)$ tel que $||x_{n(i)} - e_i|| < 1/4$ car A est dense dans E.Montrons que l'application $i \to n(i)$ est injective. Soit $n(i) = n(j)$

$$||e_i - e_j|| \leq ||e_i - x_{n(i)}|| + ||x_{n(i)} - e_j|| < 1/2 \Rightarrow e_i = e_j.$$

Cette application de I dans $\mathbb{N}$ étant injective, l'ensemble I est fini ou dénombrable.
Dans la suite l'espace E sera toujours supposé séparable

Théoreme 2.11 *Soit $(e_i)_{1\leq i\leq n}$ un système orthonormé fini de E et F le sous espace à n dimensions engendré par ces vecteurs. Pour tout x de E on a*

1) $P_F x = \sum_{i=1}^n (x|e_i)e_i$.
2) $d^2(x, F) = ||x||^2 - \sum_{i=1}^n |(x|e_i)|^2$ *où* $d(x, F) = \inf_{y\in F} d(x, y)$ $= \inf_{y\in F} ||x - y||$.

Démonstration

1) On a $\forall j$, $1 \leq j \leq n$ $(x - \sum_{i=1}^n (x|e_i)e_i|e_j) = 0$, donc $x - \sum_{i=1}^n (x|e_i)e_i \perp F$.
2) $d^2(x, F) = ||x - \sum_{i=1}^n (x|e_i)e_i||^2 = ||x||^2 - \sum_{i=1}^n |(x|e_i)|^2 - \sum_{i=1}^n |(x|e_i)|^2 + \sum_{i=1}^n |(x|e_i)|^2$. Ceci démontre que

$$\forall(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{C}^n\ ,\ ||x - \sum_{i=1}^n \lambda_i e_i|| \geq ||x - \sum_{i=1}^n (x|e_i)e_i||$$

.

Corollaire 2.11.1 *Inégalité de Bessel Soit $(e_i)_{i\in I}$, L'ensemble I étant fini ou dénombrable, un système orthonormé quelconque et $x \in E$, on a*

$$\sum_{i\in I} |(x|e_i)|^2 \leq ||x||^2.$$

La procédure qui suit, qui permet de construire des systèmes orthonormés en dimension finie s'étend aux espaces de Hilbert.

Théoreme 2.12 (Procédé d'orthogonalisation de Gram-Schmidt) *Soit E un espace de Hilbert et $(x_n, n \geq 1)$ des vecteurs linéairementindépendants de E. On définit*

$$\begin{array}{lcll} w_1 & = & x_1 & , v_1 = w_1/||w_1|| \\ w_2 & = & x_2 - (v_1|x_2)v_1 & , v_2 = w_2/||w_2|| \\ & \vdots & & \\ w_n & = & x_n - \sum_{k=1}^{n-1}(v_k|x_n)v_k & , v_n = w_n/||w_n|| \\ & \vdots & & \end{array}$$

La suite $(v_n, n \geq 1)$ est un système orthonormé et pour tout $n \geq 1$, $(v_1, v_2, \cdots, v_n)$ et $(x_1, x_2, \cdots, x_n)$ engendrent le même sous espace vectoriel de E.

La démonstration de ce théorème est aisée et laissée en exercice. Application.
Déterminer

$$\min_{a,b,c} \int_{-1}^{1} (x^3 - a - bx - cx^2)^2 dx.$$

Considérons $E = L^2([-1, 1], I\!R)$ et$F$ le sous espace vectoriel de dimension 3 engendré pa f_0, f_1, f_2 où $f_i : x \to x^i$. F est complet car c'est un sous espace de dimension finie. Le problème revient à calculer la distance d de $f : x \to x^3$ à ce sous espace (l'intégrale est égale à d^2). Déterminons tout d'abord $P_F f$, l'intégrale cherchée sera $||f - P_F f||^2$.
Construisons une base orthonormée de F par la méthode de Gram-Schmidt.
$v_0(x) = \frac{f_0(x)}{||f_0||} = \sqrt{2}/2$
$w_1(x) = x - (v_0|f_1)v_0 = x$, $v_1(x) = (\sqrt{6}/2)x$
$w_2(x) = x^2 - (v_0|f_2)v_0 - (v_1|f_2)v_1 = x^2 - 1/3$, $v_2 = (2\sqrt{2}/3)(x^2 - 1/3)$.

$$P_F f = (f|v_0)v_0 + (f|v_1)v_1 + (f|v_2)v_2 = (3/5)x \; ||f - P_F f||^2 = \int_{-1}^{1} (x^3 - 3/5x)^2 dx = 8/175$$

Définition 2.7 — *Un système orthonormé $A = (e_i, i \geq 1)$ est complet si*

$$\forall x \in E \ , \ \forall \epsilon > 0 \ , \ \exists n_0 \in I\!N \, ||x - \sum_{i=1}^{n_0} (x|e_i)e_i|| \leq \epsilon.$$

- *Un système orthonormé est fermé si pour tout $x \neq 0$ de E on peut trouver $n \in I\!N^*$ tel que $(e_n|x) \neq 0$.*
- *Un système orthonormé est maximal s'il n'est pas sous ensemble propre d'un autre système orthonormé.*

Théoreme 2.13 *Soit $A = (e_i, i \in I)$, $I = \{1, 2, \cdots, n\}$ ou $I = I\!N^*$, un système orthonormé dans un espace de Hilbert E. Les propriétés suivantes sont équivalentes.*

1) A est complet.
2) A est maximal dans E.
3) A est fermé.
4) Si $x \in E$, $(x|x) = ||x||^2 = \sum_{i\in I} |(x|e_i)|^2$. (Identité de Parseval).
5) Si $x, y \in E$, $(x|y) = \sum_{i\in I}(x|e_i)\overline{(y|e_i)}$.

<u>Démonstration</u>
$1 \Rightarrow 2$
Si A n'est pas maximal dans E $\exists e_0 \in E$ tel que $(e_i, i \in I\!N)$ est un système orthonormé. Soit $c_i = (e_0|e_i)$

$$\forall n \in I\!N\,,\ ||e_0 - \sum_{i=1}^{n} c_i e_i|| = ||e_0|| = 1$$

A n'est donc pas complet.
$2 \Rightarrow 3$
Si A n'est pas fermé, il existe $x \neq 0$, $x \in E$ tel que quel que soit $i \in I$, $(x|e_i) = 0$. En posant $e_o = x/||x||$ le système $(e_i, i \in I\!N)$ est orthonormé et A n'est donc pas maximal.
$3 \Rightarrow 4$
L'inégalité de Bessel entraine $\sum_{i\in I} |(x|e_i)|^2 \leq ||x||^2$, montrons l'inégalité inverse dans le cas le plus général $I = I\!N^*$. Posons $s_n = \sum_{i=1}^{n}(x|e_i)e_i$

$$||s_n - s_m||^2 = \left(\sum_{i=m+1}^{n}(x|e_i)e_i| \sum_{i=m+1}^{n}(x|ei)e_i\right) = \sum_{i=m+1}^{n}|(x|e_i)|^2 \leq \sum_{i=m+1}^{\infty}|(x|e_i)|^2\ (n > m).$$

Le dernier terme est le reste d'une série convergente (inégalité de Bessel) et tend donc vers zéro quand n et m tendent vers zéro, (s_n) est donc une suite de Cauchy de E, convergeant vers un élément $y \in E$ car E est complet.

$$\forall n \in I\!N^*(e_n|x) = (e_n|y) \Rightarrow \forall n\ (e_n|x - y) = 0 \Rightarrow x = y$$

car le système est fermé.
$4 \Rightarrow 5$
$(x|y)$ et $\sum_{i\in I}(x|e_i)\overline{(y|e_i)}$ sont deux formes hermitiennes prenant les mêmes valeurs sur la diagonale, donc indentiques.
$5 \Rightarrow 4$ est immédiat.
$4 \Rightarrow 1$
Soit n_0 tel que

$$n > n_o \Rightarrow ||x||^2 - \sum_{i=1}^{n}|(x|e_i)|^2 < \epsilon\ ,\ \text{alors}\ ||x - \sum_{i=1}^{n}(x|e_i)e_i||^2 = ||x||^2 - \sum_{i=1}^{n}|(x|e_i)|^2 < \epsilon.$$

Théoreme 2.14 *Un espace de Hilbert est séparable si et seulement il contient un système orthonormé complet fini ou dénombrable.*

<u>Démonstration</u>
Supposons E séparable et soit (x_n) un sous ensemble dénombrable qui est dense.

On peut en extraire une sous suite (x_{n_k}) de vecteurs linéairement indépendants telle que tout x_n soit une combinaison linéaire finie des x_{n_k}. Le procédé de Gram-Schmidt appliqué à cette suite permet de construire un système orthonormé (y_n) tel que tout x_n soit combinaison linéaire finie des y_n. Un vecteur $x \in E$ orthogonal à tous les y_n est donc orthogona à tous les x_n donc à $E = \overline{(x_n)}$ ce qui entraine que le système (y_n) est fermé.
Nous admettrons la réciproque de ce théorème.

Théoreme 2.15 *Un espace de Hilbert séparable est isométrique à l'espace $l^2 = l^2(\mathbb{N})$ s'il est de dimension infinie, à un espace $\mathbb{C}^n$ s'il est de dimension finie.*

Démonstration
Si $\dim E = n$ l'application $x = \sum_{i=1}^n (x|e_i)e_i \rightarrow ((x|e_1), (x|e_2), \cdots (x|e_n))$, où $(e_i, 1 \leq i \leq n)$ est une base orthonormée de E, est une isométrie car $||x||^2 = \sum_{i=1}^n |(x|e_i)|^2$.
Si E est de dimension infinie choisissons $A = (e_i, i \in \mathbb{N}^*)$ un système orthonormé complet et posons $x_i = (x|e_i)$. On a démontré que $||x||^2 = \sum_{i=1}^{infty} |x_i|^2$, l'application

$$\phi : E \rightarrow l^2 \;\; x \rightarrow (x_i, i \in \mathbb{N}^*)$$

est donc une isométrie, montrons qu'elle est surjective. On peut remarquer que la famille $B = (f_i, i \in \mathbb{N}^*)$, définie par $f_i(j) = \delta_i^j$, est un système orthonormé complet de l^2 car $(f_i|f_j) = \delta i^j$ et $||(a_i)||_{l^2} = \sum_{i=1}^{\infty} -(a|f_i)|^2$.
$B \subset \phi(E)$, $\phi(E)$ contenant les combinaisons linéaires finies d'éléments de B est donc dense dans l^2, or $\phi(E)$ est complet comme image isométrique d'un espace de Hilbert, il est donc fermé dans l^2, on a donc $\phi(E) = l^2$. Nous allons maintenant étudier un cas particulier important comprenant la convergence en moyenne quadratique (dans L^2) des séries de Fourier.

2.6 Bases orthonormales de $L^2([a, b])$

Théoreme 2.16 (Critère de Dalzell) *Une famille orthonormée $(f_n, n \geq 1)$ de $L^2_m([a, b])$, m étant la mesure de Lebesgue de $[a, b]$, est un système orthonormé complet si et seulement si*

$$\frac{2}{(a-b)^2} \sum_{n \geq 1} \int_a^b | \int_a^u f_n(t)dt|^2 du = 1$$

Démonstration
Rappellons que si A est un sous ensemble de E la fonction indicatrice χ_A notée aussi $\mathbf{1_A}$ est définie par $\chi_A(x) = 0$ si $x \notin A$ et $\chi_A(x) = 1$ si $x \in A$.
Supposons le système $(f_n, n \geq 1)$ orthonormé complet. $\forall u \in [a, b]$, $\chi_{[a,u]} \in L^2_m$ on peut donc écrire

$$\chi_{[a,u]} = \sum_{n=1}^{\infty} c_n f_n \; , \; c_n = (\chi_{[a,u]}|f_n) = \int_a^u \overline{f_n(t)} dt$$

L'égalité de Parseval entraine

$$\sum_{n=1}^{\infty} | \int_a^u f_n(t)dt|^2 = u - a \Rightarrow \sum_{n=1}^{\infty} \int_a^b | \int_a^u f_n(t)dt|^2 du = \frac{(b-a)^2}{2}.$$

Nous admettrons la réciproque de ce théorème.

2.6.1 Application

Corollaire 2.16.1 *La famille de fonctions*

$$e_n : x \to \frac{e^{inx}}{\sqrt{2\pi}} \ , \ n \in \mathbb{Z}$$

est un système orthonormé complet de l'espace de Hilbert $E = L^2_m([0, 2\pi])$*.*

Démonstration
Montrons que cette famille de fonctions vérifient le critère de Dalzell.

$$\frac{2}{4\pi^2} \sum_{n \in \mathbb{Z}} \int_0^{2\pi} \frac{1}{2\pi} | \int_0^u e^{int} dt|^2 du = \frac{2}{8\pi^3} (\frac{8\pi^3}{3} + \sum_{n \in \mathbb{Z}^*} \frac{4\pi}{n^2})$$

$$= 2/3 + \frac{2}{8\pi^3} (2 \sum_{n=1}^{\infty} \frac{4\pi}{n^2}) = 2/3 + \frac{2}{8\pi^3} . \frac{8\pi^3}{6} = 1.$$

Une fonction f étant élément de E on a

$$(f|e_n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx} dx \text{ et } f = \sum_{n \in \mathbb{Z}} (f|e_n) e_n \text{ dans } E$$

$$f = \sum_{n \in \mathbb{Z}} \left(\frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx \right) f_n \text{ où } f_n(x) = e^{inx}.$$

On note $c_n(f) = (1/2\pi) \int_0^{2\pi} f(x) e^{-inx} dx$, ce qui entraine $f = \sum_{n \in \mathbb{Z}} c_n(f) f_n$ dans E, ce qui signifie

$$\lim_{N \to \infty} ||f - \sum_{n=-N}^{N} c_n(f) f_n||_2 = 0.$$

La série $\sum_{n \in \mathbb{Z}} c_n(f) e^{inx}$ est appelée série de Fourier de f. En écrivant $e^{inx} = \cos(nx) + i \sin(nx)$ la série précédente peut également s'exprimer sous la forme $\sum_{n=0}^{\infty} a_n \cos(nx) + b_n \sin(nx)$. On a les relations $a_0 = c_0$, $a_n = c_n + c_{-n}$, $b_n = ic_n - ic_{-n}$.

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \ , \ a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \ (n > 0) \ , \ b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \ (n > 0).$$

Remarque.

- On peut supposer $f(0) = f(2\pi)$ car les coefficients de la série de Fourier sont inchangés en modifiant un nombre fini de valeurs prises par f, ceci permet de considérer f comme la restriction à $[0, 2\pi]$ d'une fonction 2π-périodique.
- Le calcul des coefficients peut dans ces conditions être effectué sur $[a, a + 2\pi]$ avec a quelconque.

— Si f est une fonction paire, $\forall n > 0$ $b_n = 0$ et si elle est impaire $\forall n \in I\!N$ $a_n = 0$.

Les résultats du théorème 2.13 se traduisent de la façon suivante pour les séries de Fourier.

4) Identité de Parseval

$$\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2dt = \sum_{k\in\mathbb{Z}}|c_k|^2 = |c_0|^2 + \sum_{k=1}^{\infty}(|c_k|^2 + |c_{-k}|^2)$$

$$\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2dt = |a_0|^2 + \frac{1}{2}\sum_{k=1}^{\infty}(|a_k|^2 + |b_k|^2)$$

pour une fonction $f \in L^2_m([0, 2\pi])$

5) $\forall f, g \in L^2_m([0, 2\pi])$, $\frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}dt = \sum_{n\in\mathbb{Z}} c_n(f)\overline{c_n(g)}$.

L'isométrie entre $L^2_m([0, 2\pi])$ et l^2 se traduit par
pour toute suite $(c_n, n \in \mathbb{Z})$ vérifiant $\sum_{n\in\mathbb{Z}}|c_n|^2 < +\infty$ il existe $f \in L^2_m([0, 2\pi])$ telle que la série de Fourier de f soit $\sum_{n\in\mathbb{Z}} c_n e^{inx}$.

2.6.2 Convergence ponctuelle d'une série de Fourier

Les coefficients a_n , b_n , c_n sont également définis si $f \in L^1_m([0, 2\pi])$ car $|f(t)e^{int}| = |f(t)|$. Dans le cas où f est de plus 2π-périodique le problème de la convergence ponctuelle (convergence simple ou uniforme) de $S_N(f) = \sum_{n=-N}^{N} c_n e^{inx}$ se pose. Le théorème de Dirichlet, que nous allons démontrer donne des conditions pour la convergence ponctuelle de cette série. Commençons par transformer $S_N(f)$.

L'expression $D_N(t) = \sum_{n=-N}^{N} c_n e^{inx}$ est appelé noyau de Dirichlet.

$$S_N(f)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(t)(\sum_{n=-N}^{N} e^{in(x-t)})dt = \frac{1}{2\pi}\int_0^{2\pi} f(t)D_N(x-t)dt$$

$$\begin{aligned} S_N(f)(x) &= \frac{1}{2\pi}\int_0^{2\pi} f(x+t)D_N(t)dt = \frac{1}{2\pi}\int_0^{2\pi} f(x-t)D_N(t)dt \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}(f(x+t) + f(x-t))D_N(t)dt. \end{aligned}$$

La deuxième égalité provient de la parité de la fonction D_N.

Théoreme 2.17 (Lemme de Riemann-Lebesgue.) *On considère une fonction $f \in L^1_m([0, 2\pi])$, les coefficients $c_n = (1/(2\pi))\int_0^{2\pi} f(x)e^{-inx}dx$ de la série de Fourier de f tendent vers zéro quand $|n|$ tend vers zéro.*

Si $f \in L^2_m([0, 2\pi])$ c'est immédiat car le terme général d'une série convergente (identité de Parseval) tend vers zéro quand $n \to \infty$ sinon $L^2_m([0, 2\pi])$ est dense dans $L^1_m([0, 2\pi])$ et si $f \in L^1$ il existe $g \in L^2$ telle que $||f - g||_1 < \epsilon$.

$$|c_n(f)| \le |c_n(f) - c_n(g)| + |c_n(g)| \le ||f - g||_1 + |c_n(g)|$$

et $c_n(f)$ tend bien vers zéro quand $|n|$ tend vers l'infini.

Remarque.

Ce résultat peut être démontré pour un intervalle $[a, b]$ quelconque.

Théoreme 2.18 (Critère de Dini) *Si* $f \in L^1_m([-\pi,\pi])$, $A \in I\!R$, *f 2π-périodique et si*

$$\int_0^\pi \frac{|f(x+t)+f(x-t)-2A|}{t}dt < +\infty$$

la série de Fourier de f converge en x et $\lim_{n\to\infty} S_N(f)(x) = A$

On peut remarquer que

$$D_N(t) = e^{-2iNx}(1+e^{ix}+\cdots+e^{2iNx}) = e^{-2iNx}\frac{1-e^{i(2N+1)x}}{1-e^ix} = \frac{\sin(2N+1)t/2}{\sin(t/2)}$$

$(x \neq o \mod 2\pi)$
de plus $\int_{-\pi}^{\pi} D_N(t)dt = 2\pi$ et $\int_0^\pi D_N(t)dt = \pi$ cette dernière valeur étant obtenue par parité de D_N.

$$S_N(f)(x) - A = \frac{1}{2\pi}\int_0^\pi \frac{f(x+t)+f(x-t)-2A}{t}\frac{t}{\sin(t/2)}\sin((N+1/2)t)dt$$

ceci tend vers zéro d'après le lemme de Riemann-Lebesgue car les deux premiers termes de l'intégrale sont intégrables (le premier d'après l'hypothèse) et le troisième est la partie imaginaire de $e^{i(N+1/2)t}$.

Théoreme 2.19 (Théorème de Dirichlet) *Soit f une fonction 2π-périodique et intégrable sur $[0,2\pi]$ admettant une limite à gauche et à droite en $x \in [0,2\pi]$ et dérivable à gauche et à droite en ce point. La série de Fourier de f converge en x vers la demi-somme des limites des limites à gauche et à droite.*

$$\lim_{N\to\infty} S_N(f)(x) = \frac{f(x+)+f(x-)}{2}.$$

Ceci est une conséquence immédiate du critère de Dini avec
$A = (f(x+)+f(x-))/2$ de plus

$$t \to \frac{|f(x+t)+f(x-t)-f(x+)-f(x-)|}{t}$$

est intégrablesur $[0,\pi]$, f étant dérivable à gauche et à droite de x.
Remarque.
Si les conditions du théorème sont vérifiées et si f est continue en x la série de Fourier de f converge alors vers $f(x)$ en x.

2.7 Opérateur linéaire sur un espace de Hilbert

Un opérateur linéaire sur E est une application linéaire de E dans E, sauf mention contraire ils seront continus.

Théoreme 2.20 *Une forme sesquilinéaire est continue sur $E\times E$ si et seulement si*

$$\exists M \in I\!R^+ \ , \ \forall (x,y) \in E\times E \ , \ |f(x,y)| \leq M||x||\,||y||.$$

Démonstration
Supposons f continue, la continuité en 0 se traduit par

$$\exists a, b \in \mathbb{R}_+^* \ , \ ||x|| < a \ , \ ||y|| < b \Rightarrow |f(x,y)| \leq 1$$

Par homogénéité on a

$$\forall x, y \in E \ , \ |f(x,y)| \leq \frac{1}{ab} ||x|| \, ||y||.$$

Réciproquement, si la condition est vérifiée on a $x, y, \xi, \eta \in E$

$$|f(x+\xi, y+\eta) - f(x,y)| \leq M \left(||\xi|| \, ||y+\eta|| + ||\eta|| \, ||x|| \right).$$

Ceci montre la continuité de f.

Théoreme 2.21 — *Si u est un opérateur linéaire sur E, $\phi_u : (x,y) \to (ux|y)$ définit une forme sesquilinéaire continue sur $E \times E$ telle que $||\phi_u|| = ||u||$.*
— *Réciproquement si ϕ est une forme sesquilinéaire continue sur $E \times E$ il existe $u \in \mathcal{L}(E)$ tel que $\phi = \phi_u$.*

Remarque. On définit

$$||\phi|| = \sup_{x \neq 0, y \neq 0} \frac{\phi(x,y)}{||x|| \, ||y||}$$

et donc

$$|\phi(x,y)| \leq ||\phi|| \, ||x|| \, ||y||$$

et le M du théorème est tel que $M \geq ||\phi||$.

Démonstration
La sesquilinéarité de ϕ_u est immédiate. L'inégalité de Cauchy-Schwarz entraine

$$|(ux|y)| \leq ||ux|| \, ||y|| \leq ||u|| \, ||x|| \, ||y|| \Rightarrow ||\phi_u|| \leq ||u||.$$

En prenant $y = ux$ on obtient

$$||ux||^2 = \phi_u(x, ux) \leq ||\phi_u|| \, ||x|| \, ||ux|| \Rightarrow ||ux|| \leq ||\phi_u|| \, ||x|| \Rightarrow ||u|| \leq ||\phi_u||$$

ces deux inégalités permettant d'écrire $||\phi_u|| = ||u||$.
Réciproquement soit ϕ une forme sesquilinéaire continue sur $E \times E$. L'application $g : y \to \phi(x,y)$ est linéaire continue (car $|g(y)| \leq (||\phi|| \, ||x||)||y||$). D'après le théorème 2.9 il existe $v \in E$ tel que $g(y) = (y|v)$ pour tout $y \in E$, ou encore $(v|y) = \phi(x,y)$. Posons $v = ux$, on vérifie que $u(\lambda x + \mu y) = \lambda u(x) + \mu u(y)$ si $x, y \in E \ , \ \lambda, \mu \in \mathbb{C}$, u est donc linéaire.

$$||ux||^2 = \phi(x, ux) \leq ||\phi|| \, ||x|| \, ||ux|| \Rightarrow ||ux|| \leq ||\phi|| \, ||x||$$

l'application linéaire u est donc continue.

Théoreme 2.22 *A tout opérateur $u \in \mathcal{L}(E)$ (opérateur linéaire continu sur E) il existe un élément unique $u^* \in \mathcal{L}(E)$ tel que*

$$\forall x, y \in E\ ,\ (ux|y) = (x|u^*y).$$

u^* est appelé opérateur adjoint de u.

Démonstration
Posons $\psi(y,x) = (y|ux)$. C'est une forme sesquilinéaire continue de norme $||u||$, d'après le théorème précédent il existe $u^* \in \mathcal{L}(E)$ tel que $\psi = \phi_{u^*} : (y,x) \to (u^*y|x)$, d'où

$$\forall (x,y) \in E \times E\ ,\ (x|u^*y) = (ux|y) \text{ et } ||\phi_{u^*}|| = ||u^*|| = ||\psi|| = ||u||$$

d'où $||u^*|| = ||u||$.
Propriétés.

1) $u^{**} = u$.
2) $(uv)^* = v^*u^*$.
3) $||u^*|| = ||u||$.
4) $||uu^*|| = ||u^*u|| = ||u||^2 = ||u^*||^2$.
5) $(\lambda u)^* = \overline{\lambda} u^*$.
6) $(u+v)^* = u^* + v^*$.
7) Si u est inversible, u^*est inversible et $(u^*)^{-1} = (u^{-1})^*$.

1,2,3,5,6 sont immédiats, 4 sera montré plus loin. Montrons 7. Posons $v = u^{-1}$.

$$uv = vu = I \Rightarrow v^*u^* = u^*v^* = I \Rightarrow u^* \text{ est inversible et } (u^*)^{-1} = (u^{-1})^*.$$

.

Définition 2.8 — *$u \in \mathcal{L}(E)$ est normal si $u^*u = uu^*$.*
— *u est unitaire si $u^*u = uu^* = I$.*
— *u est autoadjoint ou hermitien si $u = u^*$.*

Exemple.
$E = L^2_\mu(X)$, $u : E \to E\ f \to hf$ où h est mesurable bornée.

$$\int_X (hf)\overline{g}d\mu = \int_X f\overline{\overline{h}g}d\mu \Rightarrow u^*(g) = \overline{h}g$$

donc $uu^* = u^*u$,u est donc un opérateur normal. Si h est rélle u est autoadjoint.
Remarque.
Si u est autoadjoint $\phi_u : (x,y) \to (ux|y)$ est sesquilinéaire hermitienne (la réciproque est vraie), c'est pourquoi on dit aussi que u est hermitien.

Théoreme 2.23 —

$$\frac{1}{2}||u|| \leq \sup_{||x|| \leq 1} |(ux|x)| \leq ||u||.$$

— *si u est hermitien $||u|| = \sup_{||x||\leq 1} |(ux|x)|$.*

Démonstration

$$||u|| = ||\phi_u|| \Rightarrow ||u|| = \sup_{||x||\leq 1} ||ux|| = \sup_{||x||\leq 1,\ ||y||\leq 1} |(ux|y)| \geq \sup_{||x||\leq 1} |(ux|x)|.$$

Posons $\alpha = \sup_{||x||\leq 1} |(ux|x)|$, on a $\forall x \in E$, $|(ux|x)| \leq \alpha ||x||^2$ $||u|| = \sup_{||x||\leq 1, ||y||\leq 1} |(ux|y)|$, en utilisant 2.1 pour la forme sesquilinéaire ϕ_u on obtient

$$\begin{aligned} ||u|| &\leq \frac{1}{4}\left(\alpha||x+y||^2 + \alpha||x-y||^2 + \alpha||x+iy||^2 + \alpha||x-iy||^2\right) \\ &= \frac{\alpha}{4}\left(2(||x||^2 + ||y||^2) + 2(||x||^2 + ||y||^2)\right) \leq 2\alpha. \end{aligned}$$

on a donc $(1/2)||u|| \leq \sup_{||x||\leq 1} |(ux|x)|$.
Supposons u hermitienne. Pour $z \in \mathbb{C}$

$$\begin{aligned} |z| = \sup_{0\leq\theta\leq 2\pi} |\Re(e^{i\theta} z)| &\Rightarrow ||u|| = \sup_{||x||\leq 1, ||y||\leq 1} |(ux|y)| = \sup_{||x||\leq 1, ||y||\leq 1, 0\leq\theta\leq 2\pi} \left|\Re(e^{i\theta}(ux|y))\right| \\ &= \sup_{||x'||\leq 1, ||y||\leq 1} |\Re(u(x')|y)| \leq \frac{1}{4} \sup_{||x'||\leq 1, ||y||\leq 1} \left(\alpha||x'+y||^2 + \alpha||x'-y||^2\right) \\ &\leq \frac{\alpha}{4} \sup_{||x'||\leq 1, ||y||\leq 1} (2(||x'||^2 + ||y||^2)) = \alpha \end{aligned}$$

on a donc $||u|| \leq \alpha$ et finalement $||u|| = \sup_{||x||\leq 1} |(ux|x)|$.

Corollaire 2.23.1 *Si $u \in \mathcal{L}(E)$ u^*u est autoadjoint et $||u^*u|| = ||u||^2$.*

Démonstration
$(u^*u)^* = u^*u$, il est donc bien autoadjoint.

$$||u^*u|| = \sup_{||x||\leq 1} |(u^*ux|x)| = \sup_{||x||\leq 1} (ux|ux) = \sup_{||x||\leq 1} ||ux||^2 = ||u||^2.$$

La propriété 4 est donc démontrée.

2.8 Eléments propres d'un opérateur hermitien

Si $\lambda \in \mathbb{C}$, $E_\lambda(u) = \ker(u - \lambda I) = \{x \in E | ux = \lambda x\}$ est le sous espace propre pour la valeur propre λ si $E_\lambda(u)$ n'est pas réduit au vecteur nul. Dans ce cas un élément $x \in E_\lambda(u)$ est appelé vecteur propre associé à u (si $ux = \lambda x$ alors $|\lambda| \leq ||u||$). En dimension finie l'ensemble Λ des valeurs propres de u est constitué des racines de dét$(u - \lambda I)$, il est non vide (le corps est $\mathbb{C}$) mais E n'est somme des $E_\lambda(u)$ que si u est diagonalisable.
En dimension infinie plus rien de cela ne subsiste, Λ peut être infini

$$(x_i)_{i\in\mathbb{N}} \to (\lambda_i x_i)_{i\in\mathbb{N}} \text{ de } l^2 \text{ dans lui même } \sup_{i\in\mathbb{N}} |\lambda_i| < +\infty$$

admet tous les λ_i comme valeurs propres, mais Λ peut être vide

$$E = L^2_m([0,1])\ ,\ u : f \to hf \ h \text{ mesurable bornée } (u - \lambda I)f = 0 \Leftrightarrow hf = \lambda f$$

ceci signifie $f = 0$ p.p. dans $K = \{x \in [0,1] | h(x) \neq \lambda\}$, si $\overline{K}$ est négligeable $f = 0$ p.p. or si $h(x) = x$ $\{x \in [0,1] | h(x) = \lambda\}$ est $\emptyset$ si $\lambda \notin [0,1]$ et $\{\lambda\}$ si $\lambda \in [0,1]$, u n'a donc pas de valeur propre.

Théoreme 2.24 *Les valeurs propres d'un opérateur hermitien sont réelles et les sous espaces propres son deux à deux othogonaux.*

Démonstration
Si $ux = \lambda x$ $(ux|x) = \lambda ||x||^2 = (x|ux) \in I\!R \Rightarrow \lambda \in I\!R$.
Soient x et y deux vecteurs propres associés respectivement aux valeurs propres distinctes λ et μ.

$$\lambda(x|y) = (\lambda x|y) = (ux|y) = (x|uy) = \mu(x|y) \text{ d'où } (x|y) = 0$$

. On va terminer en montrant que sous certaines un opérateur hermitien est diagonalisable.

Définition 2.9 *L'opérateur u sur E est dit compact si $\overline{u(B)}$ est compacte, où B est la boule unité fermée de E. Il revient au même de dire que pour toute suite (x_n) bornée, on peut extraire une sous suite (x_{n_k})telle que $(u(x_{n_k}))$ converge.*

Théoreme 2.25 *Soit u un opérateur hermitien compact sur E. Pour $\lambda \neq 0$ $E_\lambda(u)$ est de dimension finie. L'ensemble Λ des valeurs propres de u est finie ou peut être rangé en une suite tendant vers zéro. De plus E est l'adhérence du sous espace engendré par tous les $E_\lambda(u)$, $\lambda \in \Lambda$.*

Démonstration
Si Λ est infini, soit $\delta > 0$. Supposons qu'il existe une infinité de valeurs propres telles que $|\lambda| \geq \delta$ $(\delta \leq \lambda \leq ||u||)$, il existe donc une suite de ces valeurs propres convergeant vers convergeant vers $\lambda \neq 0$ car la couronne $\{z \in \mathbb{C} | \delta \leq |z| \leq ||u||$ est compacte dans $\mathbb{C}$.. Soit x_n un vecteur propre associé à la valeur propre λ_n, les x_n son linéairement indépendants, on appelle E_n le sous espace engendré par $x_0, x_1, \cdots, x_n$. On a $u(E_n) \subset E_n$ et $(u - \lambda_n I)E_n \subset En - 1$ $n \geq 1$, E_{n-1} étant sous espace strict de E_n il existe $y_n \in E_n$, $y_n \perp E_{n-1}$, $||y_n|| = 1$.

$$||\frac{y_n}{\lambda_n}|| = \frac{1}{|\lambda_n|} \leq \frac{1}{\delta},$$

u étant compact on peut extraire une sous suite pour laquelle la suite $(u(y_n/\lambda_n))$ converge. Ceci est contradictoire car

$$n < m \text{ , } u(\frac{y_n}{\lambda_n}) - u(\frac{y_m}{\lambda_m}) = y_n + \underbrace{\frac{1}{\lambda_n}(u - \lambda_n I)y_n - u(\frac{y_m}{\lambda_m})}_{\in E_{n-1}}$$

cet élément est donc de norme ≥ 1 et ceci démontre que pour $\delta > 0$ il y a un nombre fini de valeurs propres de module $\geq \lambda$.
L'ensemble des valeurs s'il est infini est donc dénombrable car réunion dénombrable d'ensembles finis et de plus d'après ce qui précède $\Lambda = (\lambda_n)$ peut être rangé

en une suite tendant vers zéro. Si $E_\lambda(u)$ était de dimension infinie, on pourrait choisir un système orthonormé dénombrable $(e_n)_{n\in I\!N}$ dans $E_\lambda(u)$, mais $u(e_n) = \lambda e_n$ et $||e_n - e_m|| = \sqrt{2}$, ceci contredit la compacité de u, $E_\lambda(u)$ est donc de dimension finie. Montrons qu'il existe une valeur propre λ telle que $|\lambda| = ||u||$.
D'après le théorème 2.23 il existe une suite (x_n) d'éléments de E telle que

$$\lim_{n\to\infty} |(ux_n|x_n)| = ||u||$$

il existe une sous suite de (x_n) telle que $(ux_n|x_n)$ converge vers $\alpha = ||u||$ ou $-||u||$ car $(ux_n|x_n) \in I\!R$, u étant hermitien.

$$0 \le ||ux_n - \alpha x_n|| - ||ux_n||^2 - 2\alpha(ux_n|x_n) + \alpha^2||x_n||^2 \to_{n\to\infty} 0$$

u étant compact il existe une sous suite de (x_n) convergeant vers un élément $y \in E$ et ceci entraine pour cette même sous suite $\lim_{n\to\infty} \alpha x_n = y$, si $u \neq 0$ $\alpha \neq 0$ et $\lim_{n\to\infty} x_n = y/\alpha = x$ et $ux = \alpha x$, α est donc valeur propre pour u.
Soit $F = \overline{\cup_{\lambda\in\Lambda} E_\lambda}$ et $H = F^\perp$. H est u-invariant $(u(E) \subset E)$ et u_H est un opérateur hermitien compact sans valeur propre, d'après ce qui précède ceci n'est possible que si $H = \{0\}$, on a donc $E = F \oplus H = F$.
Après avoir choisi un système orthonormé dans chaque $E_\lambda(u)$ on obtient par leur réunion un système orthonormé complet dans E constitué de vecteurs de u et quel que soit $\delta > 0$ il y a un nombre fini de λ_n, valeurs propres de u, de module supérieur ou égal à δ.

Deuxième partie

PROBABILITES

Chapitre 3

NOTIONS FONDAMENTALES

3.1 ENSEMBLE FONDAMENTAL. EVENEMENTS

3.1.1 Exemples

Exemple 1

On lance un dé à 6 faces numérotées de 1 à 6. On cherche à décrire cette expérience aléatoire, c'est à dire due au hasard. Chaque réalisation possible sera notée sous forme de singleton, $\{1\}, \{2\}, \cdots, \{6\}$. L'ensemble des réalisations sera notée $\Omega = \{1, 2, 3, 4, 5, 6\}$ et appelé ensemble des états ou ensemble fondamental. Un observateur peut s'intéresser à un aspect partiel de l'expérience, par exemple au cas où le nombre qui sort est pair. On dira que l'évènement "nombre pair" est réalisé si le nombre obtenu est élément du sous ensemble $\{2, 4, 6\}$. Les évènements seront décrits par des sous ensembles de Ω. Les réalisations seront appelées évènements élémentaires ou éventualitées.

Exemple 2

On lance 2 pièces de monnaie, les éventualités pourront être notées $(P, P), (P, F), (F, P), (F, F)$ si les 2 pièces peuvent être distinguées. L'évènement "On obtient au moins un pile" sera représenté par le sous ensemble $\{(P, F), (F, P), (P, P)\}$.

Exemple 3

On lance une pièce de monnaie et on s'arrête lorsqu'on obtient pile pour la première fois. On peut choisir
$\Omega = \{P, FP, FFP, \cdots, FF\cdots FP, \cdots\}$
c'est un ensemble infini. L'évènement "on a joué au plus 3 fois" s'écrit
$\{P, FP, FFP\}$.

Exemple 4

La physique conduit à considérer dans certains cas le déplacement d'une particule comme un phénomène aléatoire. Si le déplacement se fait dans le plan, l'ensemble Ω des positions possibles de la particule pourra être le plan ou une partie de celui-ci. Les évènements seront des parties du plan, mais pour des raisons mathématiques il s'agira d'un sous ensemble strict de l'ensemble des parties du plan. L'ensemble des évènements sera une tribu de parties de $I\!R^2$, par exemple $\mathcal{B}(I\!R^2)$ qui est la tribu borélienne de $I\!R^2$ ou encore la plus petite tribu contenant les ouverts de $I\!R^2$.

3.1.2 Définition

Définition 3.1 *Le couple $(\Omega, \mathcal{E})$ constitué par un ensemble Ω et une tribu de parties $\mathcal{E}$ de Ω est appelé espace probabilisable. Les éléments de Ω sont appelés éventualités ou évènements élémentaires. Les éléments de $\mathcal{E}$ sont appelés évènements.*

3.2 PROBABILITE. ESPACE PROBABILISE

3.2.1 Définition

Définition 3.2 *Une probabilité P sur un espace probalisable $(\Omega, \mathcal{E})$ est une mesure sur cet espace (qui est mesurable) telle que $P(\Omega) = 1$.*

<u>Exemples</u>
<u>L'ensemble Ω est fini ou infini dénombrable.</u>
Dans ce cas on choisit en général $\mathcal{E} = \mathcal{P}(\Omega)$. Une probabilité sur Ω est définie par la donnée d'une famille
$\{P(x) = \alpha_x, x \in \Omega\}$ de nombres réels vérifiant la condition
$\sum_{x\in\Omega} \alpha_x = 1$.
<u>Dans le cas particulier où Ω est fini</u> on appelle probalité uniforme sur Ω la probabilité correspondant à $\alpha_x = 1/Card\Omega$.
Pour $A \subset \Omega, P(A) = CardA/Card\Omega$. cette probabilité correspond au tirage d'un point au hasard dans Ω.
<u>Exemple</u>
On lance n fois une pièce équilibrée. On prend
$\Omega = \{P, F\}^n,\ \alpha_x = 1/2^n$. L'évènement A="On obtient 3 piles en n lancers " a pour probabilité $P(A) = C_n^3/2^n$ pour $n \geq 3$.
<u>L'ensemble $(\Omega, \mathcal{E})$ est $(I\!R, \mathcal{B}(I\!R))$</u>
On considère une fonction mesurable positive h vérifiant $\int_{-\infty}^{+\infty} h(x)dx = 1$. Pour tout élément $A \in \mathcal{B}(I\!R)$ on définit une probabilité par $P(A) = \int_A h(x)dx$.

3.2.2 Propriétés des probabilités

$.\forall A \in \mathcal{E}\ P(\overline{A}) = 1 - P(A)$.
$.P(\emptyset) = 0$.
$.\forall A, B \in \mathcal{E}\ A \subset B \Rightarrow P(A) \leq P(B)$.
$.\forall A, B \in \mathcal{E}\ P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
Pour démontrer ce dernier résultat il suffit d'écrire
$A \cup B = (A - A \cap B) \cup (A \cap B) \cup (B - A \cap B)$ et de remarquer que si $D, C \in \mathcal{E}, C \subset D$ on a $P(C - D) = P(C) - P(D)$. Ce résultat admet la généralisation suivante, dont la démonstration se fait par récurrence.

Théoreme 3.1 *Pour n évènements $A_1, A_2, \cdots, A_n$ on a*

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} S_n^k$$

avec

$$S_n^k = \sum_{1 \leq i_1 < i_2 < \cdots \leq i_n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})$$

Exemple.
On dispose de 2 jeux de 52 cartes et on tire une carte de chaque jeu. Quelle est la probabilité d'avoir tiré au moins un roi ?

On choisit $\Omega = \{(a, b) | a\ cartedupremierjeu,\ b\ cartedudeuxiemejeu\}$, $\mathcal{E} = \mathcal{P}(\Omega)$ et $P =$ probabilité uniforme sur Ω.

Première méthode
Définissons les évènements suivants :
$A_1 =$ "On tire un roi du premier jeu et une carte autre qu'un roi du deuxième jeu".
$A_2 =$ "On tire un roi du deuxième jeu et une carte autre qu'un roi du premier jeu".
$A_3 =$ "On tire un roi du premier jeu et un roi du deuxième jeu".

$P(A_1) = P(A_2) = (4.48)/52^2 = 192/2704$ $P(A_3) = 16/2704$. $A =$ "au moins un roi" $= A_1 \cup A_2 \cup A_3$, ces 3 évenements étant 2 à 2 disjoints on a $P(A) = 400/2704$.

Deuxième méthode

$P(\overline{A}) = P($ "aucun roi"$)=48^2/2704 = 2304/2704$
$P(A) = 1 - P(\overline{A}) = 400/2704$.

3.3 PROBABILITE CONDITIONNELLE

3.3.1 Exemple.

Une entreprise achète 1000 pièces détachées à un fournisseur A_1 dont 5 pour cent sont défectueuses et 4000 à un fournisseur A_2 dont 3 pour cent sont défectueuses. Une piéce choisise au hasard est défectueuse, quelle est la probabilité qu'elle provienne du fournisseur A_1 ?

Définissons les évènements suivants :
$\Omega =$ "ensemble des pièces détachées".
$D =$ "ensemble des pièces défectueuses".
$A_i =$ "ensemble des pièces provenant du fournisseur A_i, $i = 1, 2$.

L'ensemble D devient l'évènement certain.
$Card\ D = 0,05.1000 + 0,03.4000 = 170$. A_1 contenant 50 pièces défectueuses, la probabilité cherchée est
$50/170 = \frac{50/5000}{170/5000} = P(A_1 \cap D)/P(D)$.

3.3.2 Définition.

Définition 3.3 *Soit* $(\Omega, \mathcal{E}, P)$ *un espace probabilisé et* A *un évènement de probabilité non nulle,* $F \in \mathcal{E}$*, on appelle probabilité conditionnellede* F *par rapport à* A *ou probabilité conditionnelle de* F *sachant* A *le nombre*

$$P(F|A) = \frac{P(F \cap A)}{P(A)}.$$

Conséquence. Sous les hypothèses précédentes on a $P(F \cap A) = P(F|A)P(A)$. Ce résultat peut être généralisé par récurrence à n évènements $A_1, A_2, \cdots, A_n$.

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap A_2 \cap A_{n-1}).$$

Exemples.
On tire successivement et sans remise 3 cartes d'un jeu de 52 cartes. Quelle est la probabilité d'avoir tiré 3 rois ?
Posons $A_i =$ "on tire un roi au ième tirage", $i = 1, 2, 3$.
La probabilité cherchée est

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) = 4/52.3/51.2/50.$$

On considère la composition en garçons et filles d'une famille de 2 enfants. Un ensemble fondamental est
$\Omega = \{(G, G), (G, F), (F, G), (F, F)\}$. Calculons la probabilité des évènements suivants :
$H =$ "la famille a un garçon".
$A =$ "l'aîné est un garçon".
$B =$ "les 2 enfants sont des garçons".

$$P(B|H) = \frac{P(B \cap H)}{P(H)} = 1/3\ ,\ P(B|A) = \frac{P(B \cap A)}{P(A)} = 1/2.$$

Cet exemple montre l'importance de bien poser le problème et de se méfier de l'intuition dans les calculs de probabilité conditionnelle.

3.3.3 Théorème de Bayes

Théoreme 3.2 *On considère* $(A_1, A_2, \cdots, A_n)$ n *évènements constituant une partition de* Ω *tels que* $P(A_i) > 0$, $i = 1, 2, \cdots, n$*. Pour tout évènement* A *on a*

$$P(A) = \sum_{i=1}^{n} P(A|A_i)P(A_i) \ \ \textit{et si} \ \ P(A) \neq 0$$

$$P(A_j|A) = \frac{P(A|A_j)P(A_j)}{\sum_{i=1}^{n} P(A|A_i)P(A_i)}.$$

Ce résultat est immédiat en remarquant que
$A = \cup_{i=1}^{n}(A \cap A_i)$ et que ces évènements sont deux à deux disjoints.
Exemple.
Une urne A_1 contient 70 pour cent de boules blanches et une urne A_2 80 pour

cent de boules blanches. L'urne A_1 contient trois fois plus de boules que l'urne A_2. On mélange les contenus de A_1 et de A_2 et on choisit une boule au hasard, elle est blanche, quelle est la probabilité qu'elle provienne de A_1 ?

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} = \frac{0,7.0,75}{0,7.0,75 + 0,8.0,25} = 21/29.$$

3.3.4 Evènements indépendants.

Définition 3.4 *Soit $(\Omega, \mathcal{E}, P)$ un espace probabilisé et $A, B \in \mathcal{E}$. Ces 2 évènements sont indépendants si $P(A \cap B) = P(A)P(B)$.*

<u>Conséquence.</u>
Dans le cas où ces probabilités sont définies on a $P(A|B) = P(A)$, $P(B|A) = P(A)$ quand A et B sont indépendants. Autrement dit la réalisation de l'un des évènements n'influe pas sur la probabilité qu'a l'autre de se réaliser.

Définition 3.5 *Deux tribus $\mathcal{A}$ et $\mathcal{B}$, sous ensembles de $\mathcal{E}$, sont indépendantes si*

$$\forall A \in \mathcal{A} \ , \ \forall B \in \mathcal{B} \ P(A \cap B) = P(A)P(B)$$

Théoreme 3.3 *L'indépendance des évènements A et B entraîne l'indépendance de $\overline{A}$ et B, de A et $\overline{B}$, de $\overline{A}$ et $\overline{B}$, c'est à dire des tribus $\mathcal{A} = \{\emptyset, A, \overline{A}, \Omega\}$ et $\mathcal{B} = \{\emptyset, B, \overline{B}, \Omega\}$ engendrées par A et B respectivement.*

Montrons l'indépendance de $\overline{A}$ et B, les autres cas se traitant de la même manière.
$(\overline{A} \cup A) \cap B = (\overline{A} \cap B) \cup (A \cap B) = B$.
$P(B) = P(\overline{A} \cap B) + P(A \cap B)$.
$P(\overline{A} \cap B) = P(B) - P(A \cap B) = p(B) - p(A)P(B)$.
$P(\overline{A} \cap B) = P(B)(1 - P(A)) = P(B)P(\overline{A})$.

Définition 3.6 *Les évènements $A_1, A_2, \cdots, A_n$ sont indépendants si les tribus engendrées respectivement par $A_1, A_2, \cdots, A_n$ sont indépendantes. Il revient au même de dire*

$$\forall k \in \{1, 2, \cdots, n\} \ , \ P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j})$$

pour toute suite $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.

Chapitre 4

VARIABLES ALEATOIRES DISCRETES

4.1 LOI DE PROBABILITE

4.1.1 Mesure image et mesure définie par une densité

Mesure image.

Définition 4.1 *On considère un espace mesuré* $(X, \mathcal{A}, \mu)$ *et un espace mesurable* $(Y, \mathcal{B})$*, h étant une application mesurable de* X *vers* Y*, on appelle mesure image de* μ *par* h *la mesure notée* $\nu = h(\mu)$ *définie sur* $(Y, \mathcal{B})$ *par* $\nu(B) = \mu(h^{-1}(B))$*,* $B \in \mathcal{B}$*.*

Théoreme 4.1 *Pour toute fonction* f *positive et* ν*-mesurable sur* Y *on a*

$$\int_Y f d\nu = \int_X (f \circ h) d\mu$$

<u>Démonstration</u>
Le théorème est vrai pour les fonctions indicatrices d'ensembles χ_B avec $B \in \mathcal{B}$, donc également pour toute fonction étagée, puis pour pour toute fonction mesurable positive en utilisant une suite croissante de fonctions étagées convergeant vers cette fonction.

Corollaire 4.1.1 *Une fonction* f *réelle ou complexe* $\mathcal{B}$*-mesurable sur* Y *est* ν*-intégrable si et seulement si* $f \circ h$ *est* μ*-intégrable et dans ce cas*

$$\int_Y f d\nu = \int_X (f \circ h) d\mu.$$

<u>Démonstration</u>
Le résultat est vrai pour les fonctions positives d'après le théorème précédent, si f est à valeurs dans $\mathbb{R}$ on applique le résultat à f^+ et f^-, si f est à valeurs dans $\mathbb{C}$ on utilise Ref et Imf.

Mesures de base donnée

Définition 4.2 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et g une fonction mesurable positive sur X. On appelle mesure de base (ou de poids) g par rapport à la mesure μ, la fonction d'ensembles définie sur $\mathcal{A}$ par $\nu(E) = \int_E g d\mu$, $E \in \mathcal{A}$.*

La fonction ν est bien une mesure car si les ensembles mesurables A_n sont deux à deux disjoints on obtient

$$\nu(\cup_n E_n) = \int (\chi_{\cup_n E_n}) g d\mu = \int \left(\sum_n \chi_{E_n} \right) g d\mu = \sum_n \int_{E_n} g d\mu = \sum_n \nu(E_n)$$

L'avant dernière égalité est une conséquence du théorème de Beppo-Levi. La mesure ν est notée $\nu = g\mu$.

Théoreme 4.2 *On a pour toute fonction mesurable positive f*

$$\int f d\nu = \int f g d\mu$$

où ν est la mesure de base g par rapport à μ.

La démonstration est similaire à celle du théorème précédent.

Corollaire 4.2.1 *La fonction f est ν-intégrable si et seulement si fg est μ-intégrable et*

$$\int f d\nu = \int f g d\mu$$

Remarque.
Dans le cas d'un ensemble mesurable E tel que $\mu(E) = 0$ on a alors $\nu(E) = 0$. On exprime ceci en disant que ν est absolument continue par rapport à μ.

4.1.2 Variable aléatoire

Définition 4.3 *Soit $(\Omega, \mathcal{E}, P)$ un espace probabilisé et $(E, \mathcal{B})$ un espace mesurable. Une application mesurable $X : \Omega \to E$ est appelée varible aléatoire. Si E est dénombrable (en général $E = \mathbb{Z}$ ou $E = \mathbb{N}$) on dit que X est une variable aléatoire discrète. Dans ce cas on prend en général $\mathcal{B} = \mathcal{P}(E)$.*

Notations.
$I \subset E$, $X^{-1}(I) = \{\omega \in \Omega | X(\omega) \in I\}$ est noté $[X \in I]$
$n \in E$, $X^{-1}(\{n\}) = \{\omega \in \Omega | X(\omega) = n\}$ est noté $[X = n]$.

4.1.3 Loi de probabilité.

Définition 4.4 *La loi de probabilité de la variable aléatoire discrète X est la mesure image de P par X. C'est une probabilité sur $(E, \mathcal{B})$ notée p_X. L'espace $(E, \mathcal{B}, p_X)$ est un espace probabilisé.*

Remarque.
$B \in \mathcal{B}$, $p_X(B) = P(X^{-1}(B)) = P([X \in B])$.
Dans le cas où $\mathcal{B} = \mathcal{P}(E)$ et $n \in E$, $p_X(\{n\})$ est noté $p_X(n) = P([X = n])$.
Dans ce cas
$B \subset E$, $p_X(B) = \sum_{n \in B} p_X(n)$.

Exemple.
On lance 3 dés de couleurs différentes et on appelle X le nombre de 6 obtenus. Modélisons cette expérience.

$$\Omega = \{1, 2, 3, 4, 5, 6\}^3 \ , \ \mathcal{E} = \mathcal{P}(\Omega) \ , \ P = \text{probabilité uniforme sur } \Omega$$

$$X((a, b, c)) ='' \text{nombre de six parmi } a, b, c'' \ , \ E = \{0, 1, 2, 3\} \ , \ \mathcal{B} = \mathcal{P}(E)$$

$$p_X(0) = 125/216 \ , \ p_X(1) = 75/216 \ , \ p_X(2) = 15/216 \ , \ p_X(3) = 1/216.$$

4.1.4 Fonction de répartition.

Définition 4.5 *La fonction de répartition de la variable aléatoire X est la fonction définie par $F_X(x) = P([X < x]) = P(\{\omega \in \Omega | X(\omega) < x\})$.*

Propriétés.

- $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to +\infty} F_X(x) = 1$.
- F_X est croissante.
- $\cup_{n=1}^{\infty} [X < x - 1/n] = [X < x]$ entraine
  $\lim_{n \to \infty} F_X(x - 1/n) = F_X(x)$, la fonction F_X est donc continue à gauche en tout point.
- En faisant un raisonnement analogue au précédent on obtient
  $\lim_{n \to \infty} F_X(x + 1/n) = F_X(x) + P(\{x\})$. L fonction F_X est continue à droite en x si et seulement si $P(\{x\}) = 0$.

Exercice.
Représenter graphiquement la fonction de répartition de la variable aléatoire définie dans l'exemple précédent.

4.2 ESPERANCE MATHEMATIQUE. VARIANCE.

4.2.1 Espérance mathematique.

Définition 4.6 *L'espérance mathématique de la variable aléatoire X est le nombre $E(X) = \int_E X dP$ si X est intégrable.*

Conséquence.
p_X étant la loi image de P par X on a $E(X) = \int_E x dp_X(x)$. Si X est une variable aléatoire discrète, à valeurs dans $\mathbb{Z}$ ou $\mathbb{N}$ on a $E(X) = \sum_{x \in E} x p_X(x)$, ceci correspond à la moyenne pondérée (ou barycentre) des valeurs prises par X.

4.2.2 Fonction génératrice.

Nous allons dans ce paragraphe introduire un outil qui facilitera le calcul de l'espérance mathématique et de la variance pour une variable aléatoire discrète.

Définition 4.7 *La variable aléatoire X étant à valeurs dans $\mathbb{N}$, la fonction génératrice de X est la série entière*

$$G_X(z) = \sum_{k \in E} p_X(k) z^k$$

le rayon de convergence de cette série est supérieur ou égal à 1 car $G_X(1) = 1$. Si $R > 1$ $G'_X(1) = \sum_{k \in E} k p_X(k) = E(X)$.

Exemple.
$E = \mathbb{N}^*$, $p_X(k) = q^{k-1} p$, $(p + q = 1)$.

$$G_X(z) = \sum_{k=1}^{\infty} q^{k-1} p z^k = pz \sum_{k=1}^{\infty} (qz)^{k-1} = \frac{pz}{1 - qz}$$

$E(X) = G'_X(1) = \frac{p}{(1-q)^2}$.

4.2.3 Propriétés de l'espérance mathématique.

— Si X et Y sont deux variables aléatoires on a

$$E(X + Y) = E(X) + E(Y) \ , \ E(aX) = aE(X) \ (a \in \mathbb{R})$$

— E et F étant $\mathbb{Z}$ ou $\mathbb{N}$, $g : E \to F$ une application et X une variable aléatoire discrète on a

$$E(g \circ X) = E(g(X)) = \sum_{k \in E} g(k) p_X(k)$$

Le premier résultat provient de la linéarité de l'intégrale et le deuxième des égalités

$$E(g(X)) = \int_\Omega g(X) dP = \int_E g(x) dp_X(x)$$

car p_X est la mesure image de P par X.
Exemples.
$E(X^n) = \sum_{k \in E} k^n p_X(k)$.
$E(e^X) = \sum_{k \in E} e^k p_X(k)$.

4.2.4 Variance.

Définition 4.8 *Une variable aléatoire X vérifiant $E(X^2) < +\infty$ on définit la variance de X par $Var(X) = E((X - E(X))^2)$. Cette quantité mesure la dispersion (en fait le carré de la dispersion) moyenne des valeurs de X par rapport à l'espérance mathématique $E(X)$. L'écart type de X est la quantité $\sigma(X) = \sqrt{var(X)}$.*

La fonction génératrice peut permettre de calculer la variance car
$\phi''_X(1) = \sum_{k\geq 2} k(k-1) = E(X(X-1)) = E(X^2) - E(X)$. Ceci suppose que le rayon de convergence R soit supérieur strictement à 1.
Covariance de deux variables aléatoires.

Définition 4.9 $X, Y \in L^2(\Omega, \mathcal{E}, P) \Rightarrow XY \in L^1(\Omega, \mathcal{E}, P)$. *Ceci permet de définir la covariance des variables aléatoires* x *et* Y *par*
$Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY)\text{-}E(X)E(Y)$.

L'inégalité de Cauchy-Schwarz entraine
$|\mathrm{Cov}(X,Y)| \leq \sigma(X)\sigma(Y)$.

Définition 4.10 *Le coefficient de corrélation de deux variables aléatoires* X *et* Y *qui ne sont pas presque surement constantes est*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \ , \ |\rho(X,Y)| \leq 1$$

Ce coefficient mesure le degré de dépendance linéaire de X et Y.

Théoreme 4.3 *On considère deux variables aléatoires* X *et* Y *telles que* $E(X^2) < +\infty$ *et* $E(Y^2) < +\infty$.

1. $Var(X) = 0 \Leftrightarrow X$ *est constante* P *presque surement.*
2. $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$.
3. $\rho(X,Y) = \pm 1 \Leftrightarrow (\exists (a,b,c) \in I\!R^3 - (0,0,0) | aX + bY + c = 0 \ P \ p.s.)$.

4.3 Lois de probabilités discrètes usuelles

4.3.1 Loi de Bernoulli.

Une épreuve de Bernoulli est une expérience aléatoire à deux issues, par exemple jeu de pile ou face, tirer un roi ou une autre carte dans un jeu de carte... Ces deux possibilités sont en général notées 0 et 1 et appelées échec ou succès. La variable aléatoire $X : \Omega \to \{0,1\}$ est appelée variable de Bernoulli. Sa loi est définie par $p = P([X=0])$, $q = 1 - p = P([X=1])$. On dit que p est le paramètre associé à X.
$E(X) = p$, $Var(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p) = pq$.

4.3.2 Loi Binomiale.

On répète n épreuves de Bernoulli indépendantes de paramètre p, X est le nombre de 1 obtenus ($0 \leq X \leq n$). On appelle A l'évènement "On obtient 1 ". La loi suivie par X est appelée loi binomiale de paramètres n , p et notée $B(n,p)$.
L'écriture $X \sim B(n,p)$ signifie que X suit la loi binomiale $B(n,p)$.

$p_X(k) = P([X = k]) = P($"réaliser k fois $A) = C_n^k p^k q^{n-k}$
car les répétitions sont indépendantes. $k \in \{0, 1, \cdots, n\}$.

$\phi_X(z) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} z^k = (pz + q)^n$.
$E(X) = \phi'_X(1) = np$, $E(X^2) = \phi''_X(1) + E(X) = n(n-1)p^2 + np$,
$Var(X) = E(X^2) - (E(X))^2 = np(1-p) = npq$.

4.3.3 Loi géométrique.

On répète une épreuve de Bernoulli, de paramètre p et on s'arrête dès que l'on obtient 1. On appelle X le nombre de répétitions nécessaires. La loi de X est appelée loi géométrique de paramètre p. Les valeurs prises par X sont les éléments de $\mathbb{N}^*$.

$k \in \mathbb{N}^*$, $p_X(k) = q^{k-1}p$
car ceci revient à faire $k-1$ répétitions de $\overline{A}$ la k-ième répétitions donnant A.

On a montré dans un exemple précédent que
$\phi_X(z) = pz/(1-qz)$ d'où
$\phi'_X(z) = p/(1-qz)^2$, $\phi''_X(z) = 2pq/(1-qz)^3$.
$E(X) = \phi'_X(1) = 1/p$, $E(X^2) = \phi''_X(1) + E(X) = 2q/p^2 + 1/p$
$Var(X) = q/p^2$.

4.3.4 Loi hypergéométrique.

On choisit n objets parmi $N \geq n$ objets de deux types notés I et II. X représente le nombre d'objets de type I parmi les n objets. Il existe d objets de type I.

$$p_X(k) = \frac{C_d^k C_{N-d}^{n-k}}{C_N^n} \ , \ \max(0, n-(N-d)) \leq k \leq \min(n, d)$$

4.3.5 Loi de Poisson.

C'est une loi utile pour l'étude des phénomènes de files d'attente. L'ensemble des valeurs prises par X est $\mathbb{N}$ et

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!} \ , \ k \in \mathbb{N} \ , \ \lambda > 0$$

$\phi_X(z) = e^{-\lambda(z-1)}$ d'où $E(X) = Var(X) = \lambda$.

4.4 COUPLE DE VARIABLES ALEATOIRES

Les variables aléatoires discrètes X et Y étant définies sur $(\Omega, \mathcal{E}, P)$ et à valeurs dans les ensembles dénombrables E et F respectivement, on définit une variable aléatoire discrète $Z = (X, Y)$ à valeurs dans $E \times F$ en posant $Z(\omega) = (X(\omega), Y(\omega))$. On a en effet
$\{\omega \in \Omega | Z(\omega) = (x, y)\} = \{\omega \in \Omega | X(\omega) = x\} \cap \{\omega \in \Omega | Y(\omega) = y\} \in \mathcal{E}$.

Définition 4.11 *La loi de probabilité de Z est appelée loi conjointe du couple (X,Y) et les lois des variables X et Y sont appelées lois marginales du couple de variables aléatoires.*

Il est possible de connaitre les lois marginales lorsque l'on connait la loi conjointe.

Théoreme 4.4

$$p_X(x) = \sum_{y \in F} p_Z(x,y) \ , \ x \in E \ , \ p_Y(y) = \sum_{x \in E} p_Z(x,y) \ , \ y \in F$$

Ce résultat est obtenu en remarquant que

$$p_X(x) = p_Z(\{x\} \times F) = \sum_{x=x} \sum_{y \in F} p_Z(x,y) \ , \ p_Y(y) = p_Z(E \times \{y\}) = \sum_{x \in E} \sum_{y=y} p_Z(x,y)$$

X et Y étant deux variables aléatoires discrètes, $X^{-1}(\mathcal{P}(E))$ et $Y^{-1}(\mathcal{P}(F))$ sont deux sous tribus de $\mathcal{E}$ appelées respectivement tribus engendrées par X et par Y.

Définition 4.12 *Les variables aléatoires $X_1, X_2, \cdots, X_n$ sont indépendantes si les n tribus qu'elles engendrent sont indépendantes.*

Théoreme 4.5 *Les variables aléatoires discrètes X et Y sont indépendantes si et seulement si*

$$\forall (x,y) \in E \times F \ P([Z = (x,y)]) = P([X = x])P([Y = y]) \ , \ \textit{ou encore} \ p_Z(x,y) = p_X(x)p_Y(y)$$

La vérification de ce théorème est simple et laissée en exercice. Ce résultat exprime le fait que la mesure p_Z sur $E \times F$ est le produit des mesures p_X sur E et p_Y sur F.

Théoreme 4.6 *Les variables aléatoires X et Y étant indépendantes on a $E(XY) = E(X)E(Y)$ si X et Y sont intégrables.*

Les variables aléatoires étant indépendantes on a $p_Z = p_X \otimes p_Y$.
$E(XY) = \int_{E \times F} xy dp_X(x) dp_Y(y) = \int_E x dp_X(x) \int_F y dp_Y(y)$ d'après le théorème de Fubini. Ce résultat peut être redémontré directement en écrivant la première intégrale sous forme d'une série double.
Remarque.
La réciproque de ce théorème est fausse comme le montre l'exemple suivant.
$E = F = \{-1, 0, 1\}$, $p_Z(-1,0) = p_Z(1,0) = p_Z(0,-1) = p_Z(0,1) = 1/4$, les 5 autre valeurs étant nulles.
$E(XY) = \int_{x \in E} \int_{y \in F} xy dp_Z(x,y) = \sum_{x \in E} \sum_{y \in F} xy p_Z(x,y) = 0$ car dans chaque terme de la somme un des facteurs est nul.
Par symétrie les lois marginales de X et de Y sont identiques.
$p_X(-1) = p_Z(-1,-1) + p_Z(-1,0) + p_Z(-1,1) = 1/4$
$p_X(0) = p_Z(0,-1) + p_Z(0,0) + p_Z(0,1) = 1/2$
$p_X(1) = 1 - (1/4 + 1/2) = 1/4$.
$E(X) = -1/4 + 0 + 1/4 = E(Y) = 0$. On a bien
$E(XY) = E(X)E(Y)$ et pourtant $p_Z(0,0) = 0 \neq p_X(0)p_Y(0)$, ce qui montre

que X et Y ne sont pas indépendantes.
Remarque.
Si les variables aléatoires discrètes X et Y sont indépendandantes, il en est de même de $f \circ X$ et $g \circ Y$ où f et g sont des applications de E et F respectivement vers des ensembles dénombrables. Ceci provient de l'égalité $(f \circ X)^{-1}(A) = X^{-1}(f^{-1}(A))$ qui est donc élément de la tribu engendrée par X. En particulier, dans ce cas
$Cov(X, Y) = E((X - E(X))(Y_E(Y))) = E(X - E(X))E(Y - E(Y)) = 0$
d'après le théorème précédent. Deux variables aléatoires indépendantes ayant une covariance nulle (on dit qu'elles sont non correlées) on a aussi
$Var(X + Y) = Var(X) + Var(Y)$ quand elles sont indépendantes.

Théoreme 4.7 *Soit X et Y deux variables aléatoires discrètes indépendantes. On a la relation suivante entre fonctions génératrices*

$$\phi_{X+Y}(z) = \phi_X(z)\phi_Y(z)$$

La réciproque de ce théorème est fausse.
On a en effet $\phi_{X+Y}(z) = E(z^{X+Y}) = E(z^X)E(z^Y)$ d'après une remarque précédente concernant $f(X)$ et $g(Y)$, X et Y étant indépendants.
Exemple.
X et Y étant deux variables aléatoires indépendantes suivant une même loi de Poisson de paramètre λ, $\phi_{X+Y}(z) = e^{2\lambda(z-1)}$, la variable aléatoire $X + Y$ suit donc une loi de Poisson de paramètre 2λ car une fonction génératrice est caractéristique d'une loi.

Chapitre 5

VARIABLES ALEATOIRES CONTINUES

5.1 LOI DE PROBABILITE

5.1.1 Généralités.

Définition 5.1 *Une variable aléatoire continue X est une application mesurable d'un espace probabilisé $(\Omega, \mathcal{E}, P)$ vers $(I\!R, \mathcal{B}(I\!R))$ où $\mathcal{B}(I\!R)$ est la tribu des boréliens de $I\!R$. On appelle loi de probabilité de X la mesure image p_X de P par X. Cette mesure admet en général une densité g par rapport à la mesure de Lebesgue sur $I\!R$, appelée densité de X.*

Remarques.

- Si X admet une densité g on a alors
  $\forall A \in \mathcal{B}(I\!R)$, $p_X(A) = \int_A g(x)dx$.
- Si F_X est la fonction de répartition de X on a
  $F_X(t) = P([X < t]) = p_X(] - infty, t[) = \int_{-\infty}^{t} g(t)dt$ d'où $F'_X(t) = g(t)$ si g est continue.
- Une variable aléatoire continue X peut être définie à partir de sa fonction de répartition ou de sa densité, mais par contre $P([X = x]) = 0$ en général (c'est toujours le cas quand on a une densité continue).
- une fonction g est une densité si et seulement si
  $\forall x \in I\!R\ f(x) \geq 0$, $\int_{-\infty}^{+\infty} g(x)dx = 1$.

Les proprétés de la fonction de répartition sont les mêmes que dans le cas discret.

5.1.2 Espérance mathématique et variance

Les définitions des espérance mathématiques et variance sont identiques à celles présentées dans le chapitre précédent. Dans le cas continu on obtient pour une variable intégrable

$$E(X) = \int_{\Omega} X(\omega)dP(\omega) = \int_{I\!R} x dp_X(x) = \int_{I\!R} xg(x)dx$$

d'après les propriétés d'une mesure image et d'une mesure à densité.
Si $h$est une fonction réelle continue par morceaux on a de même

$$E(h \circ X) = E(h(X)) = \int_\Omega h \circ X(\omega) dP(\omega) = \int_{I\!R} h(x)g(x)dx$$

sous réserve que $h \circ X$ soit intégrable.

5.2 Lois continues usuelles

5.2.1 Loi Exponentielle.

On appelle loi exponentielle de paramètre λ la loi de densité

$$f(x) = \lambda e^{-\lambda x} \chi_{I\!R^+} \ , \ \lambda > 0$$

elle est notée Exp(λ).

$$E(X) = \int_{I\!R} x f(x) dx = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = 1/\lambda$$

$$E(X^2) = \int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx = 1/\lambda^2 \ , \ Var(X) = E(X^2) - (E(X))^2 = 1/\lambda^2$$

5.2.2 Loi uniforme sur $[a, b]$.

La densité de cette loi est

$$f(x) = \frac{1}{b-a} \chi_{[a,b]}(x) \ , \ P([a \leq X \leq b]) = \frac{d-c}{b-a} \ , \ a \leq c < d \leq b$$

autrement dit cette probabilité est proportionnelle à la longueur de l'intervalle et ne dépend pas de la position des points dans $[a, b]$.