Notes de cours de l'ISIMA, deuxième année, http://www.isima.fr/~leborgne

Opérateur à noyau intégral, espace de Hilbert à noyau reproduisant : introduction

Gilles Leborgne

31 janvier 2024

Table des matières

La notation $f := g$ signifie : f est définie par $f = g$, encore note $f =^{\text{déf}} g$.
La notation $f =^{\text{not.}} g$ signifie : f est également noté g.

Première partie
Cadre

1 L'espace $(L^2(\Omega), (\cdot,\cdot)_{L^2})$

Soit Ω un ouvert dans $\mathbb{R}^n$. On note $L^2(\Omega)$ l'ensemble des fonctions $\Omega \to \mathbb{R}$ qui sont de carré intégrable relativement à la mesure de Lebesgue (fonctions d'énergie finie) :

$$L^2(\Omega) = \{f : \Omega \to \mathbb{R} \text{ t.q. } \int_\Omega f(\vec{x})^2\,dx < \infty\}. \tag{1.1}$$

Voir cours d'intégration. Le produit scalaire usuel de $L^2(\Omega)$ et sa norme associée sont : pour $f, g \in L^2(\Omega)$,

$$(f,g)_{L^2} = \int_\Omega f(\vec{x})\,g(\vec{x})\,dx, \quad \text{et} \quad ||f||_{L^2} = \sqrt{(f,f)_{L^2}} = \Big(\int_\Omega f(\vec{x})^2\,dx\Big)^{\frac{1}{2}}. \tag{1.2}$$

Proposition 1.1 *$(L^2(\Omega), (\cdot,\cdot)_{L^2})$ est un Hilbert (espace vectoriel muni d'un produit scalaire qui est complet pour la norme associée).*

Preuve. Voir polycopié "Intégrale de Lebesgue". ∎

2 L'espace P_1

2.1 L'espace usuel P_1 sur $[a,b]$

Dans ce polycopié, l'exemple d'espace vectoriel de dimension finie composé de fonctions sera essentiellement l'espace P_1 des fonctions continues affines par morceaux sur un intervalle de $\mathbb{R}$. Les autres exemples usuels (P_k, Q_k, ...) sont calqués sur ce modèle.

2.1.1 Définition

Soit $a,b \in \mathbb{R}$, $a<b$, et soit $C^0([a,b])$ l'ensemble des fonctions continues de $[a,b]$ dans $\mathbb{R}$.
Soit $n \in \mathbb{N}^*$, $n \geq 2$, et soit n points x_i pour $i=1,...,n$ t.q. :

$$x_1 = a, \quad x_n = b, \quad \text{et} \quad x_i < x_{i+1}, \quad \forall i = 1,...,n{-}1. \tag{2.1}$$

On considère le "maillage" de $[a,b]$ en les $n{-}1$ intervalles consécutifs (faire un dessin) :

$$[a,b] = \bigcup_{i=1}^{n-1} [x_i, x_{i+1}]. \tag{2.2}$$

Voir figure 2.1. Le maillage est dit uniforme ssi $x_{i+1}-x_i$ est constant (noté h en éléments finis).
Soit P_1 l'ensemble des fonctions continues sur $[a,b]$ qui sont affines sur chaque intervalle du maillage :

$$\begin{aligned} P_1 &= \{f_h \in C^0([a,b]) \text{ et, } \forall i=1,...,n{-}1,\ f_{h|[x_i,x_{i+1}]} \text{ affine}\} \\ &= \{f_h \in C^0([a,b]) \text{ t.q. } \forall i=1,...,n{-}1,\ \exists a_i, b_i \in \mathbb{R},\ \forall x \in [x_i,x_{i+1}],\ f_h(x) = a_i + b_i x\} \end{aligned} \tag{2.3}$$

L'indice $_h$ dans f_h fait référence à la "largeur d'intervalle" $h = \max\limits_{i=1,...,n} (x_{i+1}{-}x_i)$, cf. polycopié d'éléments finis. Une fonction $f \in L^2(]a,b[)$ sera approchée par une fonction $f_h \in P_1$, voir § 5.8.3.

2.1.2 Base usuelle des fonctions chapeau

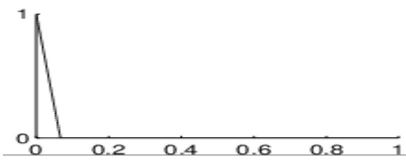
Définition 2.1 Les fonctions (de type) chapeau φ_{x_i} sont les fonctions de P_1 définies par, pour $i=1,...,n$:

$$\forall j = 1,...,n, \quad \varphi_{x_i}(x_j) = \delta_{ij} = \begin{cases} = 1 & \text{si } i = j, \\ = 0 & \text{si } i \neq j, \end{cases} \tag{2.4}$$

donc

$$\begin{cases} \varphi_{x_i}(x) = \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & \text{sur } [x_{i-1}, x_i] \quad (\text{si } i \geq 2), \\ \varphi_{x_i}(x) = \dfrac{x - x_{i+1}}{x_i - x_{i+1}} & \text{sur } [x_i, x_{i+1}] \quad (\text{si } i \leq n{-}1), \\ \varphi_{x_i}(x) = 0, \quad \text{ailleurs}, \end{cases} \tag{2.5}$$

pour $i=1,...,n$ (les φ_{x_i} sont continues et affines par morceau), voir figure 2.1. (Les fonctions extrémités φ_{x_1} et φ_{x_n} sont des "demi-chapeau".)



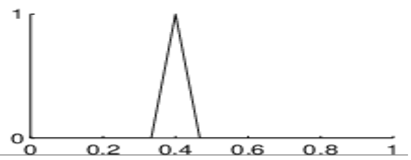


FIGURE 2.1 – Sur $[a,b]=[0,1]$, avec $x_i = \frac{i-1}{15}$ pour $i=1,...,15$ et le maillage $\bigcup_{i=1}^{14}[x_i,x_{i+1}]$: les fonctions chapeau φ_{x_1} et φ_{x_7} de la base de P_1.

Notation. Soit V un espace vectoriel de dimension n, et soit $(\vec{b}_i)_{i=1,...,n}$ une base de V.

$$\text{Si} \quad \vec{x} = \sum_{i=1}^{n} x_i \vec{b}_i \quad \text{alors on note} \quad [\vec{x}]_{|\vec{b}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{2.6}$$

la matrice colonne stockant les composantes de $\vec{x}$ dans la base $(\vec{b}_i)$.

Proposition 2.2 *L'ensemble P_1 donné en (2.3) est un espace vectoriel de dimension n, et $(\varphi_{x_i})_{i=1,\dots,n}$ est une base de P_1 (base constituée des fonctions chapeau). Et toute fonction $f_h \in P_1$ est donnée par :*

$$f_h = \sum_{i=1}^n y_i\,\varphi_{x_i}, \quad \text{où} \quad y_i = f_h(x_i), \quad \text{soit} \quad [f_h]_{|\varphi} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \overset{\text{noté}}{=} [\vec{y}]. \tag{2.7}$$

Donc $f_h(x) = \sum_{i=1}^n y_i\,\varphi_{x_i}(x)$ pour tout $x \in [a,b]$, avec $y_i = f_h(x_i)$.

Preuve. Si $f_h, g_h \in P_1$ et $\alpha \in \mathbb{R}$, alors $\alpha f_h + g_h$ est bien dans P_1 (est continue dans $[a,b]$ et affine sur chaque $[x_i, x_{i+1}]$). Donc P_1 est un sous-espace vectoriel de $(C^0([a,b]),+,.)$.

Vérifions que $(\varphi_{x_i})_{i=1,\dots,n}$ est une base de P_1, ce qui donnera aussi $\dim P_1 = n$.

• Famille libre : si $\sum_{i=1}^n \alpha_i\varphi_{x_i} = 0$ (égalité entre fonctions), i.e. si pour tout $x \in [a,b]$, $\sum_{i=1}^n \alpha_i\varphi_{x_i}(x) = 0$ (égalité entre réels), alors en particulier $\sum_{i=1}^n \alpha_i\varphi_{x_i}(x_j) = 0$, donc $\sum_{i=1}^n \alpha_i\delta_{ij} = 0$, donc $\alpha_j = 0$, vrai pour tout j : tous les α_j sont nuls.

• Famille génératrice. Soit $f_h \in P_1$. Par définition de P_1, f_h est affine sur les $[x_i, x_{i+1}]$, et de plus f_h est continue, donc f_h est déterminée ssi on connaît ses valeurs $y_i = f_h(x_i)$ pour tout $i = 1, \dots, n$. Notons $y_i := f_h(x_i)$ et $g_h := \sum_{i=1}^n y_i\varphi_{x_i}$. On a $g_h \in P_1$ car P_1 est un espace vectoriel. Et $g_h(x_j) = \sum_{i=1}^n y_i\varphi_{x_i}(x_j) = \sum_{i=1}^n y_i\delta_{ij} = y_j$, ce pour tout j, donc $f_h = g_h$ aux points x_i pour $x_i = 1, \dots, n$. Donc $f_h = g_h$ sur $[a,b]$. Donc $f_h \in \text{Vect}\{\varphi_i,\ i = 1, \dots, n\}$. ∎

Exercice 2.3 Montrer que pour tout $x \in [a,b]$ on a $\sum_{i=1}^n \varphi_{x_i}(x) = 1$.

Réponse. La fonction $f_h = 1_{[a,b]}$ est dans P_1 (trivial), vérifie $f_h(x_i) = 1$ pour tout i, et on a (2.7). ∎

2.1.3 Base non orthonormée relativement à $(\cdot,\cdot)_{L^2}$

P_1 est un sous-espace vectoriel de $L^2(]a,b[)$ (car de dimension finie et les $\varphi_{x_i} \in L^2(]a,b[)$: trivial). Donc $(\cdot,\cdot)_{L^2}$ a un sens dans P_1. On note $(P_1, (\cdot,\cdot)_{L^2}) =^{\text{noté}} P_1$.

Lemme 2.4 *La base usuelle (φ_{x_i}) de P_1 définie en (2.4) n'est pas orthogonale pour le produit scalaire $(\cdot,\cdot)_{L^2}$ (en particulier ce n'est pas une b.o.n. dans $(P_1, (\cdot,\cdot)_{L^2})$).*

Preuve.

$$(\varphi_{x_1}, \varphi_{x_2})_{L^2} = \int_{x_1}^{x_n} \varphi_{x_1}(x)\varphi_{x_2}(x)\,dx = \int_{x_1}^{x_2} \varphi_{x_1}(x)\varphi_{x_2}(x)\,dx > 0, \tag{2.8}$$

car φ_{x_1} est nulle sur $]x_2, x_n[$ et φ_{x_1} et φ_{x_2} sont strictement positives continues sur $]x_1, x_2[$. Donc $(\varphi_{x_1}, \varphi_{x_2})_{L^2} \neq 0$, donc $(\varphi_{x_i})_{i=1,\dots,n}$ n'est pas orthogonale dans $(P_1, (\cdot,\cdot)_{L^2})$. ∎

On note, pour tout $i, j = 1, \dots, n$:

$$M_{ij} = (\varphi_{x_i}, \varphi_{x_j})_{L^2} = \int_a^b \varphi_{x_i}(x)\varphi_{x_j}(x)\,dx, \quad \text{et} \quad M = [M_{ij}]_{\substack{i=1,\dots,n \\ j=1,\dots,n}}, \tag{2.9}$$

M étant la matrice du produit scalaire $(\cdot,\cdot)_{L^2}$ dans P_1 relativement à la base (φ_{x_i}) (matrice de masse).

Exercice 2.5 Montrer :

$$||\varphi_{x_1}||^2_{L^2} = \frac{x_2 - x_1}{3}, \quad ||\varphi_{x_2}||^2_{L^2} = \frac{x_3 - x_1}{3}, \quad (\varphi_{x_1}, \varphi_{x_2})_{L^2} = \frac{x_2 - x_1}{6}. \tag{2.10}$$

Réponse. $||\varphi_{x_1}||^2_{L^2} = \int_{x_1}^{x_2} \frac{(x-x_2)^2}{(x_1-x_2)^2}\,dx = [\frac{(x-x_2)^3}{3(x_1-x_2)^2}]_{x_1}^{x_2} = \frac{(x_2-x_1)}{3}$.

$||\varphi_{x_2}||^2_{L^2} = \int_{x_1}^{x_2} \frac{(x-x_1)^2}{(x_2-x_1)^2}\,dx + \int_{x_2}^{x_3} \frac{(x-x_3)^2}{(x_2-x_3)^2}\,dx = \frac{(x_2-x_1)}{3} + \frac{(x_3-x_2)}{3} = \frac{(x_3-x_1)}{3}$.

$(\varphi_{x_1}, \varphi_{x_2})_{L^2} = \int_{x_1}^{x_2} \frac{(x-x_2)}{(x_1-x_2)}\,\frac{(x-x_1)}{(x_2-x_1)}\,dx = \frac{1}{6}(x_2 - x_1)$. ∎

2.2 Rappels de définitions

2.2.1 Matrice d'un produit scalaire = matrice de masse

Soit H un espace vectoriel réel de dimension n (exemple $H = P_1$).

Définition 2.6 La matrice d'un produit scalaire sur H dans une base de H est appelée matrice de masse (ou matrice de Gram).

Donc : si $(\vec{a}_i)_{1,\dots,n}$ est une base de H, si $(\cdot,\cdot)_H$ est un produit scalaire sur H, alors la matrice :

$$M = [(\vec{a}_i, \vec{a}_j)_H] \stackrel{\text{noté}}{=} [(\cdot,\cdot)_H]_{|\vec{a}} \tag{2.11}$$

est une matrice de masse : c'est la matrice du produit scalaire $(\cdot,\cdot)_H$ dans la base $(\vec{a}_i)$. Et on a immédiatement, pour tous $\vec{x} = \sum_{i\in I} x_i\vec{a}_i$ et $\vec{y} = \sum_{i\in I} y_i\vec{a}_i$:

$$(\vec{x}, \vec{y})_H = \sum_{i,j\in I} x_i y_j (\vec{a}_i, \vec{a}_j)_H = \sum_{i,j=1}^{n} x_i M_{ij} y_j = [\vec{x}]_{|\vec{a}}^T.M.[\vec{y}]_{|\vec{a}}. \tag{2.12}$$

Proposition 2.7 *Soit $(\vec{a}_i)_{1,\dots,n}$ une base de H, $\vec{x} \in H$, $\vec{x} = \sum_{i=1}^{n} x_i\vec{a}_i$. On a :*

$$(\vec{x}, \vec{a}_j)_H = \sum_{i=1}^{n} M_{ij} x_i. \tag{2.13}$$

Donc $(\vec{x}, \vec{a}_j)_H = x_j$ pour tout j ssi $M = I$, i.e. ssi $(\vec{a}_i)$ est une $(\cdot,\cdot)_H$-b.o.n..

Preuve. $(\vec{x}, \vec{a}_j)_H = (\sum_i x_i\vec{a}_i, \vec{a}_j)_H = \sum_i x_i(\vec{a}_i, \vec{a}_j)_H = \sum_i M_{ij} x_i$. ▪

Exemple 2.8 Soit $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$ (produit cartésien n fois), soit $(\vec{e}_i)$ sa base canonique, soit $(\cdot,\cdot)_{\mathbb{R}^n}$ son produit scalaire canonique (défini par $(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n} = \delta_{ij}$ pour tout i, j). Alors la matrice de masse associée est $M = [(\vec{e}_i, \vec{e}_j)_{\mathbb{R}^n}] = [\delta_{ij}] = I$ la matrice identité. ▪

Exemple 2.9 Soit P_1 muni du produit scalaire $(\cdot,\cdot)_{L^2}$ et (φ_{x_i}) est la base des fonctions chapeau, cf. (2.4). $M = [(\varphi_{x_i}, \varphi_{x_j})_{L^2})] = [M_{ij}]$ est la matrice de masse correspondante. Exempled'un maillage régulier :

$$M = \frac{h}{6}\begin{pmatrix} 2 & 1 & 0 & \dots & & 0 \\ 1 & 4 & 1 & 0 & & \vdots \\ 0 & 1 & 4 & 1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ & & \ddots & 1 & 4 & 1 \\ 0 & \dots & & 0 & 1 & 2 \end{pmatrix}, \tag{2.14}$$

cf. (2.10). C'est la matrice usuelle qu'on retrouve dans tous les problèmes élémentaires de mécanique 1-D ou de noyaux intégraux 1-D. Ici $M \neq I$: la base n'est pas orthonormée, cf. lemme 2.4. ▪

Rappels : • Le produit cartésien $\mathbb{R}^n := \mathbb{R} \times \dots \times \mathbb{R}$ (produit n-fois) muni de l'addition interne et de la multiplication externe usuelles est un espace vectoriel réel. On note $(\vec{e}_i)$ sa base canonique, où donc $\vec{e}_1 = (1, 0, \dots, 0)$, ..., $\vec{e}_n = (0, \dots, 0, 1)$ avec 0 l'élément neutre de l'addition et 1 l'élément de la multiplication du corps $\mathbb{R}$.

• L'ensemble des matrices réelles $n * n$ est noté $\mathcal{M}_{nn}$, et l'ensemble des matrices réelles $n * 1$ (dites matrices colonnes) est noté $\mathcal{M}_{n1}$. Ce sont des espaces vectoriels réels pour les opérations d'addition interne et de multiplication externe usuelles. La base canonique $(\vec{E}_i)$ de $\mathcal{M}_{n1}$ sera systématiquement utilisée, où donc $\vec{E}_1 := [\vec{e}_1]_{|\vec{e}} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, ..., $\vec{E}_n := [\vec{e}_n]_{|\vec{E}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. Une matrice $M \in \mathcal{M}_{nn}$ est définie positive ssi $\vec{c}^T.M.\vec{c} > 0$ pour tout $\vec{c} \in \mathcal{M}_{n1} - \{\vec{0}\}$ (matrice colonne non nulle).

Proposition 2.10 *1- Une matrice de masse est une matrice symétrique définie positive.*

2- Une matrice symétrique définie positive est une matrice de masse.

Preuve. $(H,(\cdot,\cdot)_H)$ Hilbert de dimension finie.

1- Soit $M=[(\vec{a}_i,\vec{a}_j)_H]$ une matrice de masse, où donc $(\vec{a}_i)$ est une base dans H. La symétrie de M est immédiate car $(\cdot,\cdot)_H$ est un produit scalaire. Et $(\vec{x},\vec{y})_H=[\vec{x}]_{|\vec{a}}^T.M.[\vec{y}]_{|\vec{a}}$, et $0<||\vec{x}||_H^2$ pour $\vec{x}\neq\vec{0}$, donc $0<[\vec{x}]_{|\vec{a}}^T.M.[\vec{x}]_{|\vec{a}}$ pour $\vec{x}\neq\vec{0}$: M est définie positive.

2- Soit $M=[M_{ij}]$ symétrique définie positive. Soit $(\vec{a}_j)$ une base de H ; on définit la forme bilinéaire $b(\cdot,\cdot)$ sur H par $b(\vec{a}_i,\vec{a}_j)=M_{ij}$. On vérifie immédiatement que $b(\cdot,\cdot)$ est un produit scalaire dans H. ▪

Proposition 2.11 *Un espace de Hilbert $(H,(\cdot,\cdot)_H)$ de dimension n est isomorphe au produit cartésien $\mathbb{R}^n$ muni de son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$ (défini par $(\vec{e}_i,\vec{e}_n)_{\mathbb{R}^n}=\delta_{ij}$ pour tout i,j).*

Preuve. Soit $(\vec{a}_i)$ une $(\cdot,\cdot)_H$-b.o.n. dans H (où donc $(\vec{a}_i,\vec{a}_j)_H=\delta_{ij}$ pour tout i,j). Soit $J:H\to\mathbb{R}^n$ l'application linéaire définie par $J(\vec{a}_i)=\vec{e}_i$. Alors J est trivialement bijective et vérifie $(J(\vec{a}_i),J(\vec{a}_i))_{\mathbb{R}^n}=\delta_{ij}$ car $(J(\vec{a}_i),J(\vec{a}_i))_{\mathbb{R}^n}=(\vec{e}_i,\vec{e}_j)_{\mathbb{R}^n}$, donc J "conserve le produit scalaire", donc J est un isomorphisme entre deux Hilbert. ▪

Exemple 2.12 Soit $M\in\mathcal{M}_{nn}$ symétrique définie positive. Soit $(\mathbb{R}^n,(\cdot,\cdot)_M)$ l'espace $\mathbb{R}^n$ muni d'un produit scalaire $(\cdot,\cdot)_M$ (où donc $(\vec{x},\vec{y})_M:=[\vec{x}]_{|\vec{e}}^T.M.[\vec{y}]_{|\vec{e}}$). On a : $(\mathbb{R}^n,(\cdot,\cdot)_M)$ est isomorphe à $\mathbb{R}^n$ usuel, i.e. isomorphe à $(\mathbb{R}^n,(\cdot,\cdot)_{\mathbb{R}^n})=(\mathbb{R}^n,(\cdot,\cdot)_I)$ (où donc $(\vec{x},\vec{y})_I:=[\vec{x}]_{|\vec{e}}^T.[\vec{y}]_{|\vec{e}}$). ▪

Exemple 2.13 Isomorphisme $(P_1,(\cdot,\cdot)_{L^2})$ vers $(\mathbb{R}^n,(\cdot,\cdot)_M)$. Soit $f_h,g_h\in P_1$. Notons $y_1,...,y_n$ et $z_1,...,z_n$ leurs composantes sur la base (φ_{x_i}) :

$$f_h=\sum_{i=1}^n y_i\varphi_{x_i},\quad g_h=\sum_{i=1}^n z_i\varphi_{x_i},\quad\text{i.e.}\quad [f_h]_{|\varphi}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\overset{\text{noté}}{=}[\vec{y}],\quad [g_h]_{|\varphi}=\begin{pmatrix}z_1\\ \vdots\\ z_n\end{pmatrix}\overset{\text{noté}}{=}[\vec{z}].\tag{2.15}$$

Et soit $\vec{y},\vec{z}\in\mathbb{R}^n$ définis par

$$[\vec{y}]_{|\vec{e}}=[\vec{y}],\quad [\vec{z}]_{|\vec{e}}=[\vec{z}],\quad\text{i.e.}\quad \vec{y}=\sum_{i=1}^n y_i\vec{e}_i,\quad \vec{z}=\sum_{i=1}^n z_i\vec{e}_i.\tag{2.16}$$

Comme $(f_h,g_h)_{L^2}=\sum_{ij}y_iz_j(\varphi_{x_i},\varphi_{x_j})_{L^2}=\sum_{ij}y_iM_{ij}z_j$, on a :

$$(f_h,g_h)_{L^2}=[\vec{y}]^T.M.[\vec{z}]=(\vec{y},\vec{z})_M.\tag{2.17}$$

Ainsi, les calculs dans $(P_1,(\cdot,\cdot)_{L^2})$ sont traduits en calculs dans $(\mathbb{R}^n,(\cdot,\cdot)_M)$, où $(\cdot,\cdot)_M$ est un produit scalaire dans $\mathbb{R}^n$ puisque M est une matrice symétrique définie positive (matrice de masse). ▪

2.2.2 Diagonalisation d'une matrice de masse

Soit M une matrice réelle de masse et le problème spectral (problème de recherche de valeurs et vecteurs propres) :

$$\begin{cases}\text{trouver les couples } (\lambda,\vec{v})\in\mathbb{R}\times\mathcal{M}_{n1} \text{ t.q. :}\\ M.\vec{v}=\lambda\vec{v}.\end{cases}\tag{2.18}$$

En général ce problème est traduit en problème dans $\mathbb{R}^n$ cartésien muni de sa base canonique $(\vec{e}_i)$ et de son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$:

$$\begin{cases}\text{trouver les couples } (\lambda,\vec{v})\in\mathbb{R}\times\mathbb{R}^n \text{ t.q. :}\\ M.[\vec{v}]_{|\vec{e}}=\lambda[\vec{v}]_{|\vec{e}},\quad\text{réécrit abusivement}\quad M.\vec{v}=\lambda\vec{v}.\end{cases}\tag{2.19}$$

C'est cette forme traduite et sa notation abusive qu'on utilisera quand il n'y a aura pas d'ambiguïté.

Comme M est symétrique réelle, elle est diagonalisable dans une b.o.n. $(\vec{v}_i)$ de $(\mathbb{R}^n,(\cdot,\cdot)_{\mathbb{R}^n})$, voir poly diagonalisation. I.e. le problème spectral (2.19) a une infinité de solutions $(\lambda,\vec{v})$ parmi lesquelles on peut en choisir n, notées $(\lambda_i,\vec{v}_i)$, telles que $(\vec{v}_i)_{i=1,...,n}$ est une b.o.n. (de vecteurs propres), dite b.o.n. de diagonalisation, et les λ_i sont les valeurs propres associées : $\exists\lambda_1,...,\lambda_n \in \mathbb{R}$, $\exists\vec{v}_1,...,\vec{v}_n \in \mathbb{R}^n$, $\forall i,j=1,...,n$,

$$M.\vec{v}_j = \lambda_j\vec{v}_j \quad \text{et} \quad (\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n} = \delta_{ij}. \tag{2.20}$$

Notons $P=[P_{ij}]$ la matrice de passage de la base $(\vec{e}_i)$ à la base $(\vec{v}_i)$, où donc

$$\vec{v}_j = \sum_{i=1}^n P_{ij}\vec{e}_i, \quad \text{i.e.} \quad [\vec{v}_j]_{|\vec{e}} = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \quad \text{et} \quad P=[P_{ij}] = (\,[\vec{v}_1]_{\vec{e}} \quad ... \quad [\vec{v}_n]_{\vec{e}}\,). \tag{2.21}$$

Avec $D=\mathrm{diag}(\lambda_i)$ la matrice $n*n$ des valeurs propres, (2.20) s'écrit sous forme matricielle :

$$M.P = P.D \quad \text{et} \quad P^T.P = I, \quad \text{donc} \quad D = P^{-1}.M.P \quad \text{et} \quad P^{-1} = P^T. \tag{2.22}$$

Proposition 2.14 *Si M est une matrice de masse alors ses valeurs propres sont toutes >0, et M est inversible.*

Preuve. $\lambda_i = \lambda_i(\vec{v}_i,\vec{v}_i)_H = (\lambda_i\vec{v}_i,\vec{v}_i)_H = (M.\vec{v}_i,\vec{v}_i)_H > 0$, car M est définie positive, vrai pour tout i. Donc D est inversible, avec P inversible (matrice de changement de base), donc $M = P.D.P^{-1}$ est inversible. ∎

2.2.3 B.o.n. dans H par diagonalisation d'une matrice de masse

Soit $(\vec{a}_i)$ une base de H (par exemple la base (φ_{x_i}) dans P_1), soit $M=[(\vec{a}_i,\vec{a}_j)_H]$ la matrice de masse associée.

On se plonge dans $\mathbb{R}^n$: soit $(\vec{v}_i)$ une b.o.n. de $\mathbb{R}^n$ de diagonalisation de M associée aux valeurs propres λ_i de M, cf. (2.20). Et on défini $\vec{p}_j \in H$ par, pour tout j,

$$\vec{p}_j = \sum_{i=1}^n P_{ij}\vec{a}_i, \quad \text{i.e.} \quad [\vec{p}_j]_{|\vec{a}} := \vec{v}_j = \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}. \tag{2.23}$$

Proposition 2.15 *Les $\vec{p}_j$ forment une base orthogonale dans H : on a, pour tout i,j,*

$$(\vec{p}_i,\vec{p}_j)_H = \lambda_i\delta_{ij}, \quad \textit{donc} \quad (\vec{b}_i,\vec{b}_j)_H = \delta_{ij} \quad \textit{où} \quad \vec{b}_i = \frac{\vec{p}_i}{\sqrt{\lambda_i}}, \tag{2.24}$$

et $(\vec{b}_i) = (\frac{\vec{p}_i}{\sqrt{\lambda_i}})_{i=1,...,n}$ est orthonormale dans $(H,(\cdot,\cdot)_H)$.

Preuve. Par bilinéarité de $(\cdot,\cdot)_H$, pour tout i,j on a, avec (2.22) :
$$(\vec{p}_i,\vec{p}_j)_H = \sum_{k\ell} P_{ki}P_{\ell j}(\vec{a}_k,\vec{a}_\ell)_H = \sum_{k\ell}(P^T)_{ik}M_{k\ell}P_{\ell j} = (P^T.M.P)_{ij} = (P^{-1}.M.P)_{ij} = D_{ij} = \lambda_i\delta_{ij}.$$ ∎

Exercice 2.16 Soit H un espace vectoriel de dimension 2 muni d'un produit scalaire $(\cdot,\cdot)_H$. Soit $(\vec{a}_1,\vec{a}_2)$ une base de H et $M=[(\vec{a}_i,\vec{a}_j)_H] = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ la matrice de masse associée. Diagonaliser M et trouver la b.o.n. $(\vec{b}_1,\vec{b}_2)$ (à exprimer dans la base $(\vec{a}_1,\vec{a}_2)$, donc donner P).

Réponse. On a $\det(M-\lambda I) = \lambda^2 - 4\lambda + 3 = (\lambda-1)(\lambda-3)$. Donc $\lambda_1 = 1$ et $\lambda_2 = 3$ sont les deux valeurs propres. Soit $(\vec{e}_1,\vec{e}_2)$ la base canonique de $\mathbb{R}^2$. Et $\vec{v}_1 = \frac{\sqrt{2}}{2}(\vec{e}_1 - \vec{e}_2)$ et $\vec{v}_2 = \frac{\sqrt{2}}{2}(\vec{e}_1 + \vec{e}_2)$ (où donc $[\vec{v}_1] = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $[\vec{v}_1] = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$) sont deux vecteurs propres normés associés formant une b.o.n. dans $(\mathbb{R}^2,(\cdot,\cdot)_{\mathbb{R}^2})$ (immédiat). Et $P = (\,[\vec{v}_1] \quad [\vec{v}_2]\,) = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Donc dans H on prend $\vec{p}_1$ et $\vec{p}_2$ t.q. $[\vec{p}_1]_{|\vec{a}} = [\vec{v}_1]$ et $[\vec{p}_2]_{|\vec{a}} = [\vec{v}_2]$, i.e. $\vec{p}_1 = \frac{\sqrt{2}}{2}(\vec{a}_1 - \vec{a}_2)$ et $\vec{p}_2 = \frac{\sqrt{2}}{2}(\vec{a}_1 + \vec{a}_2)$. Puis on pose $\vec{b}_1 = \frac{1}{\sqrt{\lambda_1}}\vec{p}_1 = \vec{p}_1 = \frac{\sqrt{2}}{2}(\vec{a}_1 - \vec{a}_2)$ et $\vec{b}_2 = \frac{1}{\sqrt{\lambda_2}}\vec{p}_2 = \frac{1}{\sqrt{3}}\vec{p}_2 = \frac{1}{\sqrt{6}}(\vec{a}_1 + \vec{a}_2)$. Vérifications :

$(\vec{b}_1,\vec{b}_1)_H = \frac{1}{2}((\vec{a}_1,\vec{a}_1)_H - 2(\vec{a}_1,\vec{a}_2)_H + (\vec{a}_2,\vec{a}_2)_H = \frac{1}{2}(M_{11} - 2M_{12} + M_{22}) = 1$,
$(\vec{b}_1,\vec{b}_2)_H = \frac{1}{2\sqrt{3}}((\vec{a}_1,\vec{a}_1)_H - (\vec{a}_2,\vec{a}_2)_H = \frac{1}{2}(M_{11} - M_{22}) = 0$,
$(\vec{b}_2,\vec{b}_2)_H = \frac{1}{6}((\vec{a}_1,\vec{a}_1)_H + 2(\vec{a}_1,\vec{a}_2)_H + (\vec{a}_2,\vec{a}_2)_H = \frac{1}{6}(M_{11} + 2M_{12} + M_{22}) = 1$. ∎

2.2.4 Exemple : b.o.n. de diagonalisation dans P_1

Matrice de masse $M = [(\varphi_{x_i}, \varphi_{x_j})_{L^2}]$ (voir (2.14) par exemple). (2.22) donne :

$$M = P.D.P^{-1} \quad \text{avec} \quad P^{-1} = P^T. \tag{2.25}$$

La proposition 2.15 donne : une b.o.n. $(\zeta_i)_{i=1,...,n}$ de fonctions de P_1 associée aux valeurs propres λ_i est donnée par, cf. (2.24), pour $i = 1, ..., n$:

$$[\zeta_j]_{|\varphi} = \frac{1}{\sqrt{\lambda_j}} \begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}, \quad \text{i.e.} \quad \zeta_j = \frac{1}{\sqrt{\lambda_j}} \sum_{i=1}^{n} P_{ij} \varphi_{x_i}. \tag{2.26}$$

Voir figure 2.2.

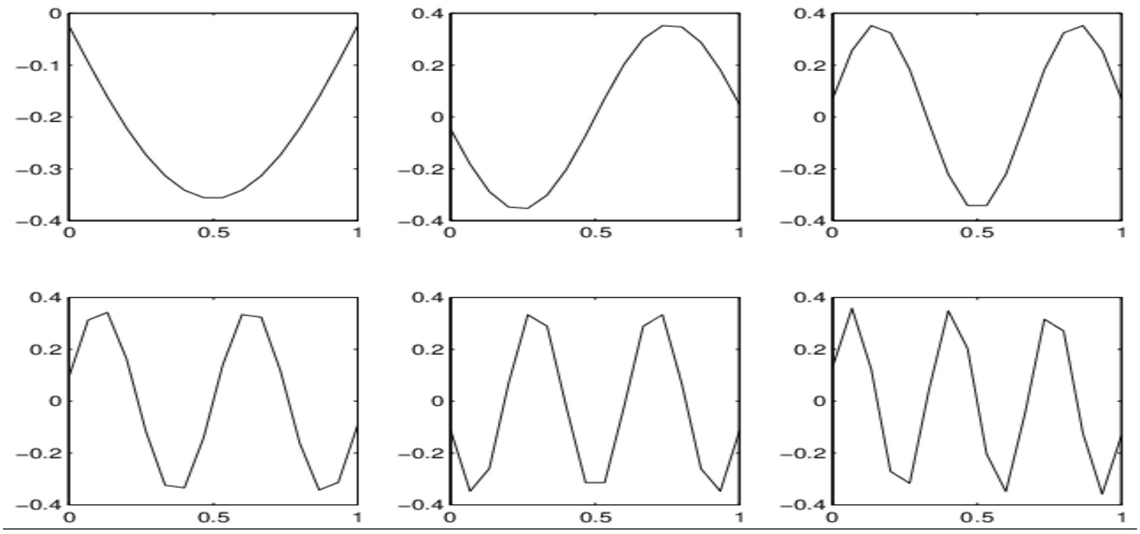


FIGURE 2.2 – Suite de la figure 3.1 (maillage de 15 intervalles réguliers). Représentation de 6 fonctions propres associées aux 6 premières valeurs propres rangées dans l'ordre décroissant, cf. (2.26).

3 Base duale vectorielle (relativement à un produit scalaire)

3.1 Définition

Soit $(H, (\cdot,\cdot)_H)$ un espace de Hilbert de dimension finie n muni d'un produit scalaire, soit $(\vec{a}_i)_{i=1,...,n}$ une base de H. Soit $M = [(\vec{a}_i, \vec{a}_j)_H] = [(\cdot,\cdot)_H]_{|\vec{a}}$ (la matrice de masse de $(\cdot,\cdot)_H$ pour la base $(\vec{a}_i)$).

Définition 3.1 La base duale vectorielle (ou base vectorielle duale) de la base $(\vec{a}_i)_{i=1,...,n}$ relativement au produit scalaire $(\cdot,\cdot)_H$ est la base des vecteurs notée $(\vec{a}_i^{\star})_{i=1,...,n}$ de H qui vérifie, pour tous $i, j = 1, ..., n$:

$$(\vec{a}_i^{\star}, \vec{a}_j)_H = \delta_{ij}. \tag{3.1}$$

(Attention, elle dépend du choix du produit scalaire.)

3.2 Composantes à l'aide de la base duale

Proposition 3.2 *La base duale vectorielle $(\vec{a}_i^{\star})_{i=1,...,n}$ existe et est unique, et la matrice $N = [(\vec{a}_i^{\star}, \vec{a}_j^{\star})_H]$ (matrice de masse de $(\cdot,\cdot)_H$ pour la base $(\vec{a}_i^{\star})$) vérifie*

$$N = M^{-1}, \quad \text{avec} \quad \vec{a}_j^{\star} = \sum_{i=1}^{n} N_{ij} \vec{a}_i. \tag{3.2}$$

Donc N est la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{a}_i^\star)$ (stocke dans sa j-ème colonne les composantes de $\vec{a}_j^\star$ dans la base $(\vec{a}_i)$). Et si $\vec{x} \in H$, $\vec{x} = \sum_{j=1}^n x_j \vec{a}_j$ alors, pour tout j,

$$x_j = (\vec{x}, \vec{a}_j^\star)_H. \tag{3.3}$$

Et la base vectorielle biduale est la base initiale : $(\vec{a}_i^\star)^\star = \vec{a}_i$ pour tout $i = 1, ..., n$.

Preuve. Supposons que la base duale $(\vec{a}_i^\star)$ existe. Notons Z_{ij} les composantes de $\vec{a}_j^\star$ dans la base $(\vec{a}_i)$, i.e. $\vec{a}_j^\star = \sum_{i=1}^n Z_{ij}\vec{a}_i$. Alors (3.1) donne $\delta_{ij} = (\vec{a}_i^\star, \vec{a}_j)_H = \sum_k Z_{kj}(\vec{a}_k, \vec{a}_i)_H = \sum_k M_{ik}Z_{kj} = (M.Z)_{ij}$ (car M est symétrique), pour tous i, j, donc $I = M.Z$, donc Z est inversible d'inverse $Z = M^{-1}$ (car M est inversible cf. prop. 2.14). Donc les Z_{ij} sont uniques, donc si la base duale existe alors elle est unique.

Puis posons $Z = M^{-1}$ et $\vec{b}_j = \sum_{i=1}^n Z_{ij}\vec{a}_i$. On a $(\vec{b}_j, \vec{a}_i)_H = \sum_k Z_{kj}(\vec{a}_k, \vec{a}_i)_H = \sum_k M_{ik}Z_{kj} = (MZ)_{ij} = \delta_{ij}$ pour tous i, j. Donc $(\vec{b}_i)$ est une base duale de $(\vec{a}_i)$. Vu l'unicité $(\vec{b}_i) =^{\text{noté}} (\vec{a}_i^\star)$ existe et est unique. Et $(\vec{b}_i, \vec{b}_j)_H = (\vec{b}_i, \sum_k Z_{kj}\vec{a}_k)_H = \sum_k Z_{kj}(\vec{b}_i, \vec{a}_k)_H = \sum_k Z_{kj}\delta_{ik} = Z_{ij}$ pour tout i, j, donc $Z = N$, et $Z = N$ est la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{a}_i^\star)$.

Puis $\vec{x} = \sum_i x_i \vec{a}_i$ donne $(\vec{x}, \vec{a}_j^\star)_H = \sum_i x_i (\vec{a}_i, \vec{a}_j^\star)_H = \sum_i x_i \delta_{ij} = x_j$.

Et, par symétrie du produit scalaire, $(\vec{a}_i, \vec{a}_j^\star)_H = \delta_{ij}$ pour tous i, j : donc $(\vec{a}_i^\star)^\star = \vec{a}_i$ pour tout i. ∎

Corollaire 3.3 *Ssi $(\vec{a}_i)$ est une b.o.n., les projections π_i sur* $\mathrm{Vect}\{\vec{a}_i\}$ *parallèlement aux directions $\vec{a}_j$ pour $j \neq i$ sont :*

$$\pi_i = (\vec{a}_i^\star, \cdot)_H : \begin{cases} H & \to \mathbb{R} \\ \vec{x} = \sum_{j=1}^n x_j \vec{a}_j & \to \pi_i(\vec{x}) \overset{\text{déf}}{=} (\vec{a}_i^\star, \vec{x})_H = x_i. \end{cases} \tag{3.4}$$

Exercice 3.4 Dans $(\mathbb{R}^2, (\cdot,\cdot)_{\mathbb{R}^2})$ avec $(\vec{e}_1, \vec{e}_2)$ la base canonique (et $(\cdot,\cdot)_{\mathbb{R}^2}$ le produit scalaire canonique). Soit $\vec{a}_1 := \vec{e}_1$ et $\vec{a}_2 := \vec{e}_1 + \vec{e}_2$. Calculer la base duale $(\vec{a}_1^\star, \vec{a}_2^\star)$.

Réponse. Le couple $(\vec{a}_1, \vec{a}_2)$ est une base (immédiat). Cherchons $\vec{a}_1^\star$ et $\vec{a}_2^\star$ sous la forme $\vec{a}_1^\star = Z_{11}\vec{e}_1 + Z_{21}\vec{e}_2$ et $\vec{a}_2^\star = Z_{12}\vec{e}_1 + Z_{22}\vec{e}_2$.

Réponse 1 : (3.1) s'écrit ici $(\vec{a}_i^\star, \vec{a}_j)_{\mathbb{R}^2} = \delta_{ij}$, donc $Z_{11}*1 + Z_{21}*0 = 1$, $Z_{11}*1 + Z_{21}*1 = 0$, $Z_{12}*1 + Z_{22}*0 = 0$ et $Z_{12}*1 + Z_{22}*1 = 1$, donc $Z_{11} = 1$, $Z_{21} = -1$, $Z_{12} = 0$, $Z_{22} = 1$, donc $\vec{a}_1^\star = \vec{e}_1 - \vec{e}_2$ et $\vec{a}_2^\star = \vec{e}_2$.

Réponse 2 : la matrice de passage de la base $(\vec{e}_i)$ à la base $(\vec{a}_i)$ est $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Notons Z la matrice de passage de la base $(\vec{e}_i)$ à la base $(\vec{a}_i^\star)$. On a $(\vec{a}_i^\star, \vec{a}_j)_{\mathbb{R}^2} = \delta_{ij}$, donc $[\vec{a}_i^\star]_{|\vec{e}}^T.[\vec{a}_j]_{|\vec{e}} = \delta_{ij}$, pour tout i, j, donc $Z^T.P = I$. Donc $Z^T = P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, donc $Z = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, donc $[\vec{a}_1^\star]_{|\vec{e}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $[\vec{a}_2^\star]_{|\vec{e}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Réponse 3 : $M = [(\vec{a}_i, \vec{a}_j)_{\mathbb{R}^n}] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, d'inverse $N = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Donc $\vec{a}_1^\star = 2\vec{a}_1 - \vec{a}_2$ et $\vec{a}_2^\star = -\vec{a}_1 + \vec{a}_2$. ∎

Exemple 3.5 Suite de l'exemple 3.4. $M = [(\vec{a}_i, \vec{a}_j)_{\mathbb{R}^n}] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, d'inverse $N = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Donc $[\vec{a}_1^\star]_{|\vec{a}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ ("première colonne" de N) et $[\vec{a}_2^\star]_{|\vec{a}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ("deuxième colonne" de N). Donc $\vec{a}_1^\star = 2\vec{a}_1 - \vec{a}_2$ et $\vec{a}_2^\star = -\vec{a}_1 + \vec{a}_2$ sont les vecteurs duaux exprimés dans la base $(\vec{a}_i)$.

Vérification : $(\vec{a}_1^\star, \vec{a}_1)_{\mathbb{R}^n} = 2(\vec{a}_1, \vec{a}_1)_{\mathbb{R}^n} - (\vec{a}_2, \vec{a}_1)_{\mathbb{R}^n} = 2M_{11} - M_{21} = 1$, $(\vec{a}_1^\star, \vec{a}_2)_{\mathbb{R}^n} = 2(\vec{a}_1, \vec{a}_2)_{\mathbb{R}^n} - (\vec{a}_2, \vec{a}_2)_{\mathbb{R}^n} = 2M_{12} - M_{22} = 0$, $(\vec{a}_2^\star, \vec{a}_1)_{\mathbb{R}^n} = -(\vec{a}_1, \vec{a}_1)_{\mathbb{R}^n} + (\vec{a}_2, \vec{a}_1)_{\mathbb{R}^n} = -M_{11} + M_{21} = 0$, $(\vec{a}_2^\star, \vec{a}_2)_{\mathbb{R}^n} = -(\vec{a}_1, \vec{a}_2)_{\mathbb{R}^n} + (\vec{a}_2, \vec{a}_2)_{\mathbb{R}^n} = -M_{12} + M_{22} = 1$. ∎

3.3 Exemple : base duale vectorielle dans P_1

Dans P_1, la base $(\cdot,\cdot)_H$-duale vectorielle $(\varphi_{x_i}^\star)_{i=1,\dots,n}$ de la base $(\varphi_{x_i})_{i=1,\dots,n}$ est donnée par

$$\varphi_{x_j}^\star = \sum_{i=1}^n N_{ij}\varphi_{x_i}. \tag{3.5}$$

Voir figure 3.1. Et puisque $\varphi_{x_i}(x_j) = \delta_{ij}$ on a donc $\varphi_{x_j}^\star(x_i) = N_{ij}$, donc, pour tout $i, j = 1, ..., n$,

$$\boxed{\varphi_{x_j}^\star(x_i) = (\varphi_{x_i}^\star, \varphi_{x_j}^\star)_{L^2}}. \tag{3.6}$$

Cette relation $\varphi_{x_j}^\star(x_i) = (\varphi_{x_i}^\star, \varphi_{x_j}^\star)_{L^2}$ sera très utilisée dans la suite (noyau reproduisant).

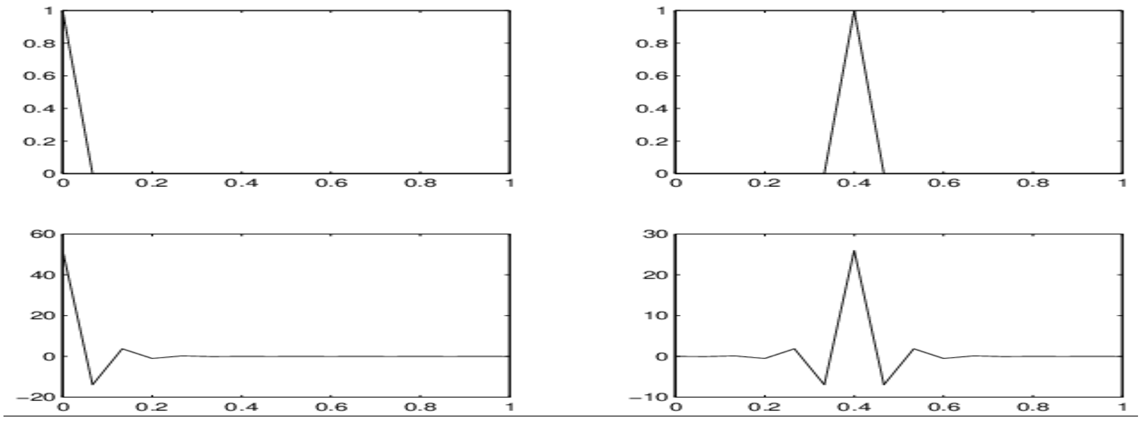


FIGURE 3.1 – Sur $[a,b]=[0,1]$, pour le maillage $\bigcup_{i=1}^{15}[\frac{i}{15},\frac{i+1}{15}]$. En haut les fonctions chapeau φ_{x_1} et φ_{x_7} de la base de P_1, comme en figure 2.1, et en bas les fonctions $\varphi^\star_{x_1}$ et $\varphi^\star_{x_7}$ de la base vectorielle duale de P_1, relativement au produit scalaire $(\cdot,\cdot)_{L^2}$.

3.4 Noyau reproduisant dans P_1

Si $f_h\in P_1$, $f_h=\sum_{i=1}^n y_i\varphi_{x_i}$, alors $y_i =^{(3.3)} (\varphi^\star_{x_i},f_h)_{L^2}$, donc

$$f_h=\sum_{i=1}^n(\varphi^\star_{x_i},f_h)_{L^2}\varphi_{x_i}. \tag{3.7}$$

Donc, pour $x\in[a,b]$:

$$f_h(x)=\sum_{i=1}^n(\int_{y=a}^b f_h(y)\varphi^\star_{x_i}(y)\,dy)\varphi_{x_i}(x)=\int_{y=a}^b(\sum_{i=1}^n\varphi^\star_{x_i}(y)\varphi_{x_i}(x))f_h(y)\,dy \tag{3.8}$$

Donc, posant

$$K=\sum_{i=1}^n\varphi_{x_i}\otimes\varphi^\star_{x_i},\quad \text{i.e.}\quad K(x,y)\overset{\text{déf}}{=}\sum_{i=1}^n\varphi_{x_i}(x)\varphi^\star_{x_i}(y),\quad \text{et avec}\quad K_x(y)\overset{\text{déf}}{=}K(x,y), \tag{3.9}$$

on a

$$f_h(x)=(K_x,f_h)_{L^2}=\int_{y=a}^b K(x,y)f_h(y)\,dy. \tag{3.10}$$

Et $K=\sum_{i=1}^n\varphi_{x_i}\otimes\varphi^\star_{x_i}$ sera notre noyau reproduisant : le calcul $(K_x,f_h)_{L^2}$ reproduit $f_h(x)$. Dessin de K figure 3.2 : sur chaque rectangle $R_{ij}=[x_i,x_{i+1}]\times[x_j,x_{j+1}]$, les φ_{x_i} et les $\varphi^\star_{x_i}$ étant affines par morceau, (3.9) indique que la fonction K est bi-affine par morceaux, i.e. de la forme

$$K_{|R_{ij}}(x,y)=(a_ix+b_i)(c_jy+d_j). \tag{3.11}$$

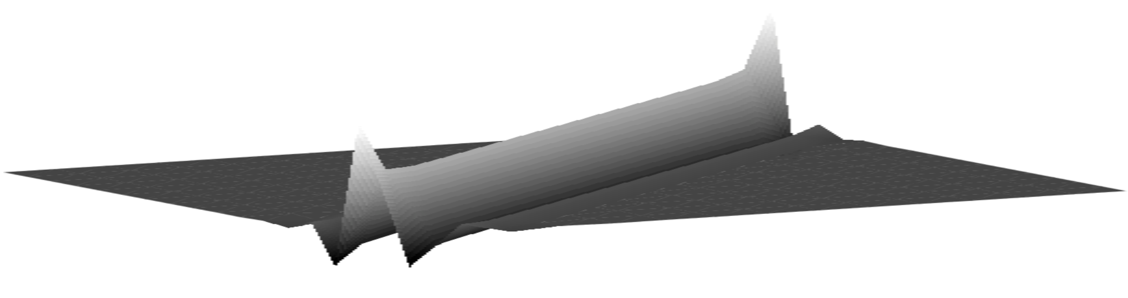


FIGURE 3.2 – Le graphe du noyau intégral reproduisant K, cf. (3.9) (suite de la figure 3.1).

4 L'espace usuel P_0 sur $[a,b]$

On conserve le maillage (2.2) : $[a,b]=\bigcup_{i=1}^{n-1}[x_i,x_{i+1}]$. On définit :

$$\begin{aligned} P_0 &\overset{\text{déf}}{=} \{f_h:[a,b]\to\mathbb{R} \text{ t.q. } \forall i=1,...,n-1,\ f_{h|]x_i,x_{i+1}[} \text{ constante}\} \\ &= \{f_h:[a,b]\to\mathbb{R} \text{ t.q. } \forall i=1,...,n-1,\ \exists a_i\in\mathbb{R},\ \forall x\in]x_i,x_{i+1}[,\ f_h(x)=a_i\}, \end{aligned} \tag{4.1}$$

l'ensemble des fonctions en escalier relativement au maillage.

P_0 est un ensemble de fonctions définies presque partout au sens de la mesure de Lebesgue.

Différence essentielle avec P_1 : les valeurs d'une $f_h\in P_0$ ne sont pas données aux x_i.

Remarque 4.1 Utilité : Stockage d'une photo sous forme de "pixels" . L'ensemble des pixels reconstitue la photo (de manière approchée). ♣

Il est immédiat que P_0 est un sous-espace vectoriel de $L^2(]a,b[)$ de dimension $n-1$, dont une base est la base $(\varphi_i)_{i=1,...,n-1}$ des fonctions définies par (les fonctions indicatrices) :

$$\varphi_i=1_{]x_i,x_{i+1}[},\quad i=1,...,n-1. \tag{4.2}$$

Et (immédiat) $M_{ij}=(\varphi_i,\varphi_j)_{L^2}=\int_a^b\varphi_i(x)\varphi_j(x)\,dx=\int_{]x_i,x_{i+1}[\bigcap]x_j,x_{j+1}[}dx=(x_{i+1}-x_i)\delta_{ij}$. (Et donc une b.o.n. dans $(P_0,(\cdot,\cdot)_{L^2})$ est donnée par les ζ_i où $\zeta_i(x)=\frac{1}{\sqrt{x_{i+1}-x_i}}\varphi_i(x)$.) Et donc la matrice de masse de $(\cdot,\cdot)_{L^2}$ pour la base (φ_i) est

$$M=\mathrm{diag}((x_{i+1}-x_i)). \tag{4.3}$$

Donc $N=M^{-1}=\mathrm{diag}((x_{i+1}-x_i)^{-1})$, et la base duale vectorielle $(\varphi_i^\star)_{i=1,...,n-1}$ relativement au produit scalaire $(\cdot,\cdot)_{L^2}$ (celle qui vérifie $(\varphi_j,\varphi_i^\star)_{L^2}=\delta_{ij}$) est donnée par, cf. (3.5) :

$$\varphi_j^\star=\sum_{i=1}^n N_{ij}\varphi_i=(x_{j+1}-x_j)^{-1}\varphi_j. \tag{4.4}$$

Et toute $f_h\in P_0$ s'écrit, pour presque tout $x\in[a,b]$ (plus précisément pour tout $x\notin\{x_1,...,x_n\}$) :

$$\begin{aligned} f_h(x) &= \sum_{j=1}^{n-1}c_j\varphi_j(x)=\sum_{j=1}^{n-1}(f_h,\varphi_j^\star)_{L^2}\varphi_j(x)=\int_{y=a}^b f_h(y)\sum_{j=1}^{n-1}\varphi_j^\star(y)\varphi_j(x)\,dy \\ &= \int_{y=a}^b K(x,y)f_h(y)\,dy=(K_x,f_h)_{L^2}, \end{aligned} \tag{4.5}$$

où

$$K(x,y):=\sum_{j=1}^{n-1}\varphi_j(x)\varphi_j^\star(y)\quad\text{et}\quad K_x(y):=K(x,y). \tag{4.6}$$

La fonction $K\in L^2(]a,b[^2)$ est notre noyau reproduisant.

Remarque 4.2 K_x n'est pas définie pour $x=x_i$, mais définir les K_{x_i} ne sert pas à grand chose puisque l'information sauvegardée est celle à "l'intérieur des pixels" (et non au bord des pixels). ♣

5 Rappels : endomorphisme

5.1 Définition

Définition 5.1 Soit H un espace vectoriel. Un endomorphisme sur H (ou dans H, ou de H) est une application linéaire de H dans lui-même. On note $\mathcal{L}(H)$ l'ensemble des endomorphismes de H. Et pour $L\in\mathcal{L}(H)$ on note :

$$L(\vec{x})\overset{\text{noté}}{=}L.\vec{x}, \tag{5.1}$$

notation d'un produit car la linéarité permet de distribuer : $L.(\vec{x}+\lambda\vec{y})=L.\vec{x}+\lambda L.\vec{y}$.

5.2 Matrice de représentation dans une base

Soit H un espace vectoriel soit de dimension finie n. Soit $(\vec{a}_i)_{i=1,\dots,n} =^{\text{not.}} (\vec{a}_i)$ une base de H. Soit $L \in \mathcal{L}(H)$ (endomorphisme). Pour $j = 1, ..., n$, on note L_{ij} les composantes du vecteur $L.\vec{a}_j$ dans la base $(\vec{a}_i)$:

$$L.\vec{a}_j = \sum_{i=1}^n L_{ij}\,\vec{a}_i, \quad \text{i.e.} \quad [L.\vec{a}_j]_{|\vec{a}} = \begin{pmatrix} L_{1j} \\ L_{2j} \\ \vdots \end{pmatrix}. \tag{5.2}$$

Définition 5.2 $[L]_{|\vec{a}} = [L_{ij}]_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ est appelée la matrice de L relativement à la base $(\vec{a}_i)$.

S'il n'y a pas d'ambiguïté sur le choix de la base on note $[L]_{|\vec{a}} = [L] = [L_{ij}]$.

Par linéarité de L, quand $\vec{x} = \sum_{j=1}^n x_j \vec{a}_j$ on a $L.\vec{x} = \sum_{j=1}^n x_j\, L.\vec{a}_j$, et donc :

$$L.\vec{x} = \sum_{i=1}^n \sum_{j=1}^n L_{ij} x_j\, \vec{a}_i, \quad \text{donc} \quad [L.\vec{x}]_{|\vec{a}} = \begin{pmatrix} \sum_{j=1}^n L_{1j} x_j \\ \sum_{j=1}^n L_{2j} x_j \\ \vdots \end{pmatrix} = [L]_{|\vec{a}}.[\vec{x}]_{|\vec{a}} \overset{\text{not.}}{=} [L.\vec{x}] = [L].[\vec{x}], \tag{5.3}$$

la dernière notation s'il n'y a pas d'ambiguïté sur le choix de la base.

Remarque 5.3 Si on dispose d'un produit scalaire $(\cdot,\cdot)_H$ sur H, alors on dispose de la base duale vectorielle $(\vec{a}_i^\star)$, cf. (3.1). Donc (5.2) donne immédiatement, pour tous i, j :

$$L_{ij} = (\vec{a}_i^\star, L.\vec{a}_j)_H, \tag{5.4}$$

cf. (3.4). ▟

Exercice 5.4 Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases dans H, et soit P la matrice de passage de la base $(\vec{a}_i)$ à la base $(\vec{b}_i)$, et soit Q la matrice de passage de la base $(\vec{b}_i)$ à la base $(\vec{a}_i)$, i.e., pour tout $j = 1, ..., n$:

$$\vec{b}_j = \sum_{i=1}^n P_{ij} \vec{a}_i, \quad \vec{a}_j = \sum_{i=1}^n Q_{ij} \vec{b}_i. \tag{5.5}$$

Vérifier $Q = P^{-1}$, et vérifier la formule de changement de base pour les endomorphismes :

$$[L]_{\vec{b}} = P^{-1}.[L]_{\vec{a}}.P. \tag{5.6}$$

Réponse. $\vec{a}_j = \sum_{i=1}^n Q_{ij}\vec{b}_i = \sum_i Q_{ij}(\sum_k P_{ki}\vec{a}_k) = \sum_k(\sum_i P_{ki}Q_{ij})\vec{a}_k = \sum_k (P.Q)_{kj}\vec{a}_k$. Donc $(P.Q)_{kj} = \delta_{kj}$, vrai pour tout j, k, donc $P.Q = I$, soit $Q = P^{-1}$.

Soit $A = [L]_{\vec{a}}$ et $B = [L]_{\vec{b}}$, où donc $L.\vec{a}_j = \sum_i A_{ij}\vec{a}_i$ et $L.\vec{b}_j = \sum_i B_{ij}\vec{b}_i$, pour tout j. Donc $\sum_i B_{ij}\vec{b}_i = L.\vec{b}_j = L.(\sum_k P_{kj}\vec{a}_k) = \sum_k P_{kj} L.\vec{a}_k = \sum_k P_{kj}(\sum_\ell A_{\ell k}\vec{a}_\ell) = \sum_{k\ell} A_{\ell k}P_{kj}(\sum_i Q_{i\ell}\vec{b}_i) = \sum_i(\sum_{k\ell} Q_{i\ell}A_{\ell k}P_{kj})\vec{b}_i = \sum_i (Q.A.P)_{ij}\vec{b}_i$. D'où $B_{ij} = (Q.A.P)_{ij}$, vrai pour tout i, j, donc $B = Q.A.P$. ▟

Exercice 5.5 Montrer : un endomorphisme L peut avoir dans une base $(\vec{a}_i)$ sa matrice $[L]_{\vec{a}}$ symétrique, et peut avoir dans une autre base $(\vec{b}_i)$ sa matrice $[L]_{\vec{b}}$ **non** symétrique.

Réponse. Soit $(\vec{a}_1, \vec{a}_2)$ une base dans $\mathbb{R}^2$, et L défini par $L.\vec{a}_1 = \vec{a}_1$ et $L.\vec{a}_2 = 2\vec{a}_2$ Donc $[L]_{\vec{a}} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ est une matrice symétrique. Soit $\vec{b}_1 = \vec{a}_1$ et $\vec{b}_2 = \vec{a}_1 + \vec{a}_2$, donc $(\vec{b}_1, \vec{b}_2)$ est une base, et $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est la matrice de passage de $(\vec{a}_i)$ à $(\vec{b}_i)$. La formule de changement de base (5.6) donne $[L]_{\vec{b}} = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$, matrice non symétrique. ▟

5.3 Notations tensorielles (relativement à un produit scalaire)

Rappel : le produit tensoriel $f \otimes g$ de deux fonctions $f, g : H \to \mathbb{R}$ est par définition la fonction $f \otimes g : H \times H \to \mathbb{R}$ définie par $(f \otimes g)(x, y) := f(x)g(y)$ pour tout $x, y \in H$. Ici on définit le produit tensoriel relatif à un produit scalaire.

Soit H un espace vectoriel muni d'un produit scalaire $(\cdot,\cdot)_H$.

Définition 5.6 Pour $\vec{u}, \vec{v} \in H$, on note $\vec{u} \otimes_H \vec{v} \in \mathcal{L}(H)$ le $(\cdot,\cdot)_H$-endomorphisme défini relativement à $(\cdot,\cdot)_H$ par :

$$\vec{u} \otimes_H \vec{v} : \begin{cases} H & \to H, \\ \vec{y} & \mapsto (\vec{u} \otimes_H \vec{v}).\vec{y} = (\vec{v}, \vec{y})_H \, \vec{u}. \end{cases} \tag{5.7}$$

(La linéarité est immédiate car $(\cdot,\cdot)_H$ est bilinéaire. Pas de définition ici sans produit scalaire.)

N.B. : attention : cette définition est très dépendante du choix du produit scalaire $(\cdot,\cdot)_H$ sur H (définition non intrinsèque). Par exemple dans $H = P_1$ on pourra prendre $(\cdot,\cdot)_H = (\cdot,\cdot)_{L^2}$, ou $= (\cdot,\cdot)_{H^1}$ ou $= (\cdot,\cdot)_{H^1_0}$, ou... suivant le problème et/ou l'utilisateur.

Exemple 5.7 Soit $(\vec{a}_i^\star)$ la base vectorielle $(\cdot,\cdot)_H$-duale de $(\vec{a}_i)$. Pour $i,j = 1,...,n$, le $(\cdot,\cdot)_H$-endomorphisme $\vec{a}_i \otimes_H \vec{a}_j^\star$ est défini par, pour tout $k = 1,...,n$:

$$(\vec{a}_i \otimes_H \vec{a}_j^\star).\vec{a}_k = \delta_{jk} \, \vec{a}_i. \tag{5.8}$$

Donc la matrice $[\vec{a}_i \otimes_H \vec{a}_j^\star]_{|\vec{a}}$ du $(\cdot,\cdot)_H$-endomorphisme $\vec{a}_i \otimes_H \vec{a}_j^\star$ dans la base $(\vec{a}_i)$: a tous ses éléments nuls sauf l'élément à l'intersection de la ligne i et de la colonne j qui vaut 1. Cela donnera une base de $\mathcal{L}(H)$, voir (5.11). ▲

Exemple 5.8 Soit $M = [(\vec{a}_i, \vec{a}_j)_H]$ la matrice de masse de $(\cdot,\cdot)_H$ pour la base $(\vec{a}_i)$. Pour $i,j = 1,...,n$, le $(\cdot,\cdot)_H$-endomorphisme $\vec{a}_i \otimes_H \vec{a}_j$ est donné par, pour tout $k = 1,...,n$:

$$(\vec{a}_i \otimes_H \vec{a}_j).\vec{a}_k = (\vec{a}_j, \vec{a}_k)_H \, \vec{a}_i = M_{jk} \, \vec{a}_i. \tag{5.9}$$

Donc le k-ème vecteur colonne de la matrice $[\vec{a}_i \otimes_H \vec{a}_j]_{|\vec{a}}$ est réduit à $\begin{pmatrix} 0 \\ . \\ 0 \\ M_{jk} \leftarrow \text{ ligne } i \\ 0 \\ . \\ 0 \end{pmatrix}$: tous ses éléments nuls sauf le i-ème qui vaut M_{jk}). Donc $[\vec{a}_i \otimes_H \vec{a}_j]_{|\vec{a}} = \begin{pmatrix} 0 & \dots & 0 \\ . & & \\ 0 & \dots & 0 \\ M_{j1} & \dots & M_{jn} \\ 0 & \dots & 0 \\ . & & \\ 0 & \dots & 0 \end{pmatrix} \leftarrow$ ligne i : la matrice $[\vec{a}_i \otimes_H \vec{a}_j]_{|\vec{a}}$ a toutes ses lignes nulles sauf la i-ème qui est la j-ème ligne de M. ▲

Exercice 5.9 Montrer que le $(\cdot,\cdot)_H$-produit tensoriel : $\otimes_H : \begin{cases} H \times H & \to \mathcal{L}(H) \\ (\vec{u}, \vec{v}) & \mapsto \otimes_H(\vec{u}, \vec{v}) \stackrel{\text{déf}}{=} \vec{u} \otimes_H \vec{v} \end{cases}$ est bilinéaire. En particulier si $(\vec{a}_i)$ et $(\vec{b}_i)$ sont deux bases, si $\vec{u} = \sum_i u_i \vec{a}_i$ et $\vec{v} = \sum_i v_i \vec{b}_i$, alors :

$$\vec{u} \otimes_H \vec{v} = \sum_{ij} u_i v_j \vec{a}_i \otimes_H \vec{b}_j. \tag{5.10}$$

Réponse. À $\vec{v}$ fixé (5.7) donne trivialement la linéarité en $\vec{u}$, et à $\vec{u}$ fixé (5.7) donne trivialement la linéarité en $\vec{v}$. D'où (5.10). ▲

5.4 Bases de $\mathcal{L}(H)$

Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux bases de H.

Proposition 5.10 *$(\vec{a}_i \otimes_H \vec{b}_j)_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ est une base de $\mathcal{L}(H)$, donc $\mathcal{L}(H)$ est de dimension n^2. En particulier $(\vec{a}_i \otimes_H \vec{a}_j^\star)_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ est une base de $\mathcal{L}(H)$. Et si $L \in \mathcal{L}(H)$ a pour composantes L_{ij} dans la base $(\vec{a}_i)$, i.e. si $L.\vec{a}_j = \sum_{i=1}^n L_{ij} \vec{a}_i$ pour tout j, cf. (5.2), alors les L_{ij} sont aussi les composantes de L dans la base $(\vec{a}_i \otimes_H \vec{a}_j^\star)$:*

$$L = \sum_{i,j=1}^{n} L_{ij} \vec{a}_i \otimes_H \vec{a}_j^\star, \quad \text{quand} \quad [L]_{|\vec{a}} = [L_{ij}]. \tag{5.11}$$

Preuve. Montrons : $(\vec{a}_i \otimes_H \vec{b}_j)_{\substack{i=1,\dots,n\\ j=1,\dots,n}} =^{\text{not.}} (\vec{a}_i \otimes_H \vec{b}_j)$ est libre dans $\mathcal{L}(H)$. Supposons $\sum_{ij} \alpha_{ij}\vec{a}_i \otimes_H \vec{b}_j = 0$ (endomorphisme nul) où les α_{ij} sont des réels. Soit $(\vec{b}_j^\star)$ la base vectorielle duale de $(\vec{b}_j)$. Alors pour $k \in [1,n]_{\mathbb{N}}$ on a $\sum_{ij} \alpha_{ij}(\vec{a}_i \otimes_H \vec{b}_j).\vec{b}_k^\star = \vec{0} = \sum_{ij} \alpha_{ij}\vec{a}_i(\vec{b}_j, \vec{b}_k^\star)_H = \sum_{ij} \alpha_{ij}\vec{a}_i\delta_{jk} = \sum_i \alpha_{ik}\vec{a}_i$, donc $\alpha_{ik} = 0$ pour tout i (car $(\vec{a}_i)$ est une famille libre), ce pour tout k. Donc $\alpha_{ik} = 0$ pour tous i,k. Donc $(\vec{a}_i \otimes_H \vec{b}_j)$ est libre. En particulier $(\vec{a}_i \otimes_H \vec{a}_j^\star)$ est libre dans $\mathcal{L}(H)$.

Montrons : $(\vec{a}_i \otimes_H \vec{a}_j^\star)$ est génératrice dans $\mathcal{L}(H)$. Soit $L \in \mathcal{L}(H)$, $L.\vec{a}_j = \sum_i L_{ij}\vec{a}_i$ pour tout j. Soit $M = \sum_{i,j} L_{ij}\vec{a}_i \otimes_H \vec{a}_j^\star \in \mathcal{L}(H)$. On a $M.\vec{a}_k = \sum_{i,j} L_{ij}\vec{a}_i(\vec{a}_j^\star, \vec{a}_k)_H = \sum_{i,j} L_{ij}\vec{a}_i\delta_{jk} = \sum_i L_{ik}\vec{a}_i = L.\vec{a}_k$, vrai pour $k \in [1,n]_{\mathbb{N}}$. Donc $M = L$, donc $L = \sum_{i,j} L_{ij}\vec{a}_i \otimes_H \vec{a}_j^\star$: $(\vec{a}_i \otimes_H \vec{a}_j^\star)$ est génératrice et on a (5.11).

Donc $(\vec{a}_i \otimes_H \vec{a}_j^\star)$ est une base dans $\mathcal{L}(H)$. Donc $\dim(\mathcal{L}(H)) = n^2$.

Donc $(\vec{a}_i \otimes_H \vec{b}_j)$ famille libre de n^2 éléments : est une base de $\mathcal{L}(H)$. ∎

Proposition 5.11 *Suite. Avec $[L]_{|\vec{a}} = [L_{ij}]$ et avec $M = [(\vec{a}_i, \vec{a}_j)_H]$ on a :*

$$L = \sum_{i,j=1}^n Z_{ij}\vec{a}_i \otimes_H \vec{a}_j \quad \text{où} \quad [Z_{ij}] = [L]_{|\vec{a}}.M^{-1}. \tag{5.12}$$

Et $[Z_{ij}] = [L_{ij}]$ ssi $(\vec{a}_i)_{i=1,\dots,n}$ est une b.o.n. de $(H, (\cdot,\cdot)_H)$.

Preuve. Pour $k \in [1,n]_{\mathbb{N}}$, d'une part $L.\vec{a}_k = \sum_i L_{ik}\vec{a}_i$, et d'autre part $(\sum_{i,j} Z_{ij}\vec{a}_i \otimes_H \vec{a}_j).\vec{a}_k = \sum_{i,j} Z_{ij}\vec{a}_i(\vec{a}_j, \vec{a}_k)_H = \sum_i(\sum_j Z_{ij}M_{kj})\vec{a}_i = \sum_i (Z.M)_{ik}\vec{a}_i$ (car M est symétrique). Vrai pour tout k, d'où $L_{ik} = (Z.M)_{ik}$ pour tous i,k, d'où (5.12). Et $[L] = Z$ ssi $M = I$. ∎

Proposition 5.12 *Avec $[L]_{|\vec{a}} = [L_{ij}]$ et $M = [(\vec{a}_i, \vec{a}_j)_H]$, on a, pour tout i,j,*

$$(L.\vec{a}_j, \vec{a}_i)_H = \sum_{k=1}^n M_{ik}L_{kj} = (M.[L]_{|\vec{a}})_{ij}. \tag{5.13}$$

Donc, pour $\vec{x} = \sum_i x_i\vec{a}_i$ et $\vec{y} = \sum_i y_i\vec{a}_i$,

$$(L.\vec{x}, \vec{y})_H = [\vec{y}]^T.M.[L]_{|\vec{a}}.[\vec{x}] \quad (= \sum_{k,j,i=1}^n x_j y_i M_{ik}L_{kj}). \tag{5.14}$$

(NB : dans notre exemple $(P_1, (\cdot,\cdot)_{L^2})$, cf. § 2.1, il ne faudra pas oublier M dans (5.14), car la base usuelle (φ_{x_i}) n'est pas $(\cdot,\cdot)_{L^2}$-orthonormée : $M \neq I$.)

Preuve. $(L.\vec{a}_j, \vec{a}_i)_H = \sum_{k=1}^n L_{kj}\,(\vec{a}_k, \vec{a}_i)_H = \sum_{k=1}^n M_{ik}L_{kj} = (M.[L])_{ij}$. D'où (5.14) par bilinéarité de $(\cdot,\cdot)_H$. ∎

Exercice 5.13 Soit $M = [(\vec{a}_i, \vec{a}_j)_H])$. Montrer : si $\vec{u} = \sum_i u_i\vec{a}_i$ et $\vec{v} = \sum_i v_i\vec{a}_i$, alors la matrice $[(\vec{u} \otimes_H \vec{v})_{ij}]$ de $\vec{u} \otimes_H \vec{v}$ dans la base $(\vec{a}_i)$ est donnée par :

$$[\vec{u} \otimes_H \vec{v}] = [u_iv_j].M, \quad \text{soit} \quad (\vec{u} \otimes_H \vec{v})_{ij} = \sum_k u_iM_{jk}v_k = u_i(M.[\vec{v}])_j. \tag{5.15}$$

Et $[u_iv_j]$ n'est la matrice de $\vec{u} \otimes_H \vec{v}$ que si $(\vec{a}_i)$ est une b.o.n..

Réponse. On a $\vec{u} \otimes_H \vec{v} = \sum_{ik} u_iv_k\vec{a}_i \otimes_H \vec{a}_k$, cf. (5.10), donc $(\vec{u} \otimes_H \vec{v}).\vec{a}_j = \sum_{ik} u_iv_k\vec{a}_i(\vec{a}_k, \vec{a}_j)_H = \sum_{ik} u_iv_kM_{kj}\vec{a}_i = \sum_i(\sum_k u_iM_{jk}v_k)\vec{a}_i$ car M est symétrique. ∎

5.5 Endomorphisme transposé (relativement à un produit scalaire)

Définition 5.14 Soit $L \in \mathcal{L}(H)$ un endomorphisme. L'application linéaire transposée de L relativement au produit scalaire $(\cdot,\cdot)_H$ est l'endomorphisme $L_H^T \in \mathcal{L}(H)$ défini par, pour tous $\vec{x}, \vec{y} \in H$:

$$(L_H^T.\vec{y}, \vec{x})_H = (\vec{y}, L.\vec{x})_H. \tag{5.16}$$

Remarque 5.15 Malheureusement la définition de L_H^T n'est pas intrinsèque : elle dépend du produit scalaire $(\cdot,\cdot)_H$ choisi dans H. Donc attention au produit scalaire choisi. ∎

Proposition 5.16 *Dans $(H,(\cdot,\cdot)_H)$, il existe un unique endomorphisme L_H^T.*

Soit $(\vec a_i)$ une base dans H et soit $M = [(\vec a_i,\vec a_j)_H]$ la matrice de masse correspondante. Et soit $[L]_{|\vec a} = [L_{ij}] =^{\text{not.}} [L]$ et $[L_H^T]_{|\vec a} = [(L_H^T)_{ij}] =^{noté} [L_H^T]$ les matrices de L et L_H^T dans cette base. On a :

$$[L_H^T] = M^{-1}.[L]^T.M. \tag{5.17}$$

Et $[L_H^T] = [L]^T$ ssi $M.[L] = [L].M$, i.e. ssi les matrices M et $[L]$ commutent.

Preuve. Analyse. Supposons que L_H^T existe. Par définition de L_H^T on a $(L_H^T.\vec a_j,\vec a_i)_H = (L.\vec a_i,\vec a_j)_H$ pour tous i,j. D'où, ayant $L_H^T.\vec a_j = \sum_k (L_H^T)_{kj}\vec a_k$ et $L.\vec a_j = \sum_\ell L_{\ell j}\vec a_\ell$:

$$\sum_k (L_H^T)_{kj}(\vec a_k,\vec a_i)_H = \sum_\ell L_{\ell i}(\vec a_\ell,\vec a_j)_H, \tag{5.18}$$

soit $\sum_k M_{ik}(L_H^T)_{kj} = \sum_\ell M_{j\ell}L_{\ell i}$ par symétrie de M, pour tous i,j. Donc :

$$(M.[L_H^T])_{ij} = (M.[L])_{ji} = ((M.[L])^T)_{ij} = ([L]^T.M)_{ij}, \tag{5.19}$$

car M est symétrique, d'où (5.17). Donc si L_H^T existe alors L_H^T est unique, donnée par (5.17).

Synthèse. Soit Z l'endomorphisme de matrice $A = M^{-1}.[L]^T.M$ dans la base $(\vec a_i)$. Alors :

$(Z.\vec a_j,\vec a_i)_H = \sum_k A_{kj}(\vec a_k,\vec a_i)_H = \sum_k A_{kj}g_{ki} = \sum_k g_{ik}A_{kj} = (M.A)_{ij} = (M.M^{-1}.[L]^T.M)_{ij} = ([L]^T.M)_{ij} = \sum_k L_{ki}M_{kj} = \sum_k L_{ki}(\vec a_k,\vec a_j)_H = (\sum_k L_{ki}\vec a_k,\vec a_j)_H = (L.\vec a_i,\vec a_j)_H$, ce pour tous i,j, d'où (5.16). D'où Z vérifie(5.16) : c'est donc L_H^T.

Si $M.[L] = [L].M$ alors $(M.[L])^T = ([L].M)^T$ et donc $[L]^T.M = M.[L]^T$ car M est symétrique, donc M et $[L]^T$ commutent. D'où $[L_H^T] = M^{-1}.[L]^T.M = M^{-1}.M.[L]^T = [L]^T$.

Si $M.[L] \ne [L].M$, alors $[L]^T.M \ne M.[L]^T$, d'où $M^{-1}.[L]^T.M \ne [L]^T$, d'où $[L_H^T] \ne [L]^T$. ∎

5.6 Endomorphisme symétrique (relativement à un produit scalaire)

Définition 5.17 L est dite $(\cdot,\cdot)_H$-symétrique, ou symétrique relativement au produit scalaire $(\cdot,\cdot)_H$ ssi, pour tous $\vec x,\vec y \in H$:

$$(L.\vec y,\vec x)_H = (\vec y, L.\vec x)_H, \tag{5.20}$$

i.e. ssi, cf. (5.16) :

$$L_H^T = L. \tag{5.21}$$

Attention : le caractère symétrique dépend du choix du produit scalaire $(\cdot,\cdot)_H$. Voir exercice 5.19.

Proposition 5.18 *Soit $(\vec a_i)$ une base. L est $(\cdot,\cdot)_H$-symétrique ssi pour tous $i,j = 1,...,n$:*

$$(L.\vec a_i,\vec a_j)_H = (\vec a_i, L.\vec a_j)_H, \quad \text{i.e. ssi} \quad M.[L]_{|\vec a} = [L]_{|\vec a}^T.M, \tag{5.22}$$

où $M = [(\cdot,\cdot)_H]_{|\vec a}$.

Preuve. On applique (5.17). ∎

Exercice 5.19 On se fixe un produit scalaire $(\cdot,\cdot)_H$.

1- Donner un exemple d'endomorphisme L $(\cdot,\cdot)_H$-symétrique dont la matrice n'est pas symétrique.

2- Montrer : si la matrice de L est $(\cdot,\cdot)_H$-symétrique dans une $(\cdot,\cdot)_H$-b.o.n., alors elle est symétrique dans toute autre $(\cdot,\cdot)_H$-b.o.n..

3- Montrer : la symétrie d'un endomorphisme L est dépendante du choix du produit scalaire : on peut avoir L symétrique pour un produit scalaire $(\cdot,\cdot)_H$ donné, et L non symétrique pour un autre produit scalaire $(\cdot,\cdot)_h$ donné :

$$L_H^T = L \quad \text{et} \quad L_h^T \ne L. \tag{5.23}$$

Réponse. Suite de l'exemple 5.5 : $[L]_{|\vec a} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $\vec b_1 = \vec a_1$ et $\vec b_2 = \vec a_1 + \vec a_2$, donc $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $[L]_{|\vec b} = P^{-1}.[L]_{|\vec a}.P = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ qui n'est pas symétrique.

1- On prend $L \in \mathcal{L}(\mathbb{R}^2)$, $(\vec a_i) = (\vec e_i)$ la base canonique et $(\cdot,\cdot)_H = (\cdot,\cdot)_{\mathbb{R}^2}$ le produit scalaire euclidien.

Ici L est $(\cdot,\cdot)_{\mathbb{R}^2}$-symétrique : sur la base canonique $(\vec{e}_1,\vec{e}_2)$ on a $(L.\vec{e}_i,\vec{e}_j)_{\mathbb{R}^2} = (L.\vec{e}_j,\vec{e}_i)_{\mathbb{R}^2} = ([L]_{|\vec{e}})_{ij}$ (la matrice $[L]_{|\vec{e}}$ est symétrique). Mais la matrice $[L]_{\vec{b}}$, de L dans la base $(\vec{b}_1,\vec{b}_2)$, n'est pas symétrique (la matrice dépend de la base), bien que L est $(\cdot,\cdot)_{\mathbb{R}^2}$-symétrique.

2- Soit $(\vec{a}_i)$ et $(\vec{b}_i)$ deux b.o.n., donc $(\vec{a}_i,\vec{a}_j)_H = \delta_{ij}$ et $(\vec{b}_i,\vec{b}_j)_H = \delta_{ij}$ pour tous i,j. Soit P la matrice de passage de $(\vec{a}_i)$ à $(\vec{b}_i)$, donc $\vec{b}_j = \sum_i P_{ij}\vec{a}_i$ pour tout j.

On a $P^T.P = I$ car $\delta_{ij} = (\vec{b}_i,\vec{b}_j)_H = \sum_{k\ell} P_{ki}P_{\ell j}(\vec{a}_k,\vec{a}_\ell)_H = \sum_{k\ell} P_{ki}P_{\ell j}\delta_{k\ell} = \sum_k P_{ki}P_{kj} = (P^T.P)_{ij}$. Donc $P^{-1} = P^T$.

Et formule de changement de base des endomorphismes : $[L]_{\vec{b}} = P^{-1}.[L]_{\vec{a}}.P$. Donc si $[L]_{\vec{a}}$ symétrique alors $[L]_{\vec{b}}^T = P^T.[L]_{\vec{a}}^T.(P^{-1})^T = P^{-1}.[L]_{\vec{a}}.P = [L]_{\vec{b}}$, et $[L]_{\vec{b}}$ est bien symétrique.

3- On reprend le cadre $\mathbb{R}^2$ du 1-, et $(\cdot,\cdot)_H = (\cdot,\cdot)_{\mathbb{R}^2}$ le produit scalaire euclidien. La matrice de $(\cdot,\cdot)_H$ dans la base $(\vec{e}_i)$ est donc la matrice identité $[(\cdot,\cdot)_H] = I$. On a vu que L est $(\cdot,\cdot)_{\mathbb{R}^2}$-symétrique.

Soit $h(\cdot,\cdot) = (\cdot,\cdot)_h$ le produit scalaire défini dans la base $(\vec{e}_i)$ par sa matrice $[h] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ dans la base canonique. Et $h(\cdot,\cdot)$ est un produit scalaire : symétrie OK car $(h(\vec{e}_i,\vec{e}_j) = h(\vec{e}_j,\vec{e}_i)$, soit $h_{ij} = h_{ji}$ pour tous $i,j)$, et valeurs propres strictement positives car $\mathrm{Tr}([h]) > 0$ et $\det([h]) > 0$.

On a $(L_h^T.\vec{e}_2,\vec{e}_1)_h = (L.\vec{e}_1,\vec{e}_2)_h = (\vec{e}_1,\vec{e}_2)_h = \begin{pmatrix} 1 & 0 \end{pmatrix}.\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}.\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1$, alors que $(L_h.\vec{e}_2,\vec{e}_1)_h = (2\vec{e}_2,\vec{e}_1)_h = 2\begin{pmatrix} 0 & 1 \end{pmatrix}.\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 0 & 1 \end{pmatrix}.\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2$, donc $L_h^T \neq L$.

Ou encore on applique (5.14) avec ici $M = [h]$, et $M.[L] = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}$ n'est pas symétrique. ∎

5.7 Diagonalisation des endomorphismes symétriques

Définition 5.20 Soit $L \in \mathcal{L}(H)$ un endomorphisme. Un couple $(\lambda,\vec{v}) \in \mathbb{R}\times H$ est un couple "(valeur propre , vecteur propre)" de L ssi :

$$\vec{v} \neq \vec{0} \quad \text{et} \quad L.\vec{v} = \lambda\vec{v}. \tag{5.24}$$

Théorème 5.21 *Soit H un espace vectoriel réel de dimension n muni d'un produit scalaire $(\cdot,\cdot)_H$. Soit $L \in \mathcal{L}(H)$. Si l'endomorphisme L est $(\cdot,\cdot)_H$-symétrique, alors L est diagonalisable dans une b.o.n. de $(H,(\cdot,\cdot)_H)$. I.e. : si $\forall\vec{x},\vec{y}\in H$ on a $(L.\vec{x},\vec{y})_H = (\vec{x},L.\vec{y})_H$, alors :*

$$\left\{\begin{array}{l} \exists\lambda_1,...,\lambda_n \in \mathbb{R}, \\ \exists\vec{p}_1,....,\vec{p}_n \in H, \end{array}\right\} \quad \text{t.q.} \quad \forall i,j = 1,...,n, \quad L.\vec{p}_j = \lambda_j\vec{p}_j \quad \text{et} \quad (\vec{p}_i,\vec{p}_j)_H = \delta_{ij}. \tag{5.25}$$

Et donc, $(\vec{p}_i)$ étant une $(\cdot,\cdot)_H$-b.o.n.,

$$L = \sum_{i=1}^n \lambda_i\,\vec{p}_i \otimes_H \vec{p}_i. \tag{5.26}$$

Preuve. Soit $(\vec{b}_i)$ une $(\cdot,\cdot)_H$-b.o.n dans H. Ainsi la matrice $[L]_{|\vec{b}}$ de L dans la base $(\vec{b}_i)$ est symétrique, cf. (5.22). Donc avec $\mathbb{R}^n$ muni de sa base canonique $(\vec{e}_i)$ et de son produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$, la matrice $A = [L]_{|\vec{b}}$ est diagonalisable dans une b.o.n. $(\vec{v}_i)$ de $\mathbb{R}^n$: $\exists\lambda_1,....,\lambda_n \in \mathbb{R}$, $\exists\vec{v}_1,...\vec{v}_n \in \mathbb{R}$ t.q. $A.\vec{v}_i = \lambda_i\vec{v}_i$ et $(\vec{v}_i,\vec{v}_j)_{\mathbb{R}^n} = \delta_{ij}$ pour tout i,j. Notons $P = [P_{ij}]$ la matrice de passage de la base $(\vec{e}_i)$ à la base $(\vec{v}_i)$, où donc $\vec{v}_j = \sum_i P_{ij}\vec{e}_i$. Posons

$$[\vec{p}_j]_{|\vec{b}} = [\vec{v}_j]_{|\vec{e}}, \quad \text{i.e.} \quad \vec{p}_j = \sum_i P_{ij}\vec{b}_i. \tag{5.27}$$

Donc :

$$[L]_{|\vec{b}}.[\vec{p}_j]_{|\vec{b}} = \lambda_j[\vec{p}_j]_{|\vec{b}} \quad \text{avec} \quad [\vec{p}_j]_{|\vec{b}}^T.[\vec{p}_j]_{|\vec{b}} = I, \tag{5.28}$$

donc, pour tout i,j :

$$L.\vec{p}_j = \lambda_j\vec{p}_j, \quad \text{avec} \quad (\vec{p}_i,\vec{p}_j)_H = \delta_{ij} \tag{5.29}$$

car $(\vec{b}_i)$ est une $(\cdot,\cdot)_H$-b.o.n.. Donc $L = \sum_{i=1}^n \lambda_i\,\vec{p}_i \otimes_H \vec{p}_i^{\star}$, et ici $\vec{p}_i^{\star} = \vec{p}_i$ (cas d'une b.o.n.), d'où (5.26). ∎

5.8 Projection et espace propre associé à la valeur propre $\lambda = 1$

5.8.1 Projection

Soit H un espace vectoriel de dimension n (finie) et $L \in \mathcal{L}(H)$. Rappel $\mathrm{Ker}L := \{\vec{x} \in H : L.\vec{x} = \vec{0}\}$ (noyau) et $\mathrm{Im}L := \{\vec{y} \in H : \exists \vec{x} \in H \text{ t.q. } L.\vec{x} = \vec{y}\}$ (image).

Définition 5.22 Une projection de H dans H est un endomorphisme $P \in \mathcal{L}(H)$ qui vérifie :

$$P \circ P = P. \tag{5.30}$$

Proposition 5.23 *Soit P une projection non nulle (il existe $\vec{v} \in H$ t.q. $P.\vec{v} \neq \vec{0}$).*
1- $\vec{v} \in \mathrm{Im}P$ ssi $P.\vec{v} = \vec{v}$ (la restriction de P à $\mathrm{Im}P$ est l'identité).
2- $\mathrm{Ker}P \cap \mathrm{Im}P = \{\vec{0}\}$.
3-

$$H = \mathrm{Ker}P \oplus \mathrm{Im}P. \tag{5.31}$$

Et on dit que P est la projection sur $\mathrm{Im}P$ parallèlement à $\mathrm{Ker}P$.
4- P est diagonalisable : P n'a que deux valeurs propres 0 et 1 associées respectivement aux espaces propres $\mathrm{Ker}P$ et $\mathrm{Im}P$.

Preuve. 1- Si $\vec{v} = P(\vec{v})$, alors $\vec{v} \in \mathrm{Im}P$. Et si $\vec{v} \in \mathrm{Im}P$ alors il existe $\vec{u} \in H$ tel que $P(\vec{u}) = \vec{v}$. Et $P(P(\vec{u})) = P(\vec{u})$ car P est une projection, donc $P(\vec{v}) = \vec{v}$. Donc $P_{|\mathrm{Im}P} = I$.

2- Soit $\vec{v} \in \mathrm{Ker}P \cap \mathrm{Im}P$: donc $P(\vec{v}) = \vec{0}$ et $\exists \vec{u}$ t.q. $\vec{v} = P(\vec{u})$. Donc $\vec{0} = P(\vec{v}) = P(P(\vec{u})) = P(\vec{u}) = \vec{v}$: on a $\vec{v} = 0$.

3- Soit $\vec{v} \in H$. On a $\vec{v} - P.\vec{v} \in \mathrm{Ker}P$ car $P.(\vec{v} - P.\vec{v}) = P.\vec{v} - (P \circ P).\vec{v} = P.\vec{v} - P.\vec{v} = 0$. Donc $\vec{v} = P\vec{v} + (\vec{v} - P.\vec{v}) \in \mathrm{Im}P + \mathrm{Ker}P$, donc $H = \mathrm{Ker}P + \mathrm{Im}P$. Et $\mathrm{Ker}P \cap \mathrm{Im}P = \{\vec{0}\}$ donne $H = \mathrm{Ker}P \oplus \mathrm{Im}P$.

4- Si $\vec{v} \in \mathrm{Ker}P$, alors $P(\vec{v}) = 0$: donc tout vecteur de $\mathrm{Ker}P$ est vecteur propre associé à la valeur propre 0. On peut en choisir $n_K = \dim(\mathrm{Ker}P)$ indépendants.

Si $\vec{v} \in \mathrm{Im}P$, alors $P.\vec{v} = \vec{v}$. Donc tout vecteur de $\mathrm{Im}P$ est vecteur propre associé à la valeur propre 1. On peut en choisir $\dim(\mathrm{Im}P) = n - n_K$ indépendants car $n = \dim(H) = \dim(\mathrm{Ker}P \oplus \mathrm{Im}P) = \dim(\mathrm{Ker}P) + \dim(\mathrm{Im}P)$: On a ainsi obtenu une base de vecteurs propres : P est diagonalisable, et il n'y a pas d'autre valeur propre (0 est valeur propre de multiplicité n_K et 1 est valeur propre de multiplicité $n - n_k$). ∎

Exercice 5.24 Vérifier : si H est de dimension finie n, si A est la matrice d'une projection P, alors $A^2 = A$, et les deux seules valeurs propres possibles de A sont 0 et 1.

Réponse. Soit $(\vec{a}_i)$ une base et A la matrice de P dans cette base, soit $P.\vec{a}_j = \sum_{i=1}^n A_{ij}\vec{a}_i$ pour tout j. Alors pour tout $j = 1, ..., n$ on a $P(P.\vec{a}_j) = P.\vec{a}_j$, soit $P(\sum_{i=1}^n A_{ij}\vec{a}_i) = \sum_{k=1}^n A_{kj}\vec{a}_k$, donc $\sum_{i,k=1}^n A_{ij}A_{ki}.\vec{a}_k = \sum_{k=1}^n A_{kj}\vec{a}_k$, donc $(A.A)_{kj} = A_{kj}$, ce pour tous k, j, donc $A^2 = A$.

Si $\lambda \in \mathbb{R}$ et $\vec{x} \in \mathbb{R}^n - \{\vec{0}\}$ vérifient $A.\vec{x} = \lambda\vec{x}$, alors $A^2.\vec{x} = A.(A.\vec{x}) = A.(\lambda\vec{x}) = \lambda A.\vec{x} = \lambda^2\vec{x}$, avec $A^2 = A$, donc $\lambda\vec{x} = \lambda^2\vec{x}$, donc $\lambda = \lambda^2$, donc $\lambda = 0$ ou 1. ∎

5.8.2 Projection orthogonale

Définition 5.25 On munit H d'un produit scalaire $(\cdot,\cdot)_H$. Une projection $(\cdot,\cdot)_H$orthogonale, ou projection orthogonale relativement au produit scalaire $(\cdot,\cdot)_H$, est une projection telle que :

$$\forall \vec{v}_k \in \mathrm{Ker}P,\ \forall \vec{v}_i \in \mathrm{Im}P,\ (\vec{v}_k, \vec{v}_i)_H = 0, \quad \text{i.e. t.q.} \quad \mathrm{Im}P \perp_H \mathrm{Ker}P, \tag{5.32}$$

et, avec (5.31), on note $H = \mathrm{Ker}P \overset{\perp_H}{\oplus} \mathrm{Im}P$.

Proposition 5.26 *Une projection $(\cdot,\cdot)_H$-orthogonale est diagonalisable dans une $(\cdot,\cdot)_H$-b.o.n..*
En particulier une projection $(\cdot,\cdot)_H$-orthogonale est un endomorphisme $(\cdot,\cdot)_H$-symétrique.

Preuve. La projection étant orthogonale, on prend une b.o.n. dans $\mathrm{Ker}L$ et une dans $\mathrm{Im}L$, et on obtient une b.o.n. faite de vecteurs propres.

Et un endomorphisme L diagonalisable dans une $(\cdot,\cdot)_H$-b.o.n. $(\vec{v}_i)$ est nécessairement $(\cdot,\cdot)_H$-symétrique : on a $(L.\vec{v}_i, \vec{v}_j)_H = (\lambda_i\vec{v}_i, \vec{v}_j)_H = \lambda_i\delta_{ij} = (\vec{v}_i, \lambda_j\vec{v}_j)_H$, donc $(L.\vec{v}_i, \vec{v}_j)_H = (\vec{v}_i, L.\vec{v}_j)_H$, ce pour tous i, j : donc L est $(\cdot,\cdot)_H$-symétrique : $L_H^T = L$. ∎

Exercice 5.27 Montrer que si P est une projection, on peut toujours trouver un produit scalaire $(\cdot,\cdot)_H$ sur H telle que P est une projection orthogonale.

Réponse. Soit $(\vec{a}_i)_{i=1,\dots,n_K}$ une base de Ker$P$ et soit $(\vec{b}_i)_{i=1,\dots,n_I}$ une base de ImP. Comme $H = \mathrm{Ker}P \oplus \mathrm{Im}P$, la famille $(\vec{c}_i)_{i=1,\dots,n}$ définie par $\vec{c}_i = \vec{a}_i$ pour tout $i = 1,\dots,n_K$ et $\vec{c}_{n_K+i} = \vec{b}_i$ pour tout $i = 1,\dots,n_I$ est une base de H. On définit alors la forme bilinéaire $g(\cdot,\cdot)$ par $g(\vec{c}_i,\vec{c}_j) = \delta_{ij}$ pour tous $i,j = 1,\dots,n$. On vérifie immédiatement que $g(\cdot,\cdot) =^{\text{noté}} (\cdot,\cdot)_H$ ainsi définie est symétrique définie positive. ▪

5.8.3 Approximée de $f \in L^2(]a,b[)$ par $f_h \in P_1$

Soit $f \in L^2(]a,b[)$ donnée. Comment trouver la fonction $f_h \in P_1$ qui approche “au mieux” f ?

Si on dispose d'une distance on veut minimiser la distance entre f et P_1. Ici on dispose de l'instrument de mesure $||.||_{L^2}$, donc : on veut trouver la fonction $f_h \in P_1$ qui minimise $||f - f_h||_{L^2}$.

P_1 étant un sous-espace vectoriel de $L^2(]a,b[)$ de dimension finie, il est fermé, donc f_h existe et est unique : c'est la projection $(\cdot,\cdot)_{L^2}$-orthogonale de f sur P_1, i.e. est solution de

$$(f_h, g_h)_{L^2} = (f, g_h)_{L^2}, \quad \forall g_h \in P_1, \tag{5.33}$$

ou si on préfère $(f - f_h, g_h)_{L^2} = 0$ pour tout $g_h \in P_1$, i.e. $f - f_h \perp_{L^2} P_1$, faire un dessin.

Calcul : soit (φ_i) une base de P_1 (par exemple la base des fonctions chapeau), et

$$f_h = \sum_{j=1}^{n} y_j \varphi_j, \tag{5.34}$$

où donc les y_j sont les composantes de f_h sur la base (φ_i). Connaître f_h c'est connaître les n y_j (les inconnues sont $y_1,\dots,y_n$).

- 1ère méthode : dans (5.33), les $g_h = \varphi_i$ donnent les n équations :

$$\sum_{j=1}^{n} y_j (\varphi_j, \varphi_i)_{L^2} = (f, \varphi_i)_{L^2}, \quad \forall i = 1,\dots,n,$$

soit

$$M.[\vec{y}] = [\vec{b}], \quad \text{où} \quad M = [(\varphi_i,\varphi_j)_{L^2}], \quad [\vec{y}] = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad [\vec{b}] = \begin{pmatrix} (f,\varphi_1)_{L^2} \\ \vdots \\ (f,\varphi_n)_{L^2} \end{pmatrix}. \tag{5.35}$$

La résolution de ce système donne les y_i, d'où $f_h = \sum_i y_i \varphi_i$.

- 2ème méthode : si on connaît la base duale vectorielle $(\varphi_i^\star)$ alors (3.1) donne directement :

$$y_i = (f_h, \varphi_i^\star)_{L^2} = (f, \varphi_i^\star)_{L^2} = \int_a^b \varphi_i^\star(x) f(x)\,dx = \int_a^b K(x_i, x) f(x)\,dx, \tag{5.36}$$

d'où $f_h = \sum_i y_i \varphi_{x_i}$.

Le coût de calcul de la base duale vectorielle est le coût d'inversion de M : C'est le coût d'une diagonalisation, d'où la :

- 3ème méthode : on connaît une b.o.n. (ζ_i) de diagonalisation, cf. (2.26). Alors, notant $f_h = \sum_{j=1}^n z_j \zeta_j$, (5.33) donne directement

$$z_i = (f_h, \zeta_i)_{L^2} \stackrel{donc}{=} (f, \zeta_i)_{L^2}, \quad \forall i = 1,\dots,n. \tag{5.37}$$

6 Application linéaire continue (opérateur borné)

Rappel : si $(E, ||.||_E)$ et $(F, ||.||_F)$ sont deux espaces de Banach (espace vectoriel normés complets), l'application linéaire $L : (E, ||.||_E) \to (F, ||.||_F)$ est continue dans E ssi :

$$\exists C > 0, \quad \forall \vec{u} \in E, \quad ||L.\vec{u}||_F \le C||\vec{u}||_E. \tag{6.1}$$

Ce qui revient à dire que, L étant linéaire, L est continue en $\vec{0}$, ou encore bornée sur la boule unité : une application linéaire continue est aussi appelée un opérateur borné.

Et on note $||L||_{\mathcal{L}(E,F)} = ||L||$ la plus petite des constantes C vérifiant (6.1), donc :

$$||L|| = \sup_{||\vec{u}||_E=1} ||L.\vec{u}||_F = \sup_{\vec{u}\neq 0} \frac{||L.\vec{u}||_F}{||\vec{u}||_E} \quad (= \sup_{\vec{u}\neq 0} ||L(\frac{\vec{u}}{||\vec{u}||_E})||_F). \tag{6.2}$$

Donc, si L est un opérateur borné, pour tout $\vec{u} \in E$ on a $||L.\vec{u}||_F \leq ||L||\,||\vec{u}||_E$.

Notation. Cas $(E, ||.||_E) = (F, ||.||_F) = (H, (\cdot,\cdot)_H)$ Hilbert. On notera :

$$\begin{aligned} \mathcal{L}(H) &= \{L : H \to H \text{ t.q. } L \text{ linéaire } \underline{\text{et}} \text{ bornée}\} \\ &= \{L : H \to H \text{ linéaire t.q. } ||L|| < \infty\} \end{aligned} \tag{6.3}$$

l'ensemble des opérateurs bornés (endomorphismes continus) de H dans H, où $||L|| = \sup_{||\vec{u}||_H=1} ||L.\vec{u}||_H$.

Rappel : si E et F sont de dimension finie alors toutes les applications linéaires sont continues. Démonstration immédiate à l'aide de bases et des matrices. C'est faux en dimension infinie : voir exemple 6.1.

Exemple 6.1 Soit $H = \ell^2 = \{(\vec{x} = (x_n)_{\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*}$ t.q. $\sum_{n=1}^\infty x_n^2 < \infty\}$ muni du produit scalaire canonique $(\vec{x},\vec{y})_{\ell^2} = \sum_{n=1}^\infty x_n y_n$ et de la norme associée $||\vec{x}||_{\ell^2} = (\sum_{i=1}^n x_n^2)^{\frac{1}{2}}$. Ainsi ℓ^2 est un Hilbert. On note $(\vec{e}_n)$ la base canonique de ℓ^2 (donc $\vec{e}_1 = (1,0,0,...)$, $\vec{e}_2 = (0,1,0,...)$, ...,et $(\vec{e}_i)$ est une $(\cdot,\cdot)_{\ell^2}$-b.o.n., trivial).

Soit $L : \ell^2 \to \ell^2$ linéaire définie par $L.\vec{e}_n = \lambda_n \vec{e}_n$ pour tout n. Sa matrice généralisée est la matrice diagonale $[L]_{|\vec{e}} = \mathrm{diag}(\lambda_n)_{n\in\mathbb{N}^*} =^{\text{noté}} \mathrm{diag}(\lambda_n)$. Et $L.\vec{x} = \sum_n \lambda_n x_n \vec{e}_n$ quand $\vec{x} = (x_n)_{\mathbb{N}^*} = \sum_n x_n \vec{e}_n$.

Si la suite $(\lambda_n)_{\mathbb{N}^*}$ est bornée, alors L est bornée de norme $||L|| = \sup_{\mathbb{N}^*}(|\lambda_n|)$, puisque $\vec{x} = \sum_n x_n \vec{e}_n$, $L.\vec{x} = \sum_n \lambda_n x_n \vec{e}_n$ et Pythagore généralisé donnent $||L.\vec{x}||^2 = \sum_{n=1}^\infty \lambda_n^2 x_n^2 \leq (\sup_{\mathbb{N}^*} \lambda_n^2) \sum_{n=1}^\infty x_n^2$, et que $||L.\vec{e}_n|| = |\lambda_n|$ permet de s'approcher aussi près que souhaité de $\sup_{\mathbb{N}^*}(|\lambda_n|)$.

Si la suite $(\lambda_n)_{\mathbb{N}^*}$ n'est pas bornée alors L n'est pas borné : on prend une sous suite $(\lambda_{n_k})_{k\in\mathbb{N}^*}$ telle que $|\lambda_{n_k}| \longrightarrow_{k\to\infty} \infty$, et on a $||L.\vec{e}_{n_k}|| \longrightarrow_{k\to\infty} \infty$. (Ici $L : \ell^2 \to \mathbb{R}^{\mathbb{N}^*}$ n'est pas un endomorphisme de ℓ^2 : il n'est pas à valeurs dans ℓ^2. Ou encore, si on veut $\mathrm{Im}(L) \subset \ell^2$, le domaine de définition de L est strictement inclus dans ℓ^2.) ∴

Remarque 6.2 Il n'existe pas d'exemple explicite d'endomorphisme non borné d'un Hilbert dans lui-même au sens où : si H est Hilbert, si $L \in \mathcal{L}(H)$ (où donc l'ensemble de définition est H tout entier et l'image ImL est incluse dans H) est donné explicitement, alors L est borné.

Pour trouver un endomorphisme non borné dans un Hilbert (donc avec $||L|| = \infty$), il faut faire appel à l'axiome du choix. Résultat démontré par Solovay (plus généralement dans un Banach), voir J. D. Maitland Wright [5].

Donc dans la pratique, quand on considèrera un endomorphisme dans un Hilbert H, il sera borné, et on pourra calculer (ou en tout cas estimer) sa norme : les seules hypothèses à vérifier seront que L est défini sur tout H, L est linéaire, et Im$L \subset H$. ∴

Deuxième partie

Noyau intégral

7 Dans ℓ^2 : noyau intégral discret = matrice généralisée

7.1 Produit scalaire dans ℓ^2

L'ensemble de suites réelles $\vec{x} = (x_i)_{i\in\mathbb{N}^*}$ est noté $\mathbb{R}^{\mathbb{N}^*} = \mathbb{R}\times\mathbb{R}\times...$ produit cartésien infini dénombrable. $(\mathbb{R}^{\mathbb{N}^*}, +, .) =^{\text{noté}} \mathbb{R}^{\mathbb{N}^*}$ est un espace vectoriel dont la base canonique est la base $(\vec{e}_i)_{i\in\mathbb{N}^*}$ donnée par $\vec{e}_1 = (1,0,0,...)$, $\vec{e}_2 = (0,1,0,...)$, ..., extension de la base canonique de $\mathbb{R}^n$ pour n infiniment grand.

ℓ^2 est le sous-ensemble de $\mathbb{R}^{\mathbb{N}^*}$ constitué des suites de carrés sommables (suites "d'énergie finie") :

$$\ell^2 = \{\vec{x} = (x_i)_{i\in\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*} \text{ t.q. } \sum_{i=1}^\infty x_i^2 < \infty\}. \tag{7.1}$$

C'est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}^*}$. Tous les $\vec{e}_i$ de la base canonique de $\mathbb{R}^{\mathbb{N}^*}$ sont trivialement dans ℓ^2, et $(\vec{e}_i)_{i\in\mathbb{N}^*}$ est la base (encore dite) canonique de ℓ^2. La condition $\sum_{i=1}^\infty x_i^2 < \infty$ permet de munir ℓ^2 du produit scalaire $(\cdot,\cdot)_{\ell^2}$ (euclidien généralisé) et de la norme $||.||_{\ell^2}$ associée (Pythagore généralisé) : pour $\vec{x} = (x_i)_{i\in\mathbb{N}^*} = \sum_{i=1}^\infty x_i\vec{e}_i$ et $\vec{y} = (y_i)_{i\in\mathbb{N}^*} = \sum_{i=1}^\infty y_i\vec{e}_i$ dans ℓ^2 :

$$(\vec{x},\vec{y})_{\ell^2} = \sum_{i=1}^\infty x_iy_i \quad \text{et} \quad ||\vec{x}||_{\ell^2} = (\sum_{i=1}^\infty x_i^2)^{\frac12}. \tag{7.2}$$

Dans la suite on utilisera exclusivement la base canonique $(\vec{e}_i)$ de ℓ^2, et on notera $[.]_{|\vec{e}} =^{\text{not.}} [L]$.

On renvoie au cours d'éléments finis pour la démonstration de : $(\ell^2, (\cdot,\cdot)_{\ell^2})$ est "un espace de Hilbert" : $(\cdot,\cdot)_{\ell^2}$ est bien un produit scalaire, et ℓ^2 est complet pour la norme associée $||.||_{\ell^2}$.

7.2 Endomorphisme de ℓ^2 : sa matrice $\to$ son noyau intégral

Soit $L \in \mathcal{L}(\ell^2)$ un endomorphisme de ℓ^2. Pour $j \in \mathbb{N}^*$, on note $L_{ij} = K(i,j)$ les composantes de $L.\vec{e}_j$ dans la base canonique $(\vec{e}_i)$:

$$L.\vec{e}_j = \sum_{i=1}^\infty L_{ij}\,\vec{e}_i = \sum_{i=1}^\infty K(i,j)\,\vec{e}_i, \quad [L.\vec{e}_j] = \begin{pmatrix} L_{1j} \\ \vdots \\ L_{nj} \\ \vdots \end{pmatrix} = \begin{pmatrix} K(1,j) \\ \vdots \\ K(n,j) \\ \vdots \end{pmatrix} = [L].[\vec{e}_j]. \tag{7.3}$$

(On rappelle que $[.]_{|\vec{e}} =^{\text{not.}} [.]$.)

Définition 7.1 La matrice (généralisée) $[L]_{|\vec{e}} = [L]$ de L dans la base $(\vec{e}_i)$ est le tableau de réels :

$$[L] = [L_{ij}]_{\substack{i\in\mathbb{N}^*\\ j\in\mathbb{N}^*}} = [K(i,j)]_{\substack{i\in\mathbb{N}^*\\ j\in\mathbb{N}^*}} = \begin{pmatrix} L_{11} & L_{12} & \dots \\ L_{21} & L_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} K(1,1) & K(1,2) & \dots \\ K(2,1) & K(2,2) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{7.4}$$

(La j-ème colonne $[L.\vec{e}_j] = [L].[\vec{e}_j]$ de $[L]$ donne les composantes de $L.\vec{e}_j$ dans la base $(\vec{e}_i)$.)

Par linéarité de L, pour $\vec{x} = (x_i)_{i\in\mathbb{N}^*} = \sum_{i=1}^\infty x_i\vec{e}_i \in \ell^2$, on a

$$L.\vec{x} = \sum_{j=1}^\infty x_j\,L.\vec{e}_j = \sum_{i=1}^\infty\sum_{j=1}^\infty L_{ij}x_j\,\vec{e}_i, \quad \text{donc} \quad [L.\vec{x}] = \begin{pmatrix} \sum_j L_{1j}x_j \\ \vdots \\ \sum_j L_{nj}x_j \\ \vdots \end{pmatrix} = [L].[\vec{x}], \tag{7.5}$$

à comparer avec (5.3). Et on note $\vec{x} = (x_i)_{i\in\mathbb{N}^*} = \sum_{i=1}^\infty x_i\vec{e}_i =^{\text{not.}} \sum_{i=1}^\infty x(i)\vec{e}_i \in \ell^2$, et $(L.\vec{x})_i =^{\text{not.}} (L.\vec{x})(i)$ la i-ème composante de $L.\vec{x}$ dans la base $(\vec{e}_i)$:

$$(L.\vec{x})(i) = \sum_{j=1}^\infty L_{ij}x_j = \sum_{j=1}^\infty K(i,j)x(j) = \sum_{j=1}^\infty K_i(j)x(j), \quad \text{donc} \quad = (\vec{K}_i, \vec{x})_{\ell^2}, \tag{7.6}$$

où

$$K_i(j) := K(i,j) = L_{ij}, \quad \text{et} \quad \vec{K}_i = (K_i(1), K_i(2), \dots) = \sum_{j=1}^\infty (K_i(j)\vec{e}_j = \sum_{j=1}^\infty L_{ij}\vec{e}_j. \tag{7.7}$$

Et la matrice $K = [K(i,j)]$ est le noyau intégral discret.

Remarque : (7.6) deviendra pour les fonctions $f \in L^2(]a,b[)$:

$$(L.f)(x) = \int_{y=a}^b K(x,y)f(\vec{y})\,dy = \int_{y=a}^b K_x(y)f(y)\,dy = (K_x, f)_{L^2}. \tag{7.8}$$

Et la fonction $K : (x,y) \to K(x,y)$ sera le noyau intégral.

Remarque 7.2 Dans (7.6) la mesure utilisée est la mesure $\mu = \sum_{j=1}^\infty \delta_j$ de comptage :
$(L.\vec{x})(i) = \sum_{j=1}^\infty K_i(j)x(j) = \sum_{j=1}^\infty (K_ix)(j) = \mu(K_i\vec{x})$.
Dans (7.8) la mesure utilisée est la mesure $\mu_\ell = d\Omega$ de Lebesgue :
$(L.f)(x) = \int_{y=a}^b K_x(y)f(y)\,dy = \int_{y=a}^b (K_xf)(y)\,dy = \mu_\ell(K_xf)$. ▪

8 Opérateur de Hilbert–Schmidt

8.1 Rappel : norme de Frobenius dans $\mathbb{R}^n$

Soit une matrice $n*n$ carrée $A = [A_{ij}]$ ett le vecteur colonne associé $\vec{V} = \begin{pmatrix} A_{11} \\ \vdots \\ A_{n1} \\ A_{12} \\ \vdots \\ A_{nn} \end{pmatrix} = \begin{pmatrix} V_1 \\ \vdots \\ V_{n^2} \end{pmatrix} \in \mathbb{R}^{n^2}$: on a mis les colonnes de L les unes à la suite des autres. La norme de Frobenius $||A||_2$ de la matrice A est la norme usuelle dans $\mathbb{R}^{n^2}$:

$$||A||_2 = ||\vec{V}||^2_{\mathbb{R}^{n^2}} = \sum_{i=1}^{n^2} V_i^2 = \sum_{i,j=1}^{n} A_{ij}^2. \tag{8.1}$$

Et de même que la généralisation simple de $\mathbb{R}^n$ est ℓ^2 (on ajoute la contrainte $\sum_{n=1}^\infty x_i^2 < \infty$), les endomorphismes simples de ℓ^2 seront ceux dont les matrices vérifieront $\sum_{i,j=1}^\infty A_{ij}^2 < \infty$.

8.2 Opérateur de Hilbert–Schmidt

Soit $(H, (\cdot,\cdot)_H)$ un espace de Hilbert séparable (i.e. qui admet une base dénombrable), et soit $(\vec{a}_i)_{\mathbb{N}^*}$ une $(\cdot,\cdot)_h$-b.o.n. (donc $(\vec{a}_i, \vec{a}_j)_H = \delta_{ij}$ pour tout i,j), et on dira simplement $(\vec{a}_i)$ est une b.o.n..

Définition 8.1 Soit $L \in \mathcal{L}(H)$ (un endomorphisme sur H) et soit L_{ij} est composantes sur la base $(\vec{a}_i)$, où donc $L.\vec{a}_j = \sum_{i=1}^\infty L_{ij}\vec{a}_i$ pour tout j. L'endomorphisme L est de Hilbert–Schmidt ssi :

$$\sum_{i,j=1}^\infty L_{ij}^2 < \infty. \tag{8.2}$$

On note

$$||L||_2 = \sqrt{\sum_{i,j=1}^\infty L_{ij}^2}, \quad \text{et} \quad \mathcal{L}_{HS} = \{L \text{ de Hilbert–Schmidt}\} \quad (\subset \mathcal{L}(H)). \tag{8.3}$$

Exemple 8.2 Dans $(\ell^2, (\cdot,\cdot)_{\ell^2})$ muni de sa base canonique, $L \in \mathcal{L}(\ell^2)$ donnée par (7.3), $L \in \mathcal{L}_{HS}$ ssi $\sum_{i,j=1}^\infty L_{ij}^2 < \infty$.

En particulier l'opérateur diagonale $L.\vec{e}_i = \lambda_i \vec{e}_i$ est de Hilbert–Schmidt ssi $\sum_{i=1}^\infty \lambda_i^2 < \infty$.

L'opérateur identité I n'est pas de Hilbert–Schmidt. ▪

Exercice 8.3 Montrer : 1- $\mathcal{L}_{HS}$ est un sous-espace vectoriel de $\mathcal{L}(H)$.

2- $||L||_2$ est un réel indépendant de la b.o.n. $(a_i)_{\mathbb{N}^*}$ choisie dans $L^2(\Omega)$ pour représenter L.

3- $||.||_2$ est une norme sur $\mathcal{L}_{HS}$.

4-

$$||L|| \le ||L||_2. \tag{8.4}$$

5-

$$||L \circ M||_2 \le ||L||_2 ||M||_2. \tag{8.5}$$

Réponse. Voir cours de 3ème année “Problèmes paraboliques...”. ▪

8.3 Opérateur de Hilbert–Schmidt symétrique : diagonalisable

Définition 8.4 1- A l'aide du produit scalaire $(\cdot,\cdot)_H$ fixé, pour $L \in \mathcal{L}(H)$ on définit l'application linéaire transposée $L_H^T =^{\text{not.}} L^T$ par, pour tous $\vec{x}, \vec{y} \in H$:

$$(L^T.\vec{y}, \vec{x})_H = (\vec{y}, L.\vec{x})_H \quad (= (L.\vec{x}, \vec{y})_H). \tag{8.6}$$

2- L est $(\cdot,\cdot)_H$-symétrique ssi $L^T = L$, i.e. ssi $(L.\vec{y}, \vec{x})_H = (\vec{y}, L.\vec{x})_H$ pour tout $\vec{x}, \vec{y}$.

3- L est défini positif ssi $\forall \vec{x} \neq 0$, $(L.\vec{x}, \vec{x})_H > 0$.

Avec $(\vec{a}_i)$ une b.o.n. de H, on note, pour tout j :

$$L.\vec{a}_j = \sum_i L_{ij}\vec{a}_i, \quad L^T.\vec{a}_j = \sum_i (L^T)_{ij}\vec{a}_i. \tag{8.7}$$

Et $(L^T.\vec{a}_j, \vec{a}_i)_H = (L.\vec{a}_i, \vec{a}_j)_H$ donne pour tout i, j :

$$(L^T)_{ij} = L_{ji} \quad \text{soit} \quad [L^T]_{|\vec{a}} = ([L]_{|\vec{a}})^T \quad \text{(cas d'une b.o.n.)}. \tag{8.8}$$

Proposition 8.5 *Si $L \in \mathcal{L}_{HS}$, cf. (8.2), alors $L^T \in \mathcal{L}_{HS}$, et :*

$$||L^T||_2 = ||L||_2. \tag{8.9}$$

Preuve. On a $\infty > \sum_{i=1}^\infty \sum_{j=1}^\infty L_{ij}^2 = \sum_{j=1}^\infty \sum_{i=1}^\infty L_{ij}^2$ grâce au théorème de Fubini (les séries étant à termes positifs), donc $= \sum_{i,j=1}^\infty (L^T)_{ij}^2$, donc $||L^T||_2 = ||L||_2 < \infty$. ▪

Théorème 8.6 *Un opérateur L de Hilbert–Schmidt symétrique est diagonalisable dans une b.o.n. : il existe $(\lambda_i)_{i\in\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*}$ (valeurs propres) et $(\vec{v}_i)_{i=1,\ldots,\infty} \in H^{\mathbb{N}^*}$ (vecteurs propres associés) t.q. :*

$$\forall i \in \mathbb{N}^*,\ L.\vec{v}_i = \lambda_i \vec{v}_i, \quad (\vec{v}_i)_{\mathbb{N}^*} \text{ base}, \quad \forall i, j \in \mathbb{N}^*,\ (\vec{v}_i, \vec{v}_j)_H = \delta_{ij}. \tag{8.10}$$

Et $||L||_2 = \sum_{i=1}^\infty \lambda_i^2$, donc $\sum_{i=1}^\infty \lambda_i^2 < \infty$ (et donc en particulier $\lambda_i \longrightarrow_{i\to\infty} 0$.)

Si de plus L est défini positif, alors toutes des valeurs propres λ_i sont > 0. Et dans ce cas, notant $L_{ij} =^{\text{noté}} K(i,j)$ les composantes de L dans une b.o.n., la forme bilinéaire $(\cdot,\cdot)_K$ définie par :

$$(\vec{x}, \vec{y})_K := (L.\vec{x}, \vec{y})_H = \sum_{ij} L_{ij} x_j y_i \overset{\text{not.}}{=} \sum_{ij} K(i,j) x(j) y(i), \tag{8.11}$$

est un produit scalaire sur H.

Preuve. Un opérateur de Hilbert–Schmidt est compact, et un opérateur compact symétrique est diagonalisable (voir cours de 3ème année : "Problèmes spectraux et équations d'évolution de type parabolique et hyperbolique"). D'où (8.10).

Et si L est défini positif, alors $\lambda_i = (L.\vec{v}_i, \vec{v}_i)_H > 0$. Et la bilinéarité de $(\cdot,\cdot)_K$ est immédiate par linéarité de L et bilinéarité de $(\cdot,\cdot)_{L^2(\Omega)}$. La symétrie est immédiate par symétrie de L et de $(\cdot,\cdot)_{L^2(\Omega)}$. Puis $(L.\vec{x}, \vec{x})_H > 0$ quand $\vec{x} \neq 0$ car L est défini positif. ▪

Ainsi dans la b.o.n. $(\vec{v}_i)$ de diagonalisation on a (comparer avec (5.26)) :

$$K = \sum_{i=1}^\infty \lambda_i\, \vec{v}_i \otimes_H \vec{v}_i, \quad ||L||_2 = \sqrt{\sum_{i=1}^\infty \lambda_i^2}. \tag{8.12}$$

Remarque 8.7 Pour les fonctions, (8.11) se lira $(f, g)_K = \iint_{x,y} K(x,y) f(x) g(y)\, dxdy$. ▪

Remarque 8.8 La trace d'un endomorphisme L de Hilbert Schmidt symétrique vérifie

$$\text{Tr}(L) = \sum_{i=1}^\infty (L\vec{v}_i, \vec{v}_i)_H = \sum_{i=1}^\infty \lambda_i. \tag{8.13}$$

Attention, cette trace peut valoir $+\infty$: prendre $\lambda_i = \frac{1}{i}$. ▪

9 Théorème de Mercer

$(\ell^2, (\cdot,\cdot)_{\ell^2})$ et $(L^2(\Omega), (\cdot,\cdot)_{L^2})$ sont séparables, i.e. admettent une b.o.n. dénombrable (cf. la base de Fourier dans $L^2(]0, 2\pi[)$ par exemple). Mais différence essentielle :

Les endomorphismes de $L^2(\Omega)$ ne sont pas tous représentables par des "matrices généralisés", exemple : l'identité de ℓ^2 a une matrice généralisée (la matrice identité), pas l'identité dans $L^2(\Omega)$ (n'a pas de noyau intégral voir § 9.8).

Dans la suite on se restreindra au cas simple des endomorphismes de Hilbert Schmidt dans $L^2(\Omega)$: ils ont des noyaux intégraux qui permettent, et on pourra alors faire des calculs matriciels.

Rem : on peut remplacer le produit scalaire usuel de $L^2(\Omega)$ donné par un produit scalaire de densité :

$$(f, g)_{L^2_\rho} = \int_\Omega f(\vec{x}) g(\vec{x}) dm(\vec{x}), \quad dm(\vec{x}) = \rho(\vec{x})\, dx \tag{9.1}$$

où ρ est une fonction positive intégrable (pour $(\cdot,\cdot)_{L^2}$ usuel $\rho = 1$).

9.1 Endomorphisme à noyau intégral : introduction

Soit :

$$L : \begin{cases} L^2(\Omega) & \to L^2(\Omega) \\ f & \mapsto L(f) \overset{\text{noté}}{=} Lf, \end{cases} \tag{9.2}$$

un endomorphisme (application linéaire de $L^2(\Omega)$ dans lui-même). Il est à noyau intégral ssi il est de la forme :

$$Lf(\vec{x}) = \int_{\vec{y}\in\Omega} K(\vec{x},\vec{y})\, f(y)\, dy = (K_{\vec{x}}, f)_{L^2}, \tag{9.3}$$

intégrale qui dépend du paramètre $\vec{x}$, (9.3) étant une extension de (7.6), les signes Σ et $\int$ étant les signes somme. Et la fonction K est appelée le "noyau intégral de L", où donc

$$K : \begin{cases} \overline{\Omega}^2 & \to \mathbb{R} \\ (\vec{x},\vec{y}) & \to K(\vec{x},\vec{y}) \end{cases} \quad \text{et} \quad K_{\vec{x}} : \begin{cases} \overline{\Omega} & \to \mathbb{R} \\ \vec{y} & \to K_{\vec{x}}(\vec{y}) := K(\vec{x},\vec{y}). \end{cases} \tag{9.4}$$

Cas particulier $K(\vec{x},\vec{y}) = k(\vec{y}-\vec{x})$: l'opérateur intégral L est à noyau de convolution k :

$$Lf(\vec{x}) = (k * f)(\vec{x}) = \int_{\vec{y}\in\Omega} f(\vec{x})k(\vec{y}-\vec{x})\, dy. \tag{9.5}$$

9.2 Noyau intégral L^2

On supposera $K \in L^2(\Omega^2)$, donc défini presque partout et t.q. :

$$||K||_{L^2(\Omega^2)} < \infty, \quad \text{i.e.} \quad \iint_{(\vec{x},\vec{y})\in\Omega^2} K(\vec{x},\vec{y})^2\, dxdy < \infty, \tag{9.6}$$

à comparer à (8.2).

Proposition 9.1 *Soit $K \in L^2(\Omega^2)$. Alors, pour tout $\vec{x} \in \Omega$, $K_{\vec{x}} \in L^2(\Omega)$. Et pour $f \in L^2(\Omega)$, on a $Lf \in L^2(\Omega)$, et $L \in \mathcal{L}(L^2(\Omega))$ est un endomorphisme continu, avec :*

$$||L||_{\mathcal{L}(L^2(\Omega))} \le ||K||_{L^2(\Omega^2)}. \tag{9.7}$$

À comparer avec (8.4).

Preuve. $K \in L^2(\Omega^2)$ ssi $K^2 \in L^1(\Omega^2)$; et $K^2 \ge 0$, donc on peut appliquer le théorème de Fubini : donc en particulier $K_{\vec{x}}^2 \in L^1(\Omega)$, donc $K_{\vec{x}} \in L^2(\Omega)$. D'où, si $f \in L^2(\Omega)$, avec (9.3) et Cauchy–Schwarz dans $L^2(\Omega)$ on obtient :

$$Lf(\vec{x})^2 = (K_x, f)_{L^2}^2 \le ||K_x||_{L^2}^2 ||f||_{L^2(\Omega)}^2,$$

d'où :

$$\int_\Omega Lf(\vec{x})^2\, dx \le ||f||_{L^2(\Omega)}^2 \int_{\vec{x}\in\Omega} ||K_x||_{L^2}^2 dx \le ||f||_{L^2(\Omega)}^2 \int_{\vec{x}\in\Omega} (\int_{\vec{y}\in\Omega} |K(\vec{x},\vec{y})|^2\, dy)dx,$$

et donc $||Lf||_{L^2(\Omega)}^2 \le ||K||_{L^2(\Omega^2)}^2 ||f||_{L^2(\Omega)}^2 < \infty$, d'où $Lf \in L^2(\Omega)$ et (9.7). ∎

Exercice 9.2 (Composition des noyaux). Soit $L, M \in \mathcal{L}(L^2(\Omega))$ opérateurs à noyaux $K_L, K_M \in L^2(\Omega^2)$. Montrer : $L \circ M \in \mathcal{L}(L^2(\Omega))$, et $L \circ M$ est à noyau $K_{LM} \in L^2(\Omega^2)$ donné par

$$K_{LM}(\vec{x},\vec{y}) = \int_{\vec{z}\in\Omega} K_L(\vec{x},\vec{z})K_M(\vec{z},\vec{y})\, dz \quad (= (K_L(\vec{x},.), K_M(.,\vec{y}))_{L^2(\Omega)}), \tag{9.8}$$

à comparer à (8.5). Ainsi $(L \circ M)(\vec{x}) = \int_\Omega K_{LM}(\vec{x},\vec{y})f(\vec{y})\, dy$ pour tout $\vec{x} \in \Omega$.

Réponse. Pour $f \in L^2(\Omega)$ et $\vec{x} \in \Omega$ on a :

$$L(M(f))(\vec{x}) = \int_{\vec{z}\in\Omega} K_L(\vec{x},\vec{z})M(f)(\vec{z})\,d\vec{z} = \int_{\vec{z}\in\Omega} K_L(\vec{x},\vec{z})(\int_{\vec{y}\in\Omega} K_M(\vec{z},\vec{y})f(\vec{y})\,d\vec{y})d\vec{z}. \tag{9.9}$$

Vérifions qu'on peut appliquer Fubini : posons K_{LM} comme en (9.8), et vérifions que K_{LM} est dans $L^2(\Omega^2)$. Avec Cauchy–Schwarz dans $L^2(\Omega)$, sachant $K_L, K_M \in L^2(\Omega^2)$ et donc $K_{L,\vec{x}}, K_{M,\vec{y}} \in L^2(\Omega)$, on a, pour $\vec{x},\vec{y} \in \Omega$:

$$K_{LM}(\vec{x},\vec{y})^2 \le ||K_L(\vec{x},.)||^2_{L^2(\Omega)}||K_M(.,\vec{y})||^2_{L^2(\Omega)} = \int_{\vec{z}} K_L(\vec{x},\vec{z})^2\,d\vec{z}\int_{\vec{t}} K_M(\vec{t},\vec{y})^2\,d\vec{t},$$

et les fonctions étant toutes positives on peut appliquer Fubini, donc

$$\begin{aligned}\iint_{(\vec{x},\vec{y})\in\Omega^2} K_{LM}(\vec{x},\vec{y})^2\,d\vec{x}d\vec{y} &\le \iint_{\vec{x},\vec{y}}\iint_{\vec{z},\vec{t}} K_L(\vec{x},\vec{z})^2K_M(\vec{t},\vec{y})^2\,d\vec{x}d\vec{y}d\vec{z}d\vec{t}\\ &\le \iint_{\vec{x},\vec{z}} K_L(\vec{x},\vec{z})^2\,d\vec{x}d\vec{z}\iint_{\vec{y},\vec{t}} K_M(\vec{t},\vec{y})^2\,d\vec{y}d\vec{t}\\ &= ||K_L||^2_{L^2(\Omega^2)}||K_M||^2_{L^2(\Omega^2)} < \infty.\end{aligned}$$

Donc $K_{LM} \in L^2(\Omega^2)$. Donc $\vec{y} \to K_{LM}(\vec{x},\vec{y})f(\vec{y}) \in L^1(\Omega)$ (car $f \in L^2(\Omega)$ et Cauchy–Schwarz). Donc on peut appliquer Fubini à (9.9) : $L(M(f))(\vec{x}) = \int_{\vec{y}\in\Omega}(\int_{\vec{z}\in\Omega} K_L(\vec{x},\vec{z})K_M(\vec{z},\vec{y})\,dz)\,d\vec{y} = \int_\Omega K_{LM}(\vec{x},\vec{y})f(\vec{y})\,d\vec{y}$. ∴

9.3 Opérateur de Hilbert–Schmidt dans $L^2(\Omega)$

$(L^2(\Omega),(\cdot,\cdot)_{L^2})$ est un Hilbert séparable, voir cours "Intégrale de Lebesgue". Soit $(a_i)_{\mathbb{N}^*}$ une b.o.n. de $(L^2(\Omega),(\cdot,\cdot)_{L^2})$ (voir cours de 3ème année, les a_i sont des fonctions $\in L^2(\Omega)$). Soit $L \in \mathcal{L}(L^2(\Omega))$ (un endomorphisme borné) de matrice généralisée $[L]_{|\vec{a}} = [L_{ij}]$ dans la b.o.n. (a_i) : pour tout $j \in \mathbb{N}^*$:

$$L.a_j = \sum_{i=1}^{\infty} L_{ij}a_i. \tag{9.10}$$

Et $L \in \mathcal{L}(L^2(\Omega))$ est un opérateur de Hilbert–Schmidt ssi, cf. (8.2) :

$$||L||_2 < \infty, \quad \text{i.e.} \quad \sum_{i,j=1}^{\infty} L_{ij}^2 < \infty. \tag{9.11}$$

Remarque 9.3 On travaille ici "comme dans ℓ^2" grâce à l'isomorphisme $\mathcal{I} : L^2(]a,b[) \to \ell^2$ explicitement donné par $\mathcal{I}(a_i) = \vec{e}_i$ pour tout $i \in \mathbb{N}^*$, où $(\vec{e}_i)$ est la base canonique de ℓ^2. Et c'est bien un isomorphisme entre deux Hilbert : il conserve le produit scalaire, à savoir :

$$(\varphi,\psi)_{L^2} = (\mathcal{I}(\varphi),\mathcal{I}(\psi))_{\ell^2}. \tag{9.12}$$

En effet : $(\vec{e}_i)$ et $(\vec{a}_i)$ b.o.n., donc $\varphi = \sum_n \varphi_n a_n$ et $\psi = \sum_n \psi_n a_n$ donnent $(\varphi,\psi)_{L^2} = \sum_{nm}\varphi_n\psi_m(a_n,a_m)_{L^2} = \sum_n \varphi_n\psi_n = \sum_{nm}\varphi_n\psi_m(\vec{e}_n,\vec{e}_m)_{\ell^2} = (\mathcal{I}(\varphi),\mathcal{I}(\psi))_{\ell^2}$, donc (9.12). ∴

9.4 Noyau d'un opérateur de Hilbert–Schmidt

Ici $(H,(\cdot,\cdot)_H) = (L^2(\Omega),(\cdot,\cdot)_{L^2})$. Soit $(a_i)_{i\in\mathbb{N}^*}$ est une b.o.n. de $L^2(\Omega)$ (donc $(a_i \otimes_H a_j)_{i,j\in\mathbb{N}^*}$ est une b.o.n. de $L^2(\Omega^2)$).

Théorème 9.4 *Si $K \in L^2(\Omega^2)$, alors l'endomorphisme $L \in \mathcal{L}(L^2(\Omega))$ associé par (9.3) est de Hilbert–Schmidt : $L \in \mathcal{L}_{HS}$.*

Réciproquement, si l'endomorphisme $L \in \mathcal{L}_{HS}$ (est de Hilbert–Schmidt), alors il a un noyau $K \in L^2(\Omega^2)$, i.e. il existe $K \in L^2(\Omega^2)$ t.q. (9.3).

Et :

$$\text{si} \quad K = \sum_{i,j=1}^{\infty} K_{ij}a_i \otimes_H a_j \quad \text{et} \quad L.a_j = \sum_{i=1}^{\infty} L_{ij}a_i, \quad \text{alors} \quad L_{ij} = K_{ij}. \tag{9.13}$$

Autrement dit la matrice $[K_{ij}]$ des composantes de la fonction K dans la b.o.n. $(a_i \otimes_H a_j)$ est égale à la matrice $[L_{ij}]$ de l'endomorphisme L dans la b.o.n. (a_i).

Preuve. $\Rightarrow$ Hyp. : $K \in L^2(\Omega^2)$, donc de la forme $K = \sum_{i,j=1}^\infty K_{ij} a_i \otimes_H a_j$, avec $||K||^2_{L^2(\Omega^2)} = \sum_{i,j=1}^\infty K_{ij}^2 < \infty$ (Pythagore généralisé car $(a_i \otimes_H a_j)_{i,j\in\mathbb{N}^*}$ est une b.o.n. de $L^2(\Omega\times\Omega)$). Et, pour $f \in L^2(\Omega)$ la proposition 9.1 donne : $L \in \mathcal{L}(L^2(\Omega))$ défini par (9.3) vérifie $L.f \in L^2(\Omega)$. Notons $L.a_j = \sum_{i=1}^\infty L_{ij} a_i$. Comme $(\vec{a}_i)$ est une b.o.n., $L_{ij} = (L.a_j, a_i)_{L^2(\Omega)} = \int_{\vec{x}} (\int_{\vec{y}} K(\vec{x},\vec{y}) a_j(\vec{y})\, dy) a_i(\vec{x})\, dx = (K, a_i \otimes_H a_j)_{L^2(\Omega^2)} = K_{ij}$ (Fubini car $\iint |K(\vec{x},\vec{y}) a_j(\vec{y}) a_i(\vec{x})|\, dxdy \le ||K||_{L^2(\Omega^2)} ||a_i \otimes_H a_j||_{L^2(\Omega^2)} < \infty$). Donc $\sum_{ij} L_{ij}^2 = \sum_{ij} K_{ij}^2 < \infty$. Donc L est de Hilbert–Schmidt.

$\Leftarrow$ Supposons $L \in \mathcal{L}_{HS}$. Notons $L.a_j = \sum_{i=1}^\infty L_{ij} a_i$, avec donc $\sum_{ij} L_{ij}^2 < \infty$. Soit $K = \sum_{ij} L_{ij} a_i \otimes_H a_j$. Donc $||K||^2_{L^2(\Omega^2)} = \sum_{ij} L_{ij}^2 < \infty$ (Pythagore), donc $K \in L^2(\Omega^2)$. Et, pour tout $\vec{x}$ et pour tout $k \in \mathbb{N}^*$:

$$(K_{\vec{x}}, a_k)_{L^2} = (\sum_{i,j=1}^\infty L_{ij} a_i(\vec{x}) a_j, a_k)_{L^2} = \sum_{i,j=1}^\infty L_{ij} a_i(\vec{x}) (a_j, a_k)_{L^2} = \sum_{i=1}^\infty L_{ik} a_i(\vec{x}),$$

la deuxième égalité par bilinéarité et continuité du produit scalaire, donc $(K_{\vec{x}}, a_k)_{L^2} = L.a_k(\vec{x})$. Donc pour $f \in L^2(\Omega)$, $f = \sum_k y_k a_k$, i.e. $f(\vec{x}) = \sum_k y_k a_k(\vec{x})$ p.p., $(K_{\vec{x}}, f)_{L^2} = L.f(\vec{x})$, donc K est le noyau intégral de L. ∎

Remarque 9.5 Soit $L \in \mathcal{L}(L^2(\Omega))$ un opérateur de Hilbert–Schmidt de noyau K. Soit $f \in L^2(\Omega)$. Pour $\vec{x} \in \Omega$ on a :

$$L.f(\vec{x}) = (K_x, f)_{L^2} = \int_\Omega f(\vec{y}) \big(K_{\vec{x}}(\vec{y})\, dy\big), \tag{9.14}$$

donc la valeur $Lf(\vec{x})$ est une “valeur moyenne de f en $\vec{x}$” pour la mesure de densité $K_x(y)\, dy$.

Et l'espace propre associé à la valeur propre 1 de L est :

$$\mathrm{Ker}(L-I) = \{f \in L^2(\Omega) : L.f = f\},$$

éventuellement réduit à $\{0\}$ (quand 1 n'est pas valeur propre de L). Supposons $\mathrm{Ker}(L-I) \ne \{0\}$, et soit :

$$H = \mathrm{Ker}(L-I).$$

Alors, pour $\vec{x} \in \Omega$, $\mathcal{K}_{\vec{x}} : \left\{ \begin{array}{ll} H & \to \mathbb{R} \\ f_h & \to \mathcal{K}_{\vec{x}}(f_h) = L.f_h(\vec{x}) = f_h(\vec{x}) = (K_{\vec{x}}, f_h)_{L^2} \end{array} \right\}$ est une fonctionnelle d'évaluation de f_h en $\vec{x}$ donnée par la mesure de densité $K_x(y)\, dy$. ∎

9.5 Cas Hilbert–Schmidt symétrique

Ici l'endomorphisme $L \in \mathcal{L}(L^2(\Omega))$ est $(\cdot,\cdot)_{L^2}$-symétrique :

$$\forall f, g \in L^2(\Omega), \quad (L.f, g)_{L^2} = (L.g, f)_{L^2}. \tag{9.15}$$

Le produit scalaire $(\cdot,\cdot)_{L^2}$ étant implicite dans la suite, on dira simplement que L est symétrique.

Proposition 9.6 *Si $L \in \mathcal{L}_{HS}$ est symétrique, alors son noyau K est symétrique :*

$$K(x,y) = K(y,x), \quad \forall x, y \in \Omega. \tag{9.16}$$

Preuve. Immédiat avec (9.13) et (9.15). ∎

Théorème 9.7 *Si $L \in \mathcal{L}_{HS}$ (de Hilbert–Schmidt) symétrique défini positif de noyau $K \in L^2(\Omega^2)$, voir théorème 9.4, alors la forme bilinéaire $(\cdot,\cdot)_K$ définie par :*

$$(f,g)_K \stackrel{\text{déf}}{=} (L.f, g)_{L^2(\Omega)} = (K, f \otimes_H g)_{L^2(\Omega^2)} = \iint_{\Omega^2} K(x,y) f(y) g(x)\, dxdy, \tag{9.17}$$

est un produit scalaire sur $L^2(\Omega)$. À comparer avec (8.11).

Preuve. La bilinéarité de $(\cdot,\cdot)_K$ est immédiate par linéarité de L et bilinéarité de $(\cdot,\cdot)_{L^2(\Omega)}$. La symétrie est immédiate par symétrie de L et de $(\cdot,\cdot)_{L^2(\Omega)}$. Puis $(L.f, f)_{L^2(\Omega)} > 0$ quand $f \ne 0$ car L est défini positif. ∎

9.6 Théorème de Mercer de diagonalisation

C'est le théorème de diagonalisation des "matrices généralisées symétriques", appliqué à $L^2(\Omega)$ dans le cas des endomorphismes à noyau intégral de Hilbert–Schmidt (voir remarque 9.3).

Théorème 9.8 *Soit $K \in L^2(\Omega^2)$ symétrique. L'opérateur $L \in \mathcal{L}_{HS}$ (de Hilbert-Schmidt) associé par (9.3) est diagonalisable dans une b.o.n. de $L^2(\Omega)$:*

$$\begin{Bmatrix} \exists(\lambda_i)_{i\in\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*}, \\ \exists(\zeta_i)_{i\in\mathbb{N}^*} \text{ base de } L^2(\Omega), \end{Bmatrix} \text{ t.q. } \forall i,j \in \mathbb{N}^*, \begin{Bmatrix} L\zeta_i = \lambda_i\zeta_i, \\ (\zeta_i,\zeta_j)_{L^2} = \delta_{ij}. \end{Bmatrix} \tag{9.18}$$

Et :

$$\sum_{i=1}^{\infty} \lambda_i^2 < \infty, \quad [L]_{|\zeta} = \mathrm{diag}(\lambda_i)_{i\in\mathbb{N}^*}, \quad K = \sum_{i=1}^{\infty} \lambda_i\, \zeta_i \otimes_H \zeta_i. \tag{9.19}$$

(Faire le parallèle avec les matrices symétriques, cf. (5.26) et (8.12).) De plus :
1- si Ω est borné et si K est borné sur la diagonale, i.e. $\sup_{\vec{x}\in\Omega} |K(\vec{x},\vec{x})| = C < \infty$, et
2- si les λ_i sont tous positifs,
alors $\sum_{\mathbb{N}^*} \lambda_i < \infty$, et la trace de K a un sens, cf. (8.13).

Preuve. On applique le théorème 8.6 : d'où (9.18) et (9.19).

Puis, si K est borné sur la diagonale, alors $\exists C > 0$, $\forall x \in \Omega$, $|K(\vec{x},\vec{x})| = |\sum_i \lambda_i\zeta_i^2(\vec{x})| < C$ (il suffit que ce soit vrai pour presque tout $\vec{x}$). Si de plus Ω est borné, alors $\infty > \int_\Omega C\,dx \geq \int_\Omega |K(\vec{x},\vec{x})|\,dx$. Si de plus $\lambda_i \geq 0$ pour tout i, on a $\infty > \int_\Omega \sum_i \lambda_i\zeta_i^2(\vec{x})\,dx = \sum_i \lambda_i ||\zeta_i^2||_{L^2}^2 = \sum_i \lambda_i$: la trace est finie. ∎

Remarque 9.9 La trace n'est pas toujours finie. Exemple $\Omega =]0,1[$ (borné) avec K non borné sur la diagonale : on prend la base de Fourier $e_k(x) = \cos(2k\pi x)$ si $k \geq 0$ est pair et $e_k(x) = \sin(2k\pi x)$ si $k \geq 1$ est impair, et on définit L par $L(e_0) = 0$ et $Le_k = \dfrac{1}{k}e_k$: la trace $\sum_k \frac{1}{k}$ est infinie. Ici $K(x,y) = \displaystyle\sum_{k>0} \frac{1}{k} e_k(x)e_k(y)$, et K n'est pas essentiellement borné sur la diagonale, sinon le théorème serait faux.

Vérifions-le dans ce cas. Notons M_0 l'ensemble des sous-ensemble de $]0,1[$ de mesure nulle. Dire que K n'est pas essentiellement borné sur $]0,1[$ signifie $\forall C > 0$, $\exists A \notin M_0$, $\sup_A |K(x,x)| \geq C$.

Soit $C > 0$, soit $N \in \mathbb{N}^*$ t.q. $\sum_{k=1}^N \frac{1}{k} > 2C$, et soit $\varepsilon \in]0, \frac{1}{2N\pi}[$ t.q. $\cos^2(2N\pi\varepsilon) > \frac{1}{2}$; alors pour tout $x \in]0,\varepsilon[$ on a $\sum_{k=1}^N \frac{1}{k}\cos^2(2k\pi x) > C$, donc $\sum_{k=1}^\infty \frac{1}{k}\cos^2(2k\pi x) > C$ (on a ajouté des termes ≥ 0), ce pour tout $x \in]0,\varepsilon[$, avec $]0,\varepsilon[$ de mesure non nulle. ∎

Remarque 9.10 Pour les fonctions à valeurs complexes, ne pas oublier le conjugué dans $(f,g)_{L^2(\Omega;\mathbb{C})} = \int_\Omega f(x)\overline{g(x)}\,dx$. $(9.19)_2$ se lit $K = \sum_{i=1}^\infty \lambda_i\, \zeta_i \otimes_H \overline{\zeta_i}$, soit $K(\vec{x},\vec{y}) = \sum_{i=1}^\infty \lambda_i\, \zeta_i(\vec{x})\overline{\zeta_i(\vec{y})}$. ∎

Notation : soit $\Phi : \Omega \to \mathbb{R}^{\mathbb{N}^*}$ la fonction donnée par $\Phi(\vec{x}) = \begin{pmatrix} \sqrt{\lambda_1}\zeta_1(\vec{x}) \\ \vdots \\ \sqrt{\lambda_n}\zeta_n(\vec{x}) \\ \vdots \end{pmatrix}$. On suppose $\Phi(\vec{x}) \in \ell^2$ pour (presque) tout $\vec{x}$. Alors, à $\vec{x}$ et $\vec{y}$ fixés, cf. (9.19) :

$$K(\vec{x},\vec{y}) = (\Phi(\vec{x}),\Phi(\vec{y}))_{\ell^2}, \tag{9.20}$$

i.e. $K(\vec{x},\vec{y})$ s'exprime comme un produit scalaire dans ℓ^2. En particulier on obtient (Cauchy–Schwartz) :

$$|K(\vec{x},\vec{y})| \leq (K(\vec{x},\vec{x}))^{\frac{1}{2}}(K(\vec{y},\vec{y}))^{\frac{1}{2}} = ||\Phi(\vec{x})||_{\ell^2}||\Phi(\vec{y})||_{\ell^2}. \tag{9.21}$$

Le seul problème (!) est qu'on ne sait pas a priori, pour un $\vec{x}$ fixé, si $\Phi(\vec{x}) \in \ell^2$, i.e. si $\sum_i \lambda_i|\zeta_i(\vec{x})|^2 < \infty$. Donc (9.20) est une notation formelle, sauf si, par exemple, on est dans le cadre du théorème 9.8.

9.7 Cas positif et opérateur racine carrée

Soit $K \in L^2(\Omega^2)$ symétrique et positif : $K(\vec{x},\vec{y}) = K(\vec{y},\vec{x})$ pour (presque) tout $\vec{x},\vec{y} \in \Omega$, et

$$\forall f \in L^2(\Omega), \quad \iint_{\Omega^2} K(\vec{x},\vec{y}) f(\vec{x}) f(\vec{y})\, dxdy \geq 0.$$

Autrement dit l'endomorphisme de Hilbert–Schmidt symétrique associé L, cf. (9.3), est positif :

$$\forall f \in L^2(\Omega), \quad (L.f, f)_{L^2} \geq 0. \tag{9.22}$$

Dans ce cas, les valeurs propres éventuelles, cf. (9.18), sont positives puisque $0 \leq (L\zeta_i, \zeta_i)_{L^2} = \lambda_i ||\zeta_i||_{L^2}^2$. Et, avec les ntotations du théorème 9.8 de Mercer :

$$K(\vec{x},\vec{y}) = \sum_i \lambda_i \, \zeta_i(\vec{x}) \zeta_i(\vec{y}). \tag{9.23}$$

Définition 9.11 Notons alors (cas K symétrique positif) :

$$k(\vec{x},\vec{y}) = \sum_i \sqrt{\lambda_i} \, \zeta_i(\vec{x}) \zeta_i(\vec{y}). \tag{9.24}$$

Définition 9.12 On appelle opérateur racine carrée de L l'opérateur $L^{\frac{1}{2}}$ défini par le noyau intégral $k(\cdot,\cdot)$:

$$L^{\frac{1}{2}} f(\vec{x}) = \int_{\vec{y}} k(\vec{x},\vec{y}) f(\vec{y})\, dy. \tag{9.25}$$

Proposition 9.13 *Pour tout $\vec{x},\vec{y} \in \Omega$ on a :*

$$K(\vec{x},\vec{y}) = (k(\vec{x},\cdot), k(\cdot,\vec{y}))_{L^2} \quad (= \int_{\vec{z}} k(\vec{x},\vec{z}) k(\vec{z},\vec{y})\, dz). \tag{9.26}$$

Et

$$(L.f, f)_{L^2} = (L^{\frac{1}{2}} f, L^{\frac{1}{2}} f)_{L^2}. \tag{9.27}$$

Preuve. La symétrie de k est évidente, cf. (9.24). Avec (9.24) on a :

$$\begin{aligned}
(k(\vec{x},\cdot), k(\cdot,\vec{y}))_{L^2} &= \sum_{ij} \sqrt{\lambda_i}\sqrt{\lambda_j} \zeta_i(\vec{x}) \zeta_j(\vec{y}) (\zeta_i(\cdot), \zeta_j(\cdot))_{L^2} \\
&= \sum_{ij} \sqrt{\lambda_i}\sqrt{\lambda_j} \zeta_i(\vec{x}) \zeta_j(\vec{y}) \delta_{ij} = \sum_i \lambda_i \zeta_i(\vec{x}) \zeta_i(\vec{y}) = K(\vec{x},\vec{y}).
\end{aligned}$$

Soit (9.26). Puis :

$$\begin{aligned}
(L^{\frac{1}{2}} f, L^{\frac{1}{2}} f)_{L^2} &= \int_{\vec{z}} L^{\frac{1}{2}} f(\vec{z}) L^{\frac{1}{2}}(\vec{z})\, dz = \int_{\vec{z}} (\int_{\vec{y}} k(\vec{z},\vec{y}) f(\vec{y})\, dy)(\int_{\vec{x}} k(\vec{z},\vec{x}) f(\vec{x})\, dx)\, dz \\
&= \int_{\vec{z}} \int_{\vec{x}} (\int_{\vec{z}} k(\vec{z},\vec{y}) k(\vec{z},\vec{x})\, dz) f(\vec{y}) f(\vec{x})\, dydx \\
&= \int_{\vec{z}} \int_{\vec{x}} K(\vec{x},\vec{y}) f(\vec{y}) f(\vec{x})\, dydx = (L.f, f).
\end{aligned}$$

D'où (9.27). ∎

Application à un noyau intégral reproduisant : c'est le cas $\lambda_i = 1$ pour tout i et donc :

$$K = \sqrt{k}, \quad K(\vec{x},\vec{y}) = (K(\vec{x},\cdot), K(\cdot,\vec{y}))_{L^2}, \tag{9.28}$$

propriété de reproduction. Ici L est diagonalisable avec $\lambda = 1$ l'unique valeur propre (multiple), donc $L = I$.

NB : Attention l'opérateur identité I n'est pas de Hilbert–Schmidt dans tout $L^2(\Omega)$, et le théorème de Mercer ne s'applique pas. Pour considérer $L = I$ il faudra se contenter de sous-espace de $L^2(\Omega)$, par exemple les sous-espaces de dimension finie comme P_1. Voir § suivant.

9.8 Existence du noyau pour l'identité dans $L^2(\Omega)$: résultat négatif

Question : est-ce que l'opérateur identité $I : L^2(\Omega) \to L^2(\Omega)$ est un opérateur qui a un noyau intégral $K \in L^2(\Omega^2)$ (qui serait donc un noyau reproduisant) ?

Réponse : NON. (Voir figure 3.2 pour l'intuition : le "pic" du noyau devient de plus en plus étroit quand $P_1 \to L^2(]a,b[)$, i.e. quand on "raffine le maillage", et le noyau tendrait vers la fonction nulle presque partout.)

Si oui, prenant $\Omega =]0,1[^2$, il existerait un noyau $K \in L^2(\Omega^2)$ tel que (9.3) soit satisfaite : pour $f \in L^2([0,1])$:

$$(If(x) =) \quad f(x) = \int_{y=0}^{1} K(x,y)\, f(y)\, dy. \tag{9.29}$$

Donc, pour tous $f, g \in L^2([0,1])$:

$$(f,g)_{L^2} = \int_0^1 f(x)g(x)\, dx = \int_{x=0}^{1} (\int_{y=0}^{1} K(x,y)\, f(y)\, dy)\, g(x)\, dx. \tag{9.30}$$

Soit $]a,b[\subset]0,1[$ et $]c,d[\subset]0,1[$ tels que $]a,b[\cap]c,d[= \emptyset$, faire un dessin : le pavé $]a,b[\times]c,d[$ est un rectangle dans le triangle supérieur $y > x$ de $]0,1[^2$, ou dans le triangle inférieur $y < x$ de $]0,1[^2$. Prenons $f = 1_{]c,d[}$ et $g = 1_{]a,b[}$. L'intégrale de gauche est alors nulle, d'où :

$$0 = \int_{x=0}^{1} (\int_{y=0}^{1} K(x,y)\, 1_{[c,d]}(y)\, dy)\, 1_{]a,b[}(x)\, dx = \int_{x=a}^{b} (\int_{y=c}^{d} K(x,y)\, dy)\, dx, \tag{9.31}$$

ce pour tout pavé $]a,b[\times]c,d[$ de $[0,1]^2$ tel que $]a,b[\cap]c,d[= \emptyset$, i.e. pour tout rectangle dans le triangle supérieur $y > x$ de $]0,1[^2$, ou dans le triangle inférieur $y < x$ de $]0,1[^2$.

Et comme $]0,1[^2$ est borné, comme $K \in L^2(\Omega^2)$ on a $K \in L^1(\Omega)$ grâce à Cauchy–Schwarz : $\int_\Omega |K|\, d\Omega = (1_\Omega, |K|)_{L^2(\Omega^2)} \le ||1_\Omega||_{L^2(\Omega^2)} ||K||_{L^2(\Omega^2)} < \infty$. Donc on peut appliquer Fubini, et (9.31) donne $\iint_C K(x,y)\, dxdy = 0$ sur tout carré C dans le triangle supérieur $y > x$ de $]0,1[^2$, ou dans le triangle inférieur $y < x$ de $]0,1[^2$. Voir figure 3.2.

Donc, la diagonale $D = \{(x,y) \in]0,1[^2 : x = y\}$ étant de mesure de Lebesgue nulle, et les carrés hors diagonale formant une base de voisinage de $[0,1]^2 - D$, K est nulle presque partout dans le carré $[0,1]^2$. Donc $If = 0 = f$ pour tout $f \in L^2([0,1])$, absurde. Donc l'endomorphisme identité $L = I$ n'a pas de noyau intégral.

Remarque 9.14 Ce cas $L^2(]a,b[)$ est très différent du cas ℓ^2 : dans ℓ^2, l'endomorphisme identité $I : \ell^2 \to \ell^2$ a un noyau intégral, de matrice la matrice identité $[I]$: tous les termes sont nuls sauf les termes pour $i = j$ qui valent 1 (termes sur la diagonale). (Bien que I ne soit pas de Hilbert–Schmidt dans ℓ^2.)

Le cas ℓ^2 correspond à la mesure de comptage $\mu = \sum_n \delta_n$, somme dénombrable des masses de Dirac en n. C'est le cas des fonctions $f =^{\text{noté}} x : n \in \mathbb{N}^* \to f(n) =^{\text{noté}} x_n \in \mathbb{R}$, i.e. le cas des suites $\vec{x} = (x(n))_{\mathbb{N}^*} = (x_n)_{\mathbb{N}^*}$. Et (9.29) donne (mesure de comptage), avec $L = I$ et donc $L_{ij} = K(i,j) = \delta_{ij}$:

$$x(i) = \sum_{j=1}^{\infty} K(i,j)x(j) = \sum_{j=1}^{\infty} \delta_{ij} x(j), \tag{9.32}$$

ce qui est vrai. (Bien que l'identité I ne soit pas de Hilbert–Schmidt dans ℓ^2).

Le cas $L^2(]a,b[)$ correspond à la mesure diffuse $\mu = dx$ de Lebesgue, $\mu(f) = \int_a^b f(x)\, dx$, et ce § montre qu'on ne peut pas passer brutalement de la mesure de comptage à la mesure diffuse de Lebesgue. ▪

Exercice 9.15 Dire pourquoi on ne peut pas appliquer le théorème de représentation de Riesz pour trouver un noyau intégral dans (9.29).

Réponse. Soit $L = I : L^2(]0,1[) \to L^2(]0,1[)$; à x fixé la fonction :

$$L_x : \begin{cases} L^2([0,1]) & \to \mathbb{R} \\ \quad f & \to L_x(f) \overset{\text{déf}}{=} L.f(x) = f(x) \end{cases}$$

est trivialement une forme linéaire. Pour pouvoir appliquer le théorème de Riesz à L_x (qui donne l'existence d'une fonction $K_x \in L^2([0,1])$ t.q. $L_x(f) = (K_x, f)_{L^2}$), il faut que L_x soit continue sur $L^2(]0,1[)$, il faut :

$$\exists c > 0, \quad \forall f \in L^2(]0,1[) \text{ t.q. } ||f||_{L^2} = 1, \quad |L_x(f)| \le c. \tag{9.33}$$

C'est trivialement faux : pour $x \in]0,1[$ fixé et n assez grand, prendre f_n la "fonction porte" qui vaut $\sqrt{n}$ sur $[x - \frac{1}{2n}, x + \frac{1}{2n}]$ et qui vaut 0 ailleurs : pour tout n on a $||f_n||_{L^2} = 1$ avec $|L_x(f_n)| = |f_n(x)| = \sqrt{n} \longrightarrow_{n\to\infty} \infty$. Donc L_x est non borné, et le théorème de représentation de Riesz ne s'applique pas. ▪

9.9 Noyau intégral borné et noyau intégral continu

Proposition 9.16 *Soit $K \in L^2(\Omega^2)$, et $L \in \mathcal{L}_{HS}$ (de Hilbert–Schmidt) donné par (9.3). On suppose $K \in L^\infty(\Omega)$ i.e. $||K||_\infty < \infty$, et Ω borné. On note $|\Omega|$ le volume de Ω.*

Alors $\mathrm{Im}(L) \subset L^\infty(\Omega)$ (i.e. les $L(f)$ sont bornés) :

$$|L.f(\vec{x})| \le \sqrt{|\Omega|}\,||K||_\infty ||f||_{L^2}, \quad \textit{pour presque tout } \vec{x} \in \Omega, \tag{9.34}$$

et donc $||L.f||_\infty \le \sqrt{|\Omega|}\,||K||_\infty ||f||_{L^2}$, et donc $||L||_{\mathcal{L}(L^2(\Omega),L^\infty(\Omega))} \le \sqrt{|\Omega|}\,||K||_\infty$.

C'est en particulier vrai si $K \in C^0(\overline{\Omega}^2)$, avec alors $\mathrm{Im}(L) \subset C^0(\overline{\Omega})$.

Preuve. $|L.f(\vec{x})| \le ||K||_\infty \int_\Omega |f(\vec{x})|\,dx \le ||K||_\infty (1_\Omega, |f|)_{L^2} \le ||K||_\infty \sqrt{|\Omega|}||f||_{L^2}$, soit (9.34).

Si $K \in C^0(\overline{\Omega}^2)$, le théorème de convergence dominée donne $L.f \in C^0(\overline{\Omega})$: l'intégrant $K(\vec{x},\vec{y})f(\vec{y})$ est continu en $\vec{x}$, et est dominé par $||K||_\infty f(\vec{y})$, domination indépendante du paramètre $\vec{x}$, avec $||K||_\infty f \in L^1(\Omega)$ car Ω est borné (Cauchy–Schwarz). ▪

Remarque 9.17 Quand $K \in L^2(\Omega^2)$, on a $L \in \mathcal{L}(L^2(\Omega))$ de Hilbert–Schmidt, donc compact.

Si $K \in C^0(\overline{\Omega}^2)$ alors $L : (L^2(\Omega), ||.||_{L^2(\Omega)}) \to (C^0(\overline{\Omega}), ||.||_\infty)$ est également compact. En effet, montrons que Lf est équicontinue, ce qui permet d'appliquer le théorème d'Ascoli, cf. cours "Problèmes spectraux...". Soit $f \in L^2(\Omega)$ et soit $\tilde{\vec{x}}, \vec{x} \in \overline{\Omega}$. On a $|L.f(\vec{x}) - L.f(\tilde{\vec{x}})| \le |(K_{\vec{x}} - K_{\tilde{\vec{x}}}, f)_{L^2}| \le ||K_{\vec{x}} - K_{\tilde{\vec{x}}}||_{L^2}||f||_{L^2} \le ||f||_{L^2}\sqrt{|\Omega|}\sup_{\vec{y}\in\overline{\Omega}} |K(\vec{x},\vec{y}) - K(\tilde{\vec{x}},\vec{y})| \longrightarrow_{\vec{x}\to\tilde{\vec{x}}} 0$, car K est uniformément continue sur $\overline{\Omega}^2$. En effet, munissant $\mathbb{R}^n$ et $\mathbb{R}^{2n}$ de la norme euclidienne, $\forall \varepsilon > 0$, $\exists \eta > 0$, $\forall \vec{x}, \tilde{\vec{x}}, \vec{y}, \tilde{\vec{y}} \in \overline{\Omega}$ t.q. $||(\vec{x},\vec{y}) - (\tilde{\vec{x}},\tilde{\vec{y}})||_{\mathbb{R}^{n2}} < \eta$ (donc $||\vec{x} - \tilde{\vec{x}}||_{\mathbb{R}^n} < \eta$), on a $|K(\vec{x},\vec{y}) - K(\tilde{\vec{x}},\tilde{\vec{y}})| < \varepsilon$, car K continue sur un compact donc uniformément continue, donc $|K(\vec{x},\vec{y}) - K(\tilde{\vec{x}},\vec{y})| < \varepsilon$. Donc $L.f$ est équicontinue sur $\overline{\Omega}$. ▪

Troisième partie

Noyau intégral reproduisant

10 Espace de Hilbert à noyau reproduisant

10.1 La masse de Dirac $\delta_{\vec{x}}$ et fonctionnelle $\mathcal{K}_{\vec{x}}$ d'évaluation en $\vec{x}$

Soit Ω un ouvert non vide dans $\mathbb{R}^n$, $n \ge 1$, $C^0(\overline{\Omega})$ l'ensemble des fonctions $\overline{\Omega} \to \mathbb{R}$ continues, $\vec{x} \in \overline{\Omega}$.

Définition 10.1 La masse de Dirac (ou mesure de Dirac) en $\vec{x}$ est la fonctionnelle (i.e. la fonction dont l'ensemble de définition est un ensemble de fonctions, ici $\mathcal{F}(\overline{\Omega})$) :

$$\delta_{\vec{x}} : \begin{cases} C^0(\overline{\Omega}) & \to \mathbb{R} \\ \quad f & \to \delta_{\vec{x}}(f) \overset{\text{déf}}{=} f(\vec{x}). \end{cases} \tag{10.1}$$

Il est immédiat que $\delta_{\vec{x}}$ est linéaire : $\delta_{\vec{x}}(f + \lambda g) = (f + \lambda g)(\vec{x}) = f(\vec{x}) + \lambda\, g(\vec{x}) = \delta_{\vec{x}}(f) + \lambda\, \delta_{\vec{x}}(g)$.

Soit H un sous-espace vectoriel fermé dans $(L^2(\Omega), ||.||_{L^2(\Omega)})$. Soit $\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|H}$ restreignant $\delta_{\vec{x}}$ à H :

$$\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|H} : \begin{cases} H & \to \mathbb{R} \\ f & \to \mathcal{K}_{\vec{x}}(f) = \delta_{\vec{x}}(f) = f(\vec{x}). \end{cases} \tag{10.2}$$

Définition 10.2 Si $\mathcal{K}_{\vec{x}}$ est continue, $\mathcal{K}_{\vec{x}}$ est appelée une fonctionnelle d'évaluation en $\vec{x}$.

Rappel : $\mathcal{K}_{\vec{x}}$ étant linéaire, sa continuité s'exprime :

$$\exists c_{\vec{x}} > 0, \quad \forall f \in H, \quad |\mathcal{K}_{\vec{x}}(f| \le c_{\vec{x}}||f||_{L^2}, \tag{10.3}$$

donc ici $|f(\vec{x})| \le c_{\vec{x}}||f||_{L^2}$. C'est en particulier vrai si : $\exists c > 0$, $\forall f \in H$, $||f||_\infty \le c||f||_{L^2}$.

Exemple 10.3 Cas $H = P_1$ cf. (2.3), $x \in]a,b[$. On a P_1 de dimension finie donc : P_1 est fermé et $\mathcal{K}_x$ continue (en dimension finie toutes les formes linéaires sont continues). Donc $\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|P_1}$ est une fonctionnelle d'évaluation, ce pour tout $x \in [a,b]$. Remarque : les fonctions de P_1 sont continues, donc $f(x)$ est bien défini. ▪

Exemple 10.4 $H = P_0$ cf. (4.1), $x \in]a,b[-\{x_1,...,x_n\}$: $\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|P_0}$ est une fonctionnelle d'évaluation.

N.B. : δ_{x_i} n'est pas défini car les fonctions de P_0 ne sont pas continues en x_i. Mais on n'en aura pas besoins (on n'aura pas besoin de définir les valeurs $f(x_i)$) : l'information qui nous intéresse est l'information “dans les pixels de la photo” (les bords des pixels ne nous intéresse pas). ▪

10.2 Fonctionnelle d'évaluation et son représentant par Riesz

Soit $H \subset C^0(\overline{\Omega})$ un sous-espace vectoriel fermé dans $(L^2(\Omega), ||.||_{L^2})$, et donc $(H, (\cdot,\cdot)_{L^2}) =^{\text{noté}} (H, (\cdot,\cdot)_H)$ est un Hilbert. Soit $\vec{x} \in \overline{\Omega}$ et la fonctionnelle d'évaluation $\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|H}$.

Comme la forme linéaire $\mathcal{K}_{\vec{x}}$ est linéaire et continue on peut appliquer le théorème de représentation de Riesz : $\mathcal{K}_{\vec{x}}$ peut être représentée, à l'aide du produit scalaire $(\cdot,\cdot)_H = (\cdot,\cdot)_{L^2}$, par un vecteur $K_{\vec{x}} \in H$ (ici un vecteur est une fonction dans l'espace vectoriel $H \subset L^2(\Omega)$), i.e.

$$\exists! K_{\vec{x}} \in H, \quad \forall f \in H, \quad (\delta_x(f) =) \quad \mathcal{K}_{\vec{x}}(f) = (K_{\vec{x}}, f)_H. \tag{10.4}$$

On a donc

$$f(\vec{x}) = (K_{\vec{x}}, f)_{L^2} = \int_{\vec{y}\in\Omega} K_{\vec{x}}(\vec{y}) f(\vec{y})\, d\vec{y} = \int_{\vec{y}\in\Omega} f(\vec{y})\, dm(\vec{y}), \tag{10.5}$$

i.e. la valeur ponctuelle $f(\vec{x})$ est donnée par une valeur moyenne (“dans le pixel”) à l'aide de la mesure de densité $dm(y) = K_{\vec{x}}(\vec{y})\, d\vec{y}$, où $K_{\vec{x}}$ est le $(\cdot,\cdot)_{L^2}$-représentant de Riesz de $\mathcal{K}_{\vec{x}} = \delta_{\vec{x}|H}$.

Exemple : $H = P_1$, $f_h \in P_1$, donc $K_{\vec{x}} \in P_1$ et $f_h(\vec{x}) = (K_{\vec{x}}, f_h)_{L^2} = \int_{\vec{y}\in\Omega} K_{\vec{x}}(\vec{y}) f_h(\vec{y})\, d\vec{y}$ (calcul simple car la fonction produit $K_{\vec{x}} f_h$ est quadratique).

10.3 Noyau reproduisant et EHNR

Définition 10.5 Avec (10.5), la fonction (définie presque partout)

$$K : \begin{cases} \overline{\Omega} \times \overline{\Omega} & \to \mathbb{R} \\ (\vec{x}, \vec{y}) & \to K(\vec{x}, \vec{y}) \overset{\text{déf}}{=} K_{\vec{x}}(\vec{y}), \end{cases} \tag{10.6}$$

est appelé le noyau reproductif sur H.

Et (10.5) s'écrit

$$f(\vec{x}) = (K_{\vec{x}}, f)_{L^2} = \int_{\vec{y}\in\Omega} K(\vec{x}, \vec{y}) f(\vec{y})\, d\vec{y}. \tag{10.7}$$

En particulier

$$K(\vec{y}, \vec{x}) = K_{\vec{y}}(\vec{x}) = (K_{\vec{x}}, K_{\vec{y}})_{L^2}, \tag{10.8}$$

et donc K est symétrique : $K(\vec{y}, \vec{x}) = K(\vec{x}, \vec{y})$ (un produit scalaire est symétrique).

Définition 10.6 (10.7) est appelée la propriété de reproduction de f par le noyau K.

Définition 10.7 Lorsque H n'est pas réduit à $\{0\}$, le triplet :

$$(H, (\cdot,\cdot)_H, (\mathcal{K}_{\vec{x}})_{\vec{x}\in\Omega}) \overset{\text{noté}}{=} (H, (\cdot,\cdot)_H, K) \overset{\text{noté}}{=} H \tag{10.9}$$

est un EHNR = Espace de Hilbert à Noyau Reproduisant.

Exemple 10.8 Dans $H = P_1$, le noyau a été donné en (3.9). ▪

Exemple 10.9 P_1 et P_0 comme vus aux exemples 10.8 et 10.10 sont des EHNR. ▪

Exemple 10.10 Dans $H = P_0$, le noyau a été donné en (4.5). ▪

Remarque 10.11 Cas d'une fonction f $C^1([a,b])$ par morceaux : sa série de Fourier $S(f)$ de f vérifie $S(f)(x) = f(x)$ aux x où f est continue, et $S(f)(x) = \frac{f(x-)+f(x+)}{2}$ valeur moyennée. Rappel : pour $[a,b] = [0,1]$ pour simplifier l'écriture,

$$S(f)(x) = \sum_{k\in\mathbb{Z}} c_k e^{i2\pi kx} \quad \text{où} \quad c_k = \int_0^1 f(t) e^{-i2\pi kt}\, dt, \tag{10.10}$$

et S est un endomorphisme de $L^2(]0,1[)$. Est-ce que S a un noyau intégral ? I.e. a-t-on S de la forme $S(f)(x) = \int K(x,y) f(y)\, dy$ (donc $= f(x)$ pour presque tout x) ? Si oui, que vaut K ?

Vue l'expression des c_k on a envie d'écrire, formellement après inversion des signes $\int$ et Σ :

$$S(f)(x) = \int_0^1 (\sum_{k\in\mathbb{Z}} e^{i2\pi k(x-t)}) f(t)\,dt, \quad \text{d'où} \quad K(x,t) = \sum_{k\in\mathbb{Z}} e^{i2\pi k(x-t)} = K_x(t).$$

Malheureusement $K_x(t)$ n'a pas de sens "au sens des fonctions" (série en général divergente), mais seulement "au sens des distributions" (converge alors vers δ_x sauf aux points x de discontinuité).

Mais on peut quand même s'en servir de manière approchée : pour un calcul numérique, on ne peut pas faire une somme infinie, et on est obligé d'approximer. Pour $N \in \mathbb{N}$ ("assez grand"), posons :

$$K_N(x,t) = \sum_{k=-N}^{N} e^{i2\pi k(x-t)}, \quad \text{puis} \quad L_N(f)(x) = \int_0^1 K_N(x,t) f(t)\,dt \overset{\text{noté}}{=} \mathcal{K}_x(f). \tag{10.11}$$

Ici on travail au sens des fonctions. Et ici L_N est bien un opérateur à noyau. Et notre fonctionnelle $\mathcal{K}_x$ d'évaluation en x donne approximativement $\mathcal{K}_x(f) \simeq \frac{f(x-)+f(x+)}{2}$. ▪

Exercice 10.12 Montrer : il n'existe pas de fonction $k \in L^2(\Omega)$ t.q. $\forall f \in L^2(\Omega)$, $\delta_{\vec{x}}(f) = (k,f)_{L^2}$, i.e. on ne peut pas représenter δ_x par une fonction $L^2(\Omega)$: donc δ_x n'est pas continue sur $L^2(\Omega)$ (sinon le théorème de Riesz serait faux).

Réponse. $[a,b] = [-1,1]$, $n \geq 2$, f_n la "fonction chapeau" définie par $f_n(0) = 1$, f_n affine sur $[-\frac{1}{n}, 0]$ et sur $[0, \frac{1}{n}]$, et nulle ailleurs, dessin. Soit x_0=0. On a $f_n \in C^0 \cap L^2(]-1,1[)$ (immédiat). La suite $(f_n)_{\mathbb{N}^*}$ converge dans $L^2(]-1,1[)$ vers la fonction nulle p.p., car $||f_n - 0||_{L^2} \longrightarrow_{n\to\infty} 0$ (immédiat). On a $\delta_0(f_n) = f_n(0) = 1$ pour tout n. S'il existait $k \in L^2(]-1,1[)$ t.q. $\delta_0(f_n) = (k, f_n)_{L^2}$, i.e. $1 = (k, f_n)_{L^2}$, pour tout n, alors par continuité du produit scalaire dans L^2 on aurait $\lim_{n\to\infty}(k, f_n)_{L^2} = (k, \lim_{n\to\infty} f_n)_{L^2} = (k, 0)_{L^2} = 0$, donc $1 = 0$, absurde. ▪

10.4 Propriété de positivité et symétrie du noyau reproduisant

Proposition 10.13 *On suppose $K \in L^2(\Omega^2)$.*

1- K est symétrique, i.e. $K(\vec{x},\vec{y}) = K(\vec{y},\vec{x})$ pour tous $\vec{x},\vec{y} \in \Omega$,

2- K est défini positif :

$$\forall f \in H - \{0\}, \quad ((K(\vec{x},\vec{y}), f(\vec{y}))_{dy}, f(\vec{x}))_{dx} > 0, \tag{10.12}$$

qui n'est autre que $(f,f)_{L^2} > 0$ pour tout $f \in H - \{0\}$. En particulier $K_{\vec{x}}$ n'est jamais l'application nulle. Et

$$(f,g)_K \overset{\text{déf}}{=} \iint_{\Omega^2} K(\vec{x},\vec{y}) f(\vec{y}) g(\vec{x})\,dxdy \quad (= (f,g)_{L^2} = (K, g\otimes f)_{L^2(\Omega^2)}) \tag{10.13}$$

définit un produit scalaire sur H.

Preuve. Symétrie avec (10.8). Puis, pour $f \in H$, $f(\vec{x}) = \delta_{\vec{x}}(f) = (K_{\vec{x}}, f)_{L^2} = \int_{\vec{y}} K_{\vec{x}}(\vec{y}) f(\vec{y})\,dy$, donc pour $f \neq 0$, $0 < (f,f)_{L^2} = \int_{\vec{x}} (\int_{\vec{y}} K_{\vec{x}}(\vec{y}) f(\vec{y})\,dy) f(\vec{x})\,dx = \int_{\vec{x},\vec{y}} K(\vec{x},\vec{y}) f(\vec{x}) f(\vec{y})\,dxdy$ (Fubini). ▪

Exemple 10.14 Dans l'EHNR $(P_1, (\cdot,\cdot)_{L^2(]a,b[)}, (\delta_x)_{x\in]a,b[})$ le caractère positif du noyau K symétrique et reproduisant s'exprime :

$\forall f_h \in P_1 - \{0\}, \quad \int_x f_h(x)^2\,dx = \int_x \int_y K(x,y) f_h(y) f_h(x)\,dxdy > 0.$ ▪

10.5 Propriété de convergence presque partout

Lemme 10.15 *Soit $(H, (\cdot,\cdot)_H, K)$ est un EHNR (avec $H \subset L^2(\Omega)$ et Ω borné).*

Si K est borné, i.e. $||K||_\infty < \infty$, alors la convergence dans H implique la convergence p.p. :

$$f \in H, \quad (f_m)_{\mathbb{N}^*} \in H^{\mathbb{N}^*}, \quad \text{et} \quad ||f - f_m||_H \underset{m\to\infty}{\longrightarrow} 0 \quad \Longrightarrow \quad f_m(x) \underset{m\to\infty}{\longrightarrow} f(x) \quad \text{p.p.}. \tag{10.14}$$

Preuve. $|(f_m - f)(\vec{x})| = |(f_m - f, K_{\vec{x}})_{L^2}| \leq ||f_m - f||_{L^2} ||K_{\vec{x}}||_{L^2} \leq ||f_m - f||_{L^2} \sqrt{|\Omega|}\, ||K||_\infty$, où $|\Omega|$ est le volume de Ω, car $||K_{\vec{x}}||^2_{L^2} = \int_{\vec{y}} K(\vec{x},\vec{y})^2\,dy \leq ||K||^2_\infty |\Omega|$. ▪

11 Annexe : théorème de Moore–Aronszajn

11.1 Position du problème

On se donne une application $K \in L^2(\Omega^2)$, et pour (presque) tout $\vec{x} \in \overline{\Omega}$ on pose, pour (presque) tout $\vec{y} \in \overline{\Omega}$:

$$K_{\vec{x}}(\vec{y}) \stackrel{\text{déf}}{=} K(\vec{x}, \vec{y}). \tag{11.1}$$

On veut savoir si on peut construire un espace de Hilbert H et un produit scalaire $(\cdot, \cdot)_H$ tel que dans H l'application K soit un noyau reproduisant, i.e., si $f \in H$, pour (presque) tout $\vec{x}$:

$$(\delta_{\vec{x}}(f) =) \quad f(\vec{x}) = (K_{\vec{x}}, f)_H. \tag{11.2}$$

En particulier on devra avoir $K_{\vec{y}}(\vec{x}) = (K_{\vec{x}}, K_{\vec{y}})_H = (K_{\vec{y}}, K_{\vec{x}})_H$, donc pour (presque) tout $\vec{x}, \vec{y}$:

$$K(\vec{x}, \vec{y}) = (K_{\vec{x}}, K_{\vec{y}})_H. \tag{11.3}$$

Exemple 11.1 Cas $H = P_1$, et K donné par (3.11). Ici K est défini sur tout $\overline{\Omega}$ (et est continu). Dans ce cas $(\cdot, \cdot)_H = (\cdot, \cdot)_{L^2}$.

Et si $\tilde{K} = 2K$, on veut $f(x) = (\tilde{K}_x, f)_H$, et $(\tilde{K}_x, f)_H = 2(K_x, f)_H$, avec $f(x) = (K_x, f)_{L^2}$, donc $2(K_x, f)_H = (K_x, f)_{L^2}$, donc $(\cdot, \cdot)_H = \frac{1}{2}(\cdot, \cdot)_{L^2}$. ∴

Exemple 11.2 Cas $H = P_0$, et K donné par (4.5). Ici K est défini presque partout dans $\overline{\Omega}$ (partout sauf aux x_i). Dans ce cas $(\cdot, \cdot)_H = (\cdot, \cdot)_{L^2}$. ∴

11.2 Analyse et pré-EHNR

11.2.1 Noyau positif

En particulier (11.2) implique $K_y(x) = (K_x, K_y)_H$, et un produit scalaire étant symétrique, ce n'est possible que si K est symétrique : pour tous $\vec{x}, \vec{y} \in \Omega$:

$$K(\vec{x}, \vec{y}) = K(\vec{y}, \vec{x}), \tag{11.4}$$

et si K est au moins positive :

Définition 11.3 Une fonction $K : \Omega \times \Omega \to \mathbb{R}$ (non nulle) est dite positive ssi pour tout $m \in \mathbb{N}^*$, pour toute suite finie $(\vec{x}_1, ..., \vec{x}_m) \in \Omega^m$ de points dans Ω, on a, pour tout vecteur $\vec{\xi} = (\xi_1, ..., \xi_m) \in \mathbb{R}^m$:

$$\sum_{i,j=1}^{m} \xi_i K(\vec{x}_i, \vec{x}_j) \xi_j \geq 0. \tag{11.5}$$

En d'autres termes, quel que soit le nombre m de points $\vec{x}_i$ pris dans Ω, la matrice $N = [K(\vec{x}_i, \vec{x}_j)]_{\substack{i=1,\dots,m \\ j=1,\dots,m}}$ est positive, i.e. $(N.\vec{\xi}, \vec{\xi})_{\mathbb{R}^m} = \vec{\xi}^t.N.\vec{\xi} \geq 0$ quand $\vec{\xi} \in \mathbb{R}^m$.

Exemple 11.4 Cas $H = P_1$, et K donné par (3.11). Comme on devra avoir $(K_{\vec{x}}, K_{\vec{y}})_H = K(\vec{x}, \vec{y})$, posant $M = [(K_{\vec{x}_i}, K_{\vec{x}_j})_H]$, (11.5) s'écrit $\sum_{i,j=1}^m \xi_i M_{ij} \xi_j = [\vec{\xi}]^T.M.[\vec{\xi}]$. Il est clair qu'ayant ici $\dim H = n < \infty$, la matrice matrice M ne peut être *définie* positive que si $(K_{\vec{x}_i})_{i=1,\dots,m}$ est une famille libre, ce qui est toujours faux pour $m > n$. ∴

11.2.2 Produit scalaire

(11.3) indique : H doit inclure le sous ensemble F de $L^2(\Omega)$ des fonctions engendrées par les K_x :

$$\begin{aligned} F &= \text{Vect}\{K_{\vec{x}} : \vec{x} \in \overline{\Omega}\} \\ &= \{f : \Omega \to \mathbb{R},\ \exists m \in \mathbb{N},\ \exists (\vec{x}_i)_{i=1,\dots,m} \in \Omega^m,\ \exists (\alpha_i)_{i=1,\dots,m} \in \mathbb{R}^m,\ f = \sum_{i=1}^m \alpha_i K_{\vec{x}_i}\}. \end{aligned} \tag{11.6}$$

De même (11.3) indique que : pour $f, g \in F$, $(f, g)_F$ et $||f||_F$ doivent être définis par, quand $f =$

$\sum_{i=1}^m \alpha_i K_{\vec{x}_i}$ et $g = \sum_{j=1}^p \beta_j K_{\vec{y}_j}$:

$$\begin{cases} (f,g)_F \stackrel{\text{déf}}{=} \displaystyle\sum_{i=1}^m \sum_{j=1}^p \alpha_i \beta_j K(\vec{x}_i, \vec{y}_j), \\ ||f||_F \stackrel{\text{déf}}{=} ((f,f)_F)^{\frac{1}{2}} = (\displaystyle\sum_{i,j=1}^m \alpha_i \alpha_j K(\vec{x}_i, \vec{x}_j))^{\frac{1}{2}}. \end{cases} \tag{11.7}$$

En particulier, pour (presque) tout $\vec{x} \in \overline{\Omega}$ et $f = \sum_{i=1}^m \alpha_i K_{\vec{x}_i} \in F$,

$$(f, K_{\vec{x}})_F = \sum_{i=1}^m \alpha_i K(\vec{x}_i, \vec{x}) = \sum_{i=1}^m \alpha_i K_{\vec{x}_i}(\vec{x}) = f(\vec{x}). \tag{11.8}$$

propriété de reproduction dans F, qui définit $\mathcal{K}_x = \delta_{\vec{x}|F} : F \to \mathbb{R}$ la fonctionnelle d'évaluation en x sur F. En particulier on retrouve (11.3). Donc $(11.7)_1$ s'écrit aussi :

$$(f,g)_F \stackrel{\text{déf}}{=} \sum_{i=1}^m \sum_{j=1}^p \alpha_i \beta_j (K_{\vec{x}_i}, K_{\vec{y}_j})_F. \tag{11.9}$$

Lemme 11.5 *Soit $f \in F$. Pour (presque) tout $\vec{x} \in \overline{\Omega}$ on a*

$$|f(\vec{x})| \le ||f||_F ||K_{\vec{x}}||_F. \tag{11.10}$$

Et donc $F \subset L^\infty(\Omega)$: les $f \in F$ sont bornés dans $\overline{\Omega}$: $||f||_\infty \le ||f||_F ||K_{\vec{x}}||_F$.

Preuve. (11.8) donne $|f(\vec{x})| = |(f, K_{\vec{x}})_F| \le ||f||_F ||K_{\vec{x}}||_F$. ∴

Lemme 11.6 *Lorsque K est symétrique positive, $(\cdot,\cdot)_F$ est un produit scalaire sur F de norme associée $||.||_F$.*

Preuve. Avec (11.9) la bilinéarité et la symétrie sont immédiates. Et K est positif, donc $(11.7)_2$ et (11.5) donnent $(f,f)_F \ge 0$. Et (11.10) et $||f||_F = 0$ donnent $|f(\vec{x})| = 0$ p.p., donc $f = 0$ dans $L^2(\Omega)$. ∴

Exemple 11.7 $K(x,y) = \sin(xy)$ pour $x, y \in [0,1]$. Donc $K_x(z) = \sin(xz)$ et $(K_x, K_y)_F = K(x,y)$, et $F = \text{Vect}\{K_x\} = \{\sum_{i=1}^m \alpha_i K_{x_i}\}$, i.e. $f \in F$ ssi de la forme $f(y) = \sum_{i=1}^m \alpha_i \sin(x_i y)$. ∴

11.2.3 Pré-EHNR

Définition 11.8 Lorsque $K \in L^2(\Omega^2)$ est symétrique positive, que F est définie par (11.6) et que $(\cdot,\cdot)_F$ est le produit scalaire associé, l'espace $(F, (\cdot,\cdot)_F, K)$ est un pré-EHNR.

Un pré-EHNR est donc un EHNR si de plus $(F, ||.||_F)$ est complet.

Proposition 11.9 *Si F est de dimension finie, alors $(F, (\cdot,\cdot)_F, K)$ est un EHNR.*

Preuve. $(F, ||.||_F)$ est un espace vectoriel réel normé de dimension finie, donc est complet. ∴

Exemple 11.10 Exemples 11.1 et 11.2. ∴

11.3 Cas de noyaux bornés ou continus

Lemme 11.11 *Quand $K \in L^\infty(\Omega^2)$ on a $F \subset L^\infty(\Omega)$ et :*

$$||K_{\vec{x}}||_F \le \sqrt{||K||_\infty}, \tag{11.11}$$

et, pour tout $f \in F$, pour presque tout $\vec{x} \in \overline{\Omega}$:

$$|f(\vec{x})| \le ||f||_F \sqrt{||K||_\infty}, \quad \text{et} \quad ||f||_\infty \le ||f||_F \sqrt{||K||_\infty}. \tag{11.12}$$

Et si $(f_m)_{\mathbb{N}^*}$ est une suite de Cauchy dans $(F, ||.||_F)$, alors $(f_m)_{\mathbb{N}^*}$ une suite de Cauchy dans $(L^\infty(\Omega), ||.||_\infty)$, donc convergente dans $L^\infty(\Omega)$ vers une fonction $f \in L^\infty(\Omega)$: pour (presque) tout $\vec{x} \in \overline{\Omega}$:

$$\lim_{m\to\infty} f_m(\vec{x}) = f(\vec{x}), \quad \text{où} \quad f \in L^\infty(\Omega). \tag{11.13}$$

Preuve. Supposons $K \neq 0$ (trivial si $K = 0$). On a $K \in L^\infty(\Omega^2)$, donc pour (presque) tout $\vec{x} \in \overline{\Omega}$ on a $K_{\vec{x}} \in L^\infty(\Omega)$; et $f \in F$ est une combinaison linéaire de fonctions $L^\infty(\Omega)$ (somme finie), donc $f \in L^\infty(\Omega)$.

Et, avec (11.3), $||K_{\vec{x}}||_F^2 = (K_{\vec{x}}, K_{\vec{x}})_F = |K(\vec{x}, \vec{x})| \leq ||K||_\infty$. D'où $(11.11)_1$. Donc pour $f \in F$ et (presque) tout $\vec{x} \in \overline{\Omega}$, (11.10) donne (11.12).

Puis soit $(f_m)_{\mathbb{N}^*}$ une suite de Cauchy dans $(F, ||.||_F)$. (11.12) donne :

$$||f_p - f_m||_\infty \leq ||f_p - f_m||_F \sqrt{||K||_\infty} \underset{p,m\to\infty}{\longrightarrow} 0. \tag{11.14}$$

Donc la suite $(f_m)_{\mathbb{N}^*}$ est de Cauchy dans $(L^\infty(\Omega), ||.||_\infty)$ qui est complet, donc (f_m) est convergente dans $L^\infty(\Omega)$. Notons $f = \lim_{n\to\infty} f_m$ dans $L^\infty(\Omega)$, avec donc $||f - f_m||_\infty \longrightarrow_{m\to\infty} 0$. D'où (11.13). ∎

Lemme 11.12 *Quand $K \in C^0(\overline{\Omega}^2)$ on a le lemme précédent où on remplace $L^\infty(\Omega)$ par $C^0(\overline{\Omega})$ et où on remplace "presque partout" par "partout". En particulier, pour tout $\vec{x} \in \overline{\Omega}$:*

$$\lim_{m\to\infty} f_m(\vec{x}) = f(\vec{x}), \quad et \quad f \in C^0(\overline{\Omega}). \tag{11.15}$$

Preuve. On remplace $L^\infty(\Omega)$ par $C^0(\overline{\Omega})$ dans la démonstration précédente. ∎

11.4 Théorème de Moore–Aronszajn

Le théorème de Moore–Aronszajn concerne les noyaux $K \in L^2(\Omega^2) \bigcap L^\infty(\Omega^2)$: donc noyaux L^2 et bornés.

Le paragraphe 11.2 a construit le pré-ENHR F qu'on va compléter pour avoir un EHNR.

On suppose F de dimension infinie (le cas F de dimension finie a été traité à la proposition 11.9). Construisons $\overline{F}$ le complété de $(F, (\cdot,\cdot)_F)$ de manière classique.

1. On considère l'ensemble des suites de Cauchy dans $(F, ||.||_F)$.
2. On dit que deux suites de Cauchy (f_m) et (g_m) sont équivalentes ssi la suite (h_m) définie par $h_{2m+1} = f_m$ et $h_{2m+2} = g_m$ est de Cauchy.
3. On vérifie que cela définit bien une relation d'équivalence (réflexive, symétrique, transitive).
4. On note $\overline{F}$ l'ensemble des classes d'équivalence. Donc un élément $\dot{f} \in \overline{F}$ est tel que si $(f_m)_{\mathbb{N}^*} \in \dot{f}$ alors $\dot{f}$ est l'ensemble de toutes les suites de Cauchy équivalentes à $(f_m)_{\mathbb{N}^*}$.
5. On définit les opérations induites $+$ et $\cdot$ dans $\overline{F}$ par : si $(f_m)_{\mathbb{N}^*} \in \dot{f}$ et si $(g_m) \in \dot{g}$, alors $\dot{f} + \dot{g}$ est la classe de $(f_m + g_m)_{\mathbb{N}^*}$, et pour $\lambda \in \mathbb{R}$, $\lambda\dot{f}$ est la classe de $(\lambda f_m)_{\mathbb{N}^*}$. (L'indépendance du choix des suites $(f_m)_{\mathbb{N}^*} \in \dot{f}$ et $(g_m) \in \dot{g}$ est immédiat.)
6. Avec les opérations induites, $(\overline{F}, +, .)$ est un espace vectoriel (exercice facile : cf. définition d'un espace vectoriel).
7. On pose $(\dot{f}, \dot{g})_{\overline{F}} = \lim_{m\to\infty}(f_m, g_m)_F$ quand $(f_m)_{\mathbb{N}^*} \in \dot{f}$ et $(g_m)_{\mathbb{N}^*} \in \dot{g}$. On vérifie que $(\cdot,\cdot)_{\overline{F}}$ est un produit scalaire sur $\overline{F}$ (forme bilinéaire symétrique définie positive : exercice facile). On note $||\dot{f}||_{\overline{F}} = \sqrt{(\dot{f}, \dot{f})_{\overline{F}}}$ la norme associée. (N.B. : la définition $(\dot{f}, \dot{g})_{\overline{F}} = \lim_{m\to\infty}(f_m, g_m)_F$ est indépendante du choix des suites $(f_m)_{\mathbb{N}^*} \in \dot{f}$ et $(g_m) \in \dot{g}$: si $(f'_m)_{\mathbb{N}^*} \in \dot{f}$ et $(g'_m) \in \dot{g}$ alors $|(f_m, g_m)_F - (f'_m, g'_m)_F| = |(f_m - f'_m, g_m)_F + (f'_m, g_m - g'_m)_F| \leq ||f_m - f'_m||_F ||g_m||_F + ||f'_m||_F ||g_m - g'_m||_F \longrightarrow_{m\to\infty} 0$, donc $\lim_{m\to\infty}(f_m, g_m)_F = \lim_{m\to\infty}(f'_m, g'_m)_F$.)
8. On identifie F à un sous-ensemble de $\overline{F}$ au sens : à tout $f \in F$ on associe la suite constante $(f_m = f)_{m\in\mathbb{N}^*} =^{\text{noté}} f \in \dot{f}$. Cette suite est constante donc de Cauchy, et sa classe $\dot{f}$ est notée abusivement f (le représentant choisi dans $\dot{f}$ pour représenter sa classe). Ainsi F est identifié au sous-ensemble de $\overline{F}$ des classes des suites constantes dans F.
9. F est dense dans $\overline{F}$: pour $\dot{f} \in \overline{F}$ et $\varepsilon > 0$, il s'agit de trouver $g \in F$ t.q. $||\dot{f} - \dot{g}||_{\overline{F}} < \varepsilon$, où $\dot{g}$ est la classe d'une suite constante $(g_m = g)_{m\in\mathbb{N}^*}$. Soit $(f_m)_{\mathbb{N}^*} \in \dot{f}$. Soit $N \in \mathbb{N}^*$ t.q. $\forall m, p \geq N$, $||f_m - f_p||_F < \varepsilon/2$. En particulier $\forall m \geq N$, $||f_m - f_N||_F < \varepsilon/2$. On pose $g = f_N$. Donc $\forall m \geq N$, $||f_m - g||_F < \varepsilon/2$, donc $\lim_{m\to\infty} ||f_m - g||_F \leq \varepsilon/2 < \varepsilon$. Donc $||\dot{f} - \dot{g}||_{\overline{F}} < \varepsilon$.
10. $(\overline{F}, ||.||_{\overline{F}})$ est complet : soit $(\dot{f}_m)$ une suite de Cauchy dans $\overline{F}$, i.e. t.q. $||\dot{f}_m - \dot{f}_p||_{\overline{F}} \longrightarrow_{m\to\infty} 0$. Soit $N \in \mathbb{N}^*$. Comme F est dense dans $\overline{F}$, il existe $g_N \in F$ t.q. $||\dot{f}_N - \dot{g}_N||_{\overline{F}} < \frac{1}{N}$. La suite $(g_N)_{N\in\mathbb{N}^*}$ ainsi construite dans F est de Cauchy, car $||g_N - g_M||_F = ||\dot{g}_N - \dot{g}_M||_{\overline{F}} \leq ||\dot{g}_N - \dot{f}_N||_{\overline{F}} + ||\dot{f}_N - \dot{f}_M||_{\overline{F}} + ||\dot{f}_M - \dot{g}_M||_{\overline{F}} \longrightarrow_{N,M\to\infty} 0$. Soit $\dot{g} \in \overline{F}$ la classe de la suite $(g_N)_{N\in\mathbb{N}^*}$. On a $||\dot{f}_m - \dot{g}||_{\overline{F}} \leq ||\dot{f}_m - \dot{f}_N||_{\overline{F}} + ||\dot{f}_N - \dot{g}||_{\overline{F}} \longrightarrow_{m,N\to\infty} 0$. Donc $(\dot{f}_m)_{\mathbb{N}^*}$ converge vers $\dot{g} \in \overline{F}$.

11. On a ainsi construit l'Hilbert $(\overline{F}, (\cdot,\cdot)_{\overline{F}})$.

Construisons l'Hilbert H de Moore–Aronszajn.

1. Pour $\dot{f} \in \overline{F}$ (une classe d'équivalence), et pour $(f_m)_{\mathbb{N}^*} \in \dot{f}$, avec (11.13) (reps. (11.15)) on pose $h(\vec{x}) = \lim_{m\to\infty} f_m(\vec{x})$ pour tout $\vec{x}$, où donc $h \in L^\infty(\Omega)$ (reps. $h \in C^0(\overline{\Omega})$).
2. La fonction h est indépendante du choix de $(f_m)_{\mathbb{N}^*} \in \dot{f}$: si $(f_m)_{\mathbb{N}^*}, (g_m)_{\mathbb{N}^*} \in \dot{f}$, si $h(\vec{x}) = \lim_{m\to\infty} f_m(\vec{x})$ et $k(\vec{x}) = \lim_{m\to\infty} g_m(\vec{x})$, donc $|k(x) - f(x)| \le |k(x) - g_m(x)| + |g_m(x) - f_m(x)| + |f_m(x) - f(x)| \longrightarrow_{m\to\infty} 0$, donc $k(\vec{x}) = h(\vec{x})$, vrai pour (presque) tout x.
3. On note H l'ensemble des fonctions obtenues. En particulier $H \subset L^\infty(\Omega)$ (resp. $H \subset C^0(\overline{\Omega})$), $F \subset H$, et H est un espace vectoriel, sous-espace vectoriel de $L^\infty(\Omega)$ (resp. $C^0(\overline{\Omega})$), immédiat.
4. Pour $h, k \in H$, avec $h = \lim_{m\to\infty} f_m \in H$ où $(f_m)_{\mathbb{N}^*} \in \dot{f}$, et $k = \lim_{m\to\infty} g_m \in H$ où $(g_m)_{\mathbb{N}^*} \in \dot{g}$, on définit $(h,k)_H = \lim_{m\to\infty}(f_m, g_m)_F = (\dot{f}, \dot{g})_{\overline{F}}$ (en particulier ne dépend pas des suites de Cauchy choisie dans $\dot{f}$ et $\dot{g}$). On vérifie immédiatement que $(\cdot,\cdot)_H$ est un produit scalaire (car $(\cdot,\cdot)_{\overline{F}}$ en est un), et on note $||.||_H$ la norme associée.
5. $F \subset H$, car pour $f \in F$ la suite constante $(f_m = f)_{\mathbb{N}^*} \in \dot{f}$ a pour limite elle-même.
6. F est dense dans H : pour $h \in H$, soit $(f_m)_{\mathbb{N}^*} \in F^{\mathbb{N}^*}$ t.q. $h(\vec{x}) = \lim_{m\to\infty} f_m(\vec{x})$ pour tout $\vec{x}$. On note $\dot{f}$ la classe de $(f_m)_{\mathbb{N}^*}$. On a $\lim_{m\to\infty} ||h - f_m||_H^2 = ||h||_H^2 - \lim_{m\to\infty} 2(h, f_m)_H + \lim_{m\to\infty} ||f_m||_H^2 = ||\dot{f}||_{\overline{F}}^2 - 2(h, \lim_{m\to\infty} f_m)_H + ||\dot{f}||_{\overline{F}}^2 = ||\dot{f}||_{\overline{F}}^2 - 2(h,h)_H + ||\dot{f}||_{\overline{F}}^2 = 0$.
7. $(H, ||.||_H)$ est complet car $(\overline{F}, ||.||_{\overline{F}})$ l'est : soit $(h_M)_{\mathbb{N}^*}$ une suite de Cauchy dans H. Donc $h_M = \lim_{m\to\infty} f_{M,m}$ pour $(f_{M,m})_{m\in\mathbb{N}^*}$ de Cauchy dans F, et on note $\dot{f}_M$ la classe de $(f_{M,m})_{m\in\mathbb{N}^*}$. Donc $||h_M - h_N||_H = ||\dot{f}_M - \dot{f}_N||_{\overline{F}}$, et $(\dot{f}_M)_{\mathbb{N}^*}$ est de Cauchy dans $\overline{F}$, donc convergente vers un $\dot{f} \in \overline{F}$. Soit $(f_m)_{\mathbb{N}^*} \in \dot{f}$. Soit $h = \lim_{m\to\infty} f_m$ dans $C^0(\overline{\Omega})$, cf. (11.13). On a $h \in H$ par définition de H. Et $\lim_{M\to\infty} ||h - h_M||_H = \lim_{M\to\infty} ||\dot{f} - \dot{f}_M||_{\overline{F}} = 0$.
8. Donc $(H, (\cdot,\cdot)_H)$ est un Hilbert.

Théorème 11.13 *(Moore–Aronszajn.) Soit Ω un ouvert borné de $\mathbb{R}^n$, soit $K \in L^2(\Omega^2) \bigcap L^\infty(\Omega^2)$, soit F défini par (11.6). On suppose K symétrique (cf.(11.4)) et positive sur F (cf. (11.5)).*

Alors l'espace de Hilbert H défini ci-dessus est un EHNR :

(i) pour tout $\vec{x} \in \overline{\Omega}$ on a $K_{\vec{x}} \in H$,

(ii) pour tous $\vec{x}, \vec{y} \in \overline{\Omega}$ on a $(K_{\vec{x}}, K_{\vec{y}})_H = K(\vec{x}, \vec{y})$,

(iii) pour $h \in H$ et (presque) tout $\vec{x} \in \overline{\Omega}$ on a $h(\vec{x}) = (K_{\vec{x}}, h)_H = \delta_{\vec{x}}(h)$, propriété de noyau reproduisant.

Preuve. Par construction de H, (i), (ii) sont vérifiées.

Et $h(\vec{x}) = \lim f_m(\vec{x}) = \lim(K_{\vec{x}}, f_m)_H = (K_{\vec{x}}, \lim f_m)_H = (K_{\vec{x}}, h)_H$. ∎

Références

[1] Aronszajn N. : *Theory of Reproducing Kernels*. Transactions of the American Mathematical Society, Vol. 68, No. 3, May 1950, pp. 337-404.
`http://icl.pku.edu.cn/member/yujs/papers/pdf/ReproducingKernels.pdf`

[2] Barra Vincent : *Apprentissage statistique*. Cours de l'ISIMA.

[3] Canu Stéphane : *Machines à noyaux pour l'apprentissage statistique*. Editions T.I. 2008.

[4] Guichardet A. : *Intégration, analyse hilbertienne*. Cours de l'Ecole Polytechnique, 1983.

[5] Maitland Wright J. D. : *All operators on a Hilbert space are bounded*. Bull. Amer. Math. Soc., Volume 79, Number 6 (1973), 1247-1250.

[6] Riesz F., Nagy S. : *Leçons d'analyse fonctionnelle*. 3ème édition 1955, reprint, éditions Jacques Gabay.

[7] Rudin W. : *Principles of Mathematical Analysis*. MCGraw-Hill 1976.

[8] Sejdinovic D., Gretton A. : *What is an RKHS ?*. March 11, 2012.
`http://www.gatsby.ucl.ac.uk/~gretton/coursefiles/RKHS_Notes1.pdf`

[9] Van der Oord E. : *Indice d'une application linéaire. Théorème de Riesz. Opérateurs compacts dans L^2*. Séminaire Choquet. Initiation à l'analyse, tome 7, n°1 (1967-1968), exp. n°A1, p. A1-A8.
`http://archive.numdam.org/article/SC_1967-1968__7_1_A2_0.pdf`

[10] Vert J.P., Tsuda K., Schölkopf B. : *A primer on Kernel Methods*. 2004.