Chapitre II

ESPACES DE HILBERT

II.1. Généralités

II.1.1. Définitions

Définition II.1.1. *Soit H un espace vectoriel réel, resp. complexe. On appelle* **produit scalaire** *sur H toute* forme bilinéaire symétrique, *resp.* hermitienne, *qui est* définie positive.

On notera $(x \mid y)$ le produit scalaire des vecteurs $x, y \in H$.

Cela signifie que l'application :

$$\begin{array}{rccc} (\,.\mid .\,): & H \times H & \longrightarrow & \mathbb{K} = \mathbb{R} \text{ ou } \mathbb{C} \\ & (x,y) & \longmapsto & (x \mid y) \end{array}$$

vérifie :

1) pour tout $y \in H$, l'application $x \in H \mapsto (x \mid y) \in \mathbb{K}$ est une forme linéaire ;

2) pour tous $x, y \in H$, on a :

$$\begin{cases} (y \mid x) = \overline{(x \mid y)} & \text{si l'espace est réel} \\ (y \mid x) = \overline{(x \mid y)} & \text{si l'espace est complexe;} \end{cases}$$

3) pour tout $x \in H$, on a $(x \mid x) \geqslant 0$ et $(x \mid x) = 0$ si et seulement si $x = 0$.

Remarque. Notons que dans le cas complexe, on a donc, pour $x, y \in H$ et $\lambda \in \mathbb{C}$:

$$\boxed{(x \mid \lambda y) = \bar{\lambda}\,(x \mid y)}\,.$$

Définition II.1.2. *Si l'espace vectoriel H est muni d'un produit scalaire, on dit que c'est un* **espace préhilbertien**.

Exemples.

1) a) Le produit scalaire usuel de $\mathbb{R}^n$ est défini par :

$$\boxed{(x \mid y) = x_1 y_1 + \cdots + x_n y_n}$$

pour $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$.

Le produit scalaire usuel de $\mathbb{C}^n$ est défini par :

$$\boxed{(x \mid y) = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n}$$

pour $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{C}^n$.

b) On peut définir d'autres produits scalaires sur $\mathbb{K}^n$ en se donnant des *poids*, c'est-à-dire des nombres $w_1, \ldots, w_n > 0$, et en posant :

$$\begin{cases} (x \mid y) = \displaystyle\sum_{k=1}^{n} w_k x_k y_k \,, & \text{si } \mathbb{K} = \mathbb{R}\,; \\ (x \mid y) = \displaystyle\sum_{k=1}^{n} w_k x_k \overline{y}_k \,, & \text{si } \mathbb{K} = \mathbb{C}\,. \end{cases}$$

2) Si $(S, \mathscr{T}, m)$ est un espace mesuré, on munit $H = L^2(m)$ d'un produit scalaire (que l'on qualifiera de *naturel*) en posant, pour $f, g \in L^2(m)$:

$$\boxed{(f \mid g) = \int_S fg \, dm} \quad \text{dans le cas réel,}$$

et :

$$\boxed{(f \mid g) = \int_S f\overline{g} \, dm} \quad \text{dans le cas complexe.}$$

En particulier, sur ℓ_2, on a le produit scalaire naturel défini par :

$$\boxed{(x \mid y) = \sum_{n=1}^{\infty} x_n y_n} \quad \text{dans le cas réel,}$$

et :

$$\boxed{(x \mid y) = \sum_{n=1}^{\infty} x_n \overline{y}_n} \quad \text{dans le cas complexe.}$$

pour $x = (x_n)_{n \geqslant 1}, y = (y_n)_{n \geqslant 1} \in \ell_2$.

II.1.2. Propriétés élémentaires

Notation. Puisque $(x \mid x) \geqslant 0$, on peut poser :

$$\boxed{\|x\| = \sqrt{(x \mid x)}}\,.$$

Proposition II.1.3. *Pour tous $x, y \in H$:*

a) $\boxed{\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,(x \mid y)}$ *(cas réel) ;*

b) $\boxed{\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\,(x \mid y)}$ *(cas complexe).*

Preuve. Il suffit de développer :

$$\|x+y\|^2 = (x+y \mid x+y) = (x \mid x) + (y \mid y) + (x \mid y) + (y \mid x),$$

et utiliser le fait que $(x \mid y) + (y \mid x) = (x \mid y) + \overline{(x \mid y)} = 2(x \mid y)$ dans le cas réel, et $= 2\,\mathrm{Re}\,(x \mid y)$ dans le cas complexe. □

Théorème II.1.4 (inégalité de Cauchy-Schwarz). *Pour tous $x, y \in H$:*

$$\boxed{|(x \mid y)| \leqslant \|x\|\,\|y\|}.$$

Exemple. Dans le cas où $H = L^2(m)$, elle est équivalente à l'inégalité de Cauchy-Schwarz pour les intégrales :

$$\left|\int_S fg\,dm\right| \leqslant \int_S |fg|\,dm \leqslant \left(\int_S |f|^2\,dm\right)^{1/2} \left(\int_S |g|^2\,dm\right)^{1/2}$$

Preuve. On ne la fera que dans le cas complexe ; c'est un peu plus facile dans le cas réel (on considère le signe du produit scalaire au lieu de son argument). En fait la preuve est valable même pour les ***semi-produits scalaires***, c'est-à-dire si la forme bilinéaire symétrique (resp. hermitienne) est seulement positive (c'est-à-dire que l'on ne demande pas que $(x \mid x) = 0$ entraîne $x = 0$).

Soit $\theta \in \mathbb{R}$ tel que :

$$(\mathrm{e}^{-i\theta}x \mid y) = \mathrm{e}^{-i\theta}(x \mid y) \in \mathbb{R}_+$$

(si $(x \mid y) \neq 0$, θ est l'argument du nombre complexe $(x \mid y)$). Posons $x' = \mathrm{e}^{-i\theta}x$.

Pour tout $t \in \mathbb{R}$, on a, par la Proposition II.1.3 :

$$\|x'\|^2 + 2\,\mathrm{Re}\,(x' \mid y)\,t + \|y\|^2 t^2 = \|x' + ty\|^2 \geqslant 0.$$

Si $\|y\| = 0$, on a $\|x'\|^2 + 2\,\mathrm{Re}\,(x' \mid y)\,t \geqslant 0$ pour tout $t \in \mathbb{R}$; cela n'est possible que si $\mathrm{Re}\,(x' \mid y) = 0$. Si $\|y\| \neq 0$, on a un trinôme du second degré en t, qui est toujours positif ou nul ; son discriminant doit être négatif ou nul :

$$\mathrm{Re}\,(x' \mid y) - \|x'\|^2\|y\|^2 \leqslant 0.$$

Comme :

$$(x' \mid y) = \mathrm{e}^{-i\theta}(x \mid y) = |(x \mid y)| \in \mathbb{R}_+,$$

on a :

$$\mathrm{Re}\,(x' \mid y) = (x' \mid y) = |(x \mid y)|.$$

Comme, de plus, $\|x'\| = \|x\|$, on obtient l'inégalité annoncée. □

Corollaire II.1.5. *L'expression* $\|x\| = \sqrt{(x \mid x)}$ *définit une* <u>*norme*</u> *sur* H, *appelée* **norme hilbertienne.**

Preuve. Il suffit de vérifier l'inégalité triangulaire :

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\,(x \mid y) \leqslant \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| = (\|x\| + \|y\|)^2,$$

grâce à l'inégalité de Cauchy-Schwarz. □

Corollaire II.1.6. *Pour chaque* $y \in H$, *la forme linéaire :*

$$\boxed{\begin{array}{llll} \Phi_y : & H & \longrightarrow & \mathbb{K} = \mathbb{R} \textit{ ou } \mathbb{C} \\ & x & \longmapsto & (x \mid y) \end{array}}$$

est <u>*continue*</u>. *Sa norme dans* H^* *est* $\boxed{\|\Phi_y\| = \|y\|}$.

Preuve. On peut supposer $y \neq 0$. L'inégalité de Cauchy-Schwarz dit que :

$$|\Phi_y(x)| = |(x \mid y)| \leqslant \|y\|\,\|x\|\,;$$

cela prouve que Φ_y est continue et que $\|\Phi_y\| \leqslant \|y\|$.

Comme $\Phi_y(y) = \|y\|^2$, on a $\|\Phi_y\| \geqslant \dfrac{|\Phi_y(y)|}{\|y\|} = \|y\|$. □

Remarque importante. *Cas d'égalité dans l'inégalité de Cauchy-Schwarz.* Lorsque l'on regarde la preuve de l'inégalité (dans le cas d'un *produit scalaire*), on voit que l'on a $|(x \mid y)| = \|x\|\,\|y\|$ si et seulement si $y = 0$ ou bien si $y \neq 0$ et le discriminant du trinôme du second degré en t est nul ; cela signifie que ce trinôme possède une racine (double) : il existe $t_0 \in \mathbb{R}$ tel que $\|x' + t_0 y\| = 0$; autrement dit $\mathrm{e}^{-i\theta}x + t_0 y = 0$: les vecteurs x et y sont **linéairement liés**.

Inversement, si x et y sont linéairement dépendants, il est clair que l'on a égalité.

II.1.3. Orthogonalité

Définition II.1.7. *On dit que deux vecteurs* x *et* y *d'un espace préhilbertien* H *sont* orthogonaux *si* $(x \mid y) = 0$. *On note* $\boxed{x \perp y}$.

Exemple. Dans $H = \mathbb{R}^2$, pour le produit scalaire usuel, on a $(-1,1) \perp (1,1)$.

Notons que la relation d'orthogonalité est symétrique : si $x \perp y$, alors $y \perp x$ (car $(y \mid x) = \overline{(x \mid y)}$).

D'après la Proposition II.1.3, on a, **dans le cas réel** :

$$\boxed{x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2},$$

ce que l'on peut appeler le "*Théorème de Pythagore*".

Dans le **cas complexe** :

$$x \perp y \iff \Big[\ \|x+y\|^2 = \|x\|^2 + \|y\|^2 \quad \text{et} \quad \|x+iy\|^2 = \|x\|^2 + \|y\|^2\ \Big].$$

En effet, pour tout nombre complexe a, on a $\mathrm{Im}\,(a) = \mathrm{Re}\,(-ia)$ et par conséquent $\mathrm{Im}\,(x \mid y) = \mathrm{Re}\,(x \mid iy)$.

Des parties $A, B \subseteq H$ sont dites **orthogonales** si tout $x \in A$ est orthogonal à tout $y \in B$:

$$x \perp y, \qquad \forall x \in A,\ \forall y \in B.$$

On dit aussi que l'une est orthogonale à l'autre.

Définition II.1.8. *L'***orthogonal** *d'une partie $A \subseteq H$ est l'ensemble :*

$$\boxed{A^\perp = \{y \in H\,;\ y \perp x\,,\ \forall x \in A\}}.$$

On a $B^\perp \subseteq A^\perp$ si $A \subseteq B$; donc en particulier $\left(\overline{A}\right)^\perp \subseteq A^\perp$; mais la continuité des applications $\Phi_y \colon x \mapsto (x \mid y)$ entraîne que $\boxed{\left(\overline{A}\right)^\perp = A^\perp}$.

Proposition II.1.9. *Pour toute partie A de H, $A^\perp$ est orthogonal à A ; c'est la plus grande partie orthogonale à A.*

De plus $A^\perp$ est un sous-espace vectoriel **fermé** *de H.*

Preuve. Le début est clair. Pour le reste, remarquons que :

$$\boxed{A^\perp = \bigcap_{x \in A} \ker \Phi_x}$$

et que chaque sous-espace vectoriel $\ker \Phi_x = \Phi_x^{-1}(\{0\})$ est fermé puisque Φ_x est continue. □

II.1.4. Espaces de Hilbert

Définition II.1.10. *Si un espace préhilbertien est* complet, *pour sa norme hilbertienne, on dit que c'est un* **espace de Hilbert**.

C'est donc un cas particulier d'espace de Banach.

Exemples. 1) Tout espace préhilbertien de dimension finie est un espace de Hilbert. Lorsque le corps de base est réel, on dit que c'est un *espace euclidien*, et que c'est un *espace hermitien* lorsque le corps de base est complexe.

2) Pour toute mesure positive m, $\boxed{L^2(m) \text{ est un espace de Hilbert}}$, en vertu du Théorème de Riesz-Fisher, puisque la norme $\|\,.\,\|_2$:

$$\|f\|_2 = \left(\int_S |f(t)|^2\, dm(t)\right)^{1/2}$$

est la norme hilbertienne associée au produit scalaire usuel :

$$(f \mid g) = \int_S f(t)\overline{g(t)}\, dm(t).$$

En particulier, $\boxed{\ell_2 \text{ est un espace de Hilbert}}$.

II.2. Le Théorème de projection et ses conséquences

II.2.1. Le Théorème de projection

C'est grâce à ce théorème que l'on obtient toutes les "*bonnes*" propriétés des espaces de Hilbert.

Rappelons d'abord qu'une partie C d'un espace vectoriel est dite **convexe** si le segment $[x, y]$ est contenu dans C dès lors que $x, y \in C$:

$$x, y \in C \quad \Longrightarrow \quad [x, y] \subseteq C,$$

où $[x, y] = \{tx + (1-t)y\,;\ t \in [0, 1]\}$.

Tout sous-espace vectoriel est convexe ; toute boule est convexe.

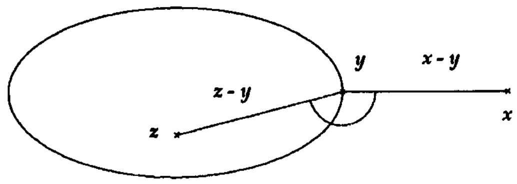
Théorème II.2.1 (Théorème de projection). *Soit H un espace de* **Hilbert** *et soit C une partie* **convexe** *et* **fermée,** *non vide, de H. Alors, pour tout $x \in H$, il existe un* *<u>unique</u>* *$y \in C$ tel que :*

$$\|x - y\| = \operatorname{dist}(x, C).$$

On dit que $y = P_C(x)$ est la projection de x sur C. *Il est caractérisé par la propriété :*

$$y \in C \quad \textit{et} \quad \operatorname{Re}(x - y \mid z - y) \leqslant 0\,, \ \forall z \in C. \qquad (*)$$

Dans le cas réel, l'inégalité dans la caractérisation $(*)$ signifie que l'angle $\alpha = \widehat{(x - y, z - y)}$ est obtus.



Notons que la complétude de H n'est pas absolument indispensable : on peut la supprimer, mais en supposant que c'est C qui est complet.

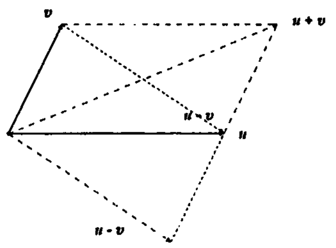
Preuve.

1) *Existence.* On aura besoin du lemme suivant, dont la preuve est immédiate, avec la Proposition II.1.3.

Lemme II.2.2 (identité du parallélogramme). *Pour tous $u, v \in H$:*

$$\boxed{\|u + v\|^2 + \|u - v\|^2 = 2\big(\|u\|^2 + \|v\|^2\big)}\,.$$

Cela signifie que la somme des carrés des diagonales d'un parallélogramme est égale à la somme des carrés des quatre côtés.



Soit $d = \operatorname{dist}(x, C) = \inf_{z \in C} \|x - z\|$.

Notons que si $d = 0$, alors $x \in C$ (car C est fermé), et $y = x$ est l'unique point de C tel que $\|x - y\| = d$.

Pour tout $n \geqslant 1$, il existe $z_n \in C$ tel que :

$$\|x - z_n\|^2 \leqslant d^2 + \frac{1}{n}\cdot$$

Appliquons alors, pour $n, p \geqslant 1$, l'identité du parallélogramme à $u = x - z_n$ et $v = x - z_p$; on obtient :

$$4\left\|x - \frac{z_n + z_p}{2}\right\|^2 + \|z_n - z_p\|^2 = 2\big(\|x - z_n\|^2 + \|x - z_p\|^2\big)\,.$$

Mais, C étant *convexe*, on a $\dfrac{z_n + z_p}{2} \in C$; donc :

$$\left\|x - \frac{z_n + z_p}{2}\right\| \geqslant d\,;$$

de sorte que l'on obtient :

$$\|z_n - z_p\|^2 \leqslant 2\left(d^2 + \frac{1}{n} + d^2 + \frac{1}{p}\right) - 4d^2 = 2\left(\frac{1}{n} + \frac{1}{p}\right)\cdot$$

La suite $(z_n)_n$ est par conséquent une suite de Cauchy. Comme H est complet, elle converge donc vers un élément $y \in H$. Mais comme C est fermé, on a en fait, puisque les z_n sont dans C, $y \in C$.

De plus, le fait que $\|x - z_n\|^2 \leqslant d^2 + 1/n$ entraîne, en passant à la limite, que $\|x - y\| \leqslant d$. On a donc $\|x - y\| = d$, puisque $y \in C$.

2) *Unicité.* Si $\|x - y_1\| = \|x - y_2\| = d$, avec $y_1, y_2 \in C$, alors, comme ci-dessus, l'identité du parallélogramme donne :

$$\begin{aligned}4d^2 + \|y_1 - y_2\|^2 &\leqslant 4\left\|x - \frac{y_1 + y_2}{2}\right\|^2 + \|y_1 - y_2\|^2\\ &= 2\left(\|x - y_1\|^2 + \|x - y_2\|^2\right) = 2\left(d^2 + d^2\right);\end{aligned}$$

d'où $\|y_1 - y_2\|^2 \leqslant 0$, ce qui n'est possible que si $y_1 = y_2$.

3) *Preuve de* $(*)$.

a) Si $z \in C$, on a $(1-t)y + tz \in C$ pour $0 \leqslant t \leqslant 1$, par la convexité de C ; donc :

$$\|x - (1-t)y - tz\|^2 \geqslant \|x - y\|^2 ,$$

soit en développant $\|x-(1-t)y-tz\|^2 = \|(x-y)+t(y-z)\|^2$ avec la Proposition II.1.3 :

$$t^2 \|y - z\|^2 + 2t \operatorname{Re}(x - y \mid y - z) \geqslant 0.$$

Pour $t \neq 0$, divisons par t, puis faisons ensuite tendre t vers 0 ; il vient $\operatorname{Re}(x - y \mid y - z) \geqslant 0$, soit :

$$\operatorname{Re}(x - y \mid z - y) \leqslant 0 .$$

b) Réciproquement, si y vérifie $(*)$, on a, pour tout $z \in C$:

$$\begin{aligned}\|x - z\|^2 &= \|(x - y + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 + 2 \operatorname{Re}(x - y \mid y - z) \\ &= \|x - y\|^2 + \|y - z\|^2 - 2 \operatorname{Re}(x - y \mid z - y) \geqslant \|x - y\|^2 ;\end{aligned}$$

donc $y = P_C(x)$, par unicité. □

II.2.2. Conséquences

Proposition II.2.3. *L'application $P_C \colon H \to C$ est continue ; plus précisément, on a, pour tous $x_1, x_2 \in H$:*

$$\|P_C(x_1) - P_C(x_2)\| \leqslant \|x_1 - x_2\| .$$

Preuve. Posons $y_1 = P_C(x_1)$ et $y_2 = P_C(x_2)$; la condition $(*)$ donne :

$$\begin{cases} \operatorname{Re}(x_1 - y_1 \mid z - y_1) \leqslant 0 & \forall z \in C; \\ \operatorname{Re}(x_2 - y_2 \mid z' - y_2) \leqslant 0 & \forall z' \in C. \end{cases}$$

En prenant $z = y_2$ et $z' = y_1$, et en additionnant, il vient :

$$\operatorname{Re}\big([x_1 - y_1] - [x_2 - y_2] \mid y_2 - y_1\big) \leqslant 0.$$

On obtient donc :

$$\begin{aligned}\|y_1 - y_2\|^2 &= \operatorname{Re}\|y_1 - y_2\|^2 = \operatorname{Re}\big([y_2 - x_2] + [x_2 - x_1] + [x_1 - y_1] \mid y_2 - y_1\big) \\ &= \operatorname{Re}\big([x_1 - y_1] - [x_2 - y_2] \mid y_2 - y_1\big) + \operatorname{Re}(x_2 - x_1 \mid y_2 - y_1) \\ &\leqslant \operatorname{Re}(x_2 - x_1 \mid y_2 - y_1) \\ &\leqslant |(x_2 - x_1 \mid y_2 - y_1)| \leqslant \|x_2 - x_1\| \, \|y_2 - y_1\|\end{aligned}$$

par l'inégalité de Cauchy-Schwarz. Il en résulte, en divisant par $\|y_2 - y_1\|$ (que l'on peut supposer non nul, car sinon le résultat est évident), que l'on a bien $\|y_1 - y_2\| \leqslant \|x_2 - x_1\|$. □

Dans le cas où le convexe C est un sous-espace vectoriel, on a de meilleures propriétés.

Théorème II.2.4. *Si F est un* **sous-espace vectoriel fermé** *de l'espace de* **Hilbert** *H, alors l'application $P_F \colon H \to F$ est une application* <u>linéaire</u> continue, *et $P_F(x)$ est l'unique point $y \in F$ tel que :*

$$\boxed{y \in F \quad \text{et} \quad x - y \in F^\perp}.$$

Preuve. D'abord, si $y \in F$ et $x - y \in F^\perp$, on a :

$$\operatorname{dist}(x, F)^2 = \inf_{z \in F} \|x - z\|^2 = \inf_{z \in F} \left[\|x - y\|^2 + \|y - z\|^2\right] = \|x - y\|^2;$$

donc $\|x - y\| = \operatorname{dist}(x, F)$ et $y = P_F(x)$.

La réciproque résulte de la condition $(*)$:

$$\operatorname{Re}(x - y \mid z - y) \leqslant 0, \quad \forall z \in F;$$

en effet, comme F est un sous-espace vectoriel, on a :

$$z = y + \lambda w \in F, \quad \forall w \in F \quad \text{et} \quad \forall \lambda \in \mathbb{K}.$$

Lorsque H est réel, on a donc, pour tout $w \in F$:

$$\lambda(x - y \mid w) = (x - y \mid \lambda w) \leqslant 0, \quad \forall \lambda \in \mathbb{R},$$

ce qui n'est possible que si $(x - y \mid w) = 0$.

Lorsque l'espace H est complexe, on a, de même, pour tout $w \in F$:

$$\lambda \operatorname{Re}(x - y \mid w) = \operatorname{Re}(x - y \mid \lambda w) \leqslant 0, \quad \forall \lambda \in \mathbb{R},$$

et, avec $z = y + i\lambda w$:

$$\lambda \operatorname{Im}(x - y \mid w) = \operatorname{Re}(x - y \mid i\lambda w) \leqslant 0, \quad \forall \lambda \in \mathbb{R},$$

ce qui, de nouveau, n'est possible que si $(x - y \mid w) = 0$.

La linéarité de P_F est alors facile à voir, grâce à l'unicité ; en effet, si $y_1 = P_F(x_1)$, $y_2 = P_F(x_2)$, alors $(x_1 - y_1), (x_2 - y_2) \in F^\perp$; donc, pour $a_1, a_2 \in \mathbb{K}$, $(a_1x_1 + a_2x_2) - (a_1y_1 + a_2y_2) \in F^\perp$; donc $P_F(a_1x_1 + a_2x_2) = a_1y_1 + a_2y_2$. □

Notons que la continuité a été vue à la Proposition II.2.3, et qu'en prenant $x_2 = 0$ dans cette proposition, on a : $\|P_F(x)\| \leqslant \|x\|$ pour tout $x \in H$; la norme de P_F est donc $\leqslant 1$. Mais comme $P_F(x) = x$ pour tout $x \in F$, on obtient, si $F \neq \{0\}$, que $\boxed{\|P_F\| = 1}$.

À titre d'exercice, on pourra montrer que, pour un convexe fermé C, P_C est linéaire si et seulement si C est un sous-espace vectoriel.

Théorème II.2.5. *Si H est un espace de* **Hilbert**, *alors, pour tout sous-espace vectoriel* **fermé**, *on a :*

$$\boxed{H = F \oplus F^{\perp}},$$

et la projection sur F parallèlement à $F^{\perp}$ associée est P_F. Elle est donc continue, de sorte que la somme directe est une somme directe topologique.

On dit que P_F est la **projection orthogonale** *sur F.*

Le fait que H soit la somme directe de F et $F^{\perp}$ signifie que tout $x \in H$ s'écrit, de façon unique, $\boxed{x = y + z}$, avec $\boxed{y \in F,\ z \in F^{\perp}}$. Notons que, puisque F et $F^{\perp}$ sont orthogonaux, on a : $\|x\|^2 = \|y\|^2 + \|z\|^2$; en d'autres termes :

$$\boxed{\|x\|^2 = \|P_F(x)\|^2 + \|x - P_F(x)\|^2}.$$

On retrouve le fait que P_F est continue et de norme 1, si $F \neq \{0\}$. On voit aussi que $\|Id_H - P_F\| = 1$, si $F^{\perp} \neq \{0\}$; mais on verra juste après qu'en fait $Id_H - P_F$ est la projection orthogonale sur $F^{\perp}$.

Preuve. On a $x = P_F(x) + \big(x - P_F(x)\big)$, avec $x - P_F(x) \in F^{\perp}$, par le Théorème II.2.4. D'autre part, si $x \in F \cap F^{\perp}$, on a, en particulier, $(x \mid x) = 0$; donc $x = 0$. □

Remarque. Le Théorème II.2.5 est vraiment spécifique aux espaces de Hilbert ; en effet, J. Lindenstraus et L. Tzafriri ont montré en 1971 que si E est un espace de Banach dans lequel tout sous-espace vectoriel fermé est l'image d'une projection continue, alors cet espace E est isomorphe à un espace de Hilbert. La preuve repose sur le Théorème de Dvoretzky, disant que tout sous-espace vectoriel de dimension finie n d'un espace normé contient un sous-espace vectoriel, de dimension "*assez grande*", de l'ordre de $\log n$, qui est très proche d'un espace de Hilbert (voir le Chapitre 8 du livre : D. Li - H. Queffélec, *Introduction à l'étude des espaces de Banach - Analyse et Probabilités*, Cours Spécialisés 12, Société Mathématique de France, 2004).

Le résultat suivant peut être montré directement, mais il est facilement obtenu à partir du Théorème II.2.5.

Corollaire II.2.6. *On a $F^{\perp\perp} = \overline{F}$ pour tout sous-espace vectoriel F de l'espace de Hilbert H.*

Preuve. Comme $F^{\perp}$ est un sous-espace vectoriel fermé, par la Proposition II.1.9, on peut lui appliquer le Théorème II.2.5 : $H = F^{\perp} \oplus F^{\perp\perp}$, que l'on peut aussi écrire : $H = F^{\perp\perp} \oplus F^{\perp}$.

D'autre part, on peut aussi appliquer ce théorème au sous-espace vectoriel fermé $\overline{F}$: $H = \overline{F} \oplus \big(\overline{F}\big)^{\perp} = \overline{F} \oplus F^{\perp}$.

Il en résulte, puisque l'on sait que $\overline{F} \subseteq F^{\perp\perp}$, que $F^{\perp\perp} = \overline{F}$. □

Notons qu'en général un sous-espace vectoriel a une infinité de supplémentaires ; mais il n'a qu'un seul supplémentaire orthogonal.

On en déduit, puisque $H^{\perp} = \{0\}$ et $\{0\}^{\perp} = H$, le critère **très pratique** suivant de densité.

Corollaire II.2.7. *Soit H un espace de Hilbert, et F un sous-espace vectoriel de H. Alors F est* **dense** *dans H si et seulement si $F^{\perp} = \{0\}$.*

Ainsi, pour montrer qu'un sous-espace vectoriel F est **dense** dans H, il suffit de vérifier que :

$$\boxed{[(x \mid y) = 0\,, \quad \forall x \in F] \quad \Longrightarrow \quad y = 0}.$$

Voyons un exemple d'application. Rappelons que le *support* de $f \colon \mathbb{R} \to \mathbb{C}$, noté $\operatorname{supp} f$, est l'adhérence de $\{x \in \mathbb{R}\,;\ f(x) \neq 0\}$.

Théorème II.2.8. *L'espace $\mathscr{K}(\mathbb{R})$ des fonctions continues sur $\mathbb{R}$ à support compact est dense dans $L^2(\mathbb{R})$.*

Ce théorème se démontre, sous une forme plus générale d'ailleurs, dans tout cours d'Intégration (voir aussi le Théorème III.1.2) ; mais il s'agit ici, même si le résultat est important par lui-même, de voir comment appliquer le Corollaire II.2.7.

Notons que $\mathscr{K}(\mathbb{R})$ n'est pas réellement contenu dans $L^2(\mathbb{R})$, puisque ce dernier est un espace de classes d'équivalence de fonctions, mais, comme deux applications continues qui sont égales presque partout, pour la mesure de Lebesgue, le sont en fait partout, l'application canonique $\jmath \colon \mathscr{K}(\mathbb{R}) \to L^2(\mathbb{R})$, qui associe à chaque fonction sa classe d'équivalence, est *injective* ; on peut donc identifier chaque $f \in \mathscr{K}(\mathbb{R})$ à sa classe d'équivalence $\jmath(f)$, c'est-à-dire $\mathscr{K}(\mathbb{R})$ à $\jmath\big[\mathscr{K}(\mathbb{R})\big]$.

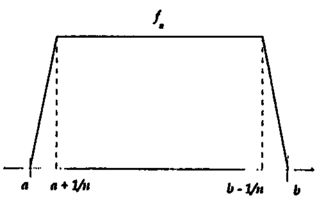
Preuve. Soit $g \in L^2(\mathbb{R})$ telle que :

$$(f \mid g) = \int_{\mathbb{R}} f\bar{g}\, d\lambda = 0\,, \quad \forall f \in \mathscr{K}(\mathbb{R}).$$

On veut montrer que $g = 0$.

En prenant les parties réelles et imaginaires, on peut supposer que g est à valeurs réelles, et l'on écrit $g = g^+ - g^-$. On a, pour toute $f \in \mathscr{K}(\mathbb{R})$:

$$\int_{\mathbb{R}} f(t)\, g^+(t)\, dt = \int_{\mathbb{R}} f(t)\, g^-(t)\, dt\,.$$



Soit $a < b$. Il existe des $f_n \in \mathscr{K}(\mathbb{R})$ telles que :

$$\left\{ \begin{array}{l} 0 \leqslant f_n \leqslant \mathbb{1}_{]a,b[}, \\ f_n(t) \xrightarrow[n \to \infty]{} \mathbb{1}_{]a,b[}(t) \text{ pour } t \in \mathbb{R}, \end{array} \right.$$

et telles que la suite $(f_n)_n$ soit croissante.

Le Théorème de convergence monotone donne :

$$\int_a^b g^+(t)\,dt = \lim_{n\to\infty}\uparrow \int_{\mathbb{R}} f_n(t)\,g^+(t)\,dt = \lim_{n\to\infty}\uparrow \int_{\mathbb{R}} f_n(t)\,g^-(t)\,dt = \int_a^b g^-(t)\,dt.$$

Cela veut dire que les mesures positives $\mu = g^+.\lambda$ et $\nu = g^-.\lambda$ sont égales sur tous les intervalles $]a,b[$ et y prennent des valeurs finies :

$$\int_a^b g^+(t)\,dt \leqslant \int_a^b |g(t)|\,dt = \int_{\mathbb{R}} |g(t)|\,\mathbb{1}_{[a,b]}(t)\,dt \leqslant \sqrt{b-a}\,\|g\|_2 < +\infty\,,$$

par l'inégalité de Cauchy-Schwarz. Le Théorème d'unicité des mesures dit alors que $\mu = \nu$. Cela signifie que $g^+ = g^-$ presque partout, c'est-à-dire $g = 0$ dans $L^2(\mathbb{R})$. □

Corollaire II.2.9. $\mathscr{C}([0,1])$ *est dense dans* $L^2(0,1)$.

Preuve. Soit $f \in L^2(0,1)$. Prolongeons-la en $\tilde{f}$ sur $\mathbb{R}$ par 0 en dehors de $[0,1]$. On a $\tilde{f} \in L^2(\mathbb{R})$. Pour tout $\varepsilon > 0$, il existe $g \in \mathscr{K}(\mathbb{R})$ telle que $\|f - g\|_{L^2(\mathbb{R})} \leqslant \varepsilon$. Soit $h = g_{|[0,1]}$ la restriction de g à $[0,1]$. On a, d'une part, $h \in \mathscr{C}([0,1])$ et, d'autre part, $\|f - h\|_{L^2(0,1)} \leqslant \|\tilde{f} - g\|_{L^2(\mathbb{R})} \leqslant \varepsilon$. □

II.2.3. Représentation du dual

Rappelons que le dual est :

$$H^* = \{\Phi\colon H \to \mathbb{K};\ \Phi \text{ linéaire continue}\}\,,$$

où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est le corps de base.

Savoir donner une représentation "*concrète*" du dual d'un espace fonctionnel permet souvent de résoudre des problèmes sur l'espace lui-même. Dans le cas des espaces de Hilbert, c'est particulièrement simple.

Rappelons d'abord que nous avons vu que, pour tout $y \in H$, la forme linéaire $\Phi_y\colon x \in H \mapsto (x \mid y)$ est continue, c'est-à-dire est un élément du dual H^*, et que $\|\Phi_y\| = \|y\|$. Il s'avère que *tous* les éléments du dual sont de cette forme.

Théorème II.2.10 (Théorème de représentation de Fréchet-Riesz).
Soit H un espace de **Hilbert**. *Pour toute $\Phi \in H^*$, il* **existe** *un (unique) $y \in H$ tel que* $\boxed{\Phi(x) = (x \mid y)}$ *pour tout $x \in H$.*

Ce théorème a été prouvé, de façon indépendante, par M. Fréchet et F. Riesz en 1907, pour $H = L^2(0,1)$; les deux articles ont été publiés, par coïncidence, dans le même numéro des Notes aux Compte-rendus de l'Académie des Sciences.

Une autre façon de voir ce théorème est de dire que l'application :

$$\begin{array}{rccl} J\colon & H & \longrightarrow & H^* \\ & y & \longmapsto & \Phi_y = J(y) \end{array}$$

est **surjective**. Elle est donc bijective car c'est une isométrie (au sens des espaces métriques) : $\|J(y) - J(y')\| = \|\Phi_y - \Phi_{y'}\| = \|\Phi_{y-y'}\| = \|y - y'\|$.

Notons que dans le cas réel, J est linéaire, mais que dans le cas complexe, elle n'est que *semi-linéaire*.

Preuve. Nous savons déjà que J est une isométrie métrique ; cela prouve l'unicité. Ce qu'il faut voir, c'est la surjectivité.

Soit $\Phi \in H^*$, non nulle. Comme Φ est continue, le sous-espace vectoriel $F = \ker \Phi$ est fermé. Donc :

$$H = (\ker \Phi) \oplus (\ker \Phi)^{\perp}.$$

Mais comme Φ est une forme linéaire non nulle, $\ker \Phi$ est de codimension 1 ; donc $(\ker \Phi)^{\perp}$ est de dimension 1.

Soit $u \in (\ker \Phi)^{\perp}$, de norme 1, et posons $y = \overline{\Phi(u)}\, u$. Alors, comme $y \in (\ker \Phi)^{\perp}$, Φ_y est nulle sur $\ker \Phi$; mais, d'autre part :

$$\Phi_y(u) = (u \mid y) = \Phi(u)\,(u \mid u) = \Phi(u)\,\|u\|^2 = \Phi(u).$$

Ainsi l'on a bien $\Phi = \Phi_y$. □

Remarque. La valeur $y = \overline{\Phi(u)}\, u$ peut sembler "tomber du ciel". En fait, si l'on veut avoir $\Phi(x) = (x \mid y)$ pour tout $x \in H$, on doit l'avoir pour $x \in \ker \Phi$; donc y doit être dans $(\ker \Phi)^{\perp}$. Ainsi $y = cu$, et l'égalité $\Phi(u) = (u \mid y)$ entraîne $\Phi(u) = \bar{c}\,(u \mid u) = \bar{c}\,\|u\|^2 = \bar{c}$. On a donc forcément $y = \overline{\Phi(u)}\, u$.

II.2.4. Adjoint d'un opérateur

On appelle *opérateur* sur H toute application linéaire continue $T\colon H \to H$.

Proposition II.2.11. *Soit H un espace de Hilbert. Pour tout $T \in \mathscr{L}(H)$, il existe un autre opérateur, noté T^*, et appelé l'**adjoint** de T, tel que :*

$$\boxed{(Tx \mid y) = (x \mid T^*y)}, \quad \forall x, y \in H.$$

De plus $\|T^*\| = \|T\|$.

Preuve. Soit $y \in H$. L'application :

→K

$$\Phi_y \circ T\colon \quad \begin{array}{ccc} H & \longrightarrow & H \\ x & \longmapsto & (Tx \mid y) \end{array}$$

est une forme linéaire continue sur H ; il existe donc, par le Théorème de Fréchet-Riesz, un unique élément de H, que l'on notera T^*y, tel que :

$$(x \mid T^*y) = (Tx \mid y), \quad \forall x \in H.$$

A cause de l'unicité, l'application $T^*\colon y \in H \mapsto T^*y \in H$ est clairement linéaire : si $y_1, y_2 \in H$ et $a_1, a_2 \in \mathbb{K}$, on a, pour tout $x \in H$:

$$\begin{aligned}(x \mid T^*(a_1y_1 + a_2y_2)) &= (Tx \mid a_1y_1 + a_2y_2) = \bar{a}_1(Tx \mid y_1) + \bar{a}_2(Tx \mid y_2) \\ &= \bar{a}_1(x \mid T^*y_1) + \bar{a}_2(x \mid T^*y_2) = (x \mid a_1T^*y_1 + a_2T^*y_2);\end{aligned}$$

donc $T^*(a_1y_1 + a_2y_2) = a_1T^*y_1 + a_2T^*y_2$.

D'autre part, l'inégalité de Cauchy-Schwarz donne :

$$|(\Phi_y \circ T)(x)| = |(Tx \mid y)| \leqslant \|Tx\| \, \|y\| \leqslant \|T\| \, \|x\| \, \|y\| \,;$$

donc $\|T^* y\| = \|\Phi_y \circ T\| \leqslant \|T\| \, \|y\|$. Cela prouve que l'application linéaire T^* est continue et que $\|T^*\| \leqslant \|T\|$.

Pour voir que $\|T\| \leqslant \|T^*\|$, remarquons que T^* a lui-même un adjoint T^{**}, et que l'on a $T^{**} = T$:

$$(y \mid T^{**} x) = (T^* y \mid x) = (y \mid Tx)$$

pour tous $x, y \in H$; cela implique que $T^{**} x = Tx$ pour tout $x \in H$. Alors $\|T\| = \|T^{**}\| \leqslant \|T^*\|$. □

II.3. Bases orthonormées

Pour éviter de parler de familles sommables, on se restreindra aux espaces **séparables**. Pour le cas général, on pourra se reporter, par exemple, au livre de G. Choquet, *Cours d'Analyse*, Masson.

II.3.1. Espaces séparables

Définition II.3.1. *Un espace topologique E est dit* séparable *s'il existe une partie $D \subseteq E$ qui est dénombrable et dense dans E : $\overline{D} = E$.*

Dans le cas des espaces normés, on a une notion équivalente.

Proposition II.3.2. *Soit E un espace vectoriel normé. Pour que E soit séparable, il faut et il suffit qu'il existe dans E une partie Δ qui soit dénombrable et* **totale** *dans E.*

On dit qu'une partie Δ d'un espace vectoriel normé E est *totale* lorsque le *sous-espace vectoriel* vect (Δ) *engendré* par cette partie est dense.

Preuve. Le $\mathbb{Q}$-sous-espace vectoriel (respectivement le $(\mathbb{Q}+i\mathbb{Q})$-sous-espace vectoriel) engendré par Δ est dénombrable et son adhérence est la même que celle de vect (Δ). □

Exemples.

1) Tout espace vectoriel de dimension finie est séparable.

2) Les espaces c_0 et ℓ_p, pour $1 \leqslant p < \infty$, sont séparables, car si

$$e_n = (0, \ldots, 0, \underset{\substack{\uparrow \\ n^{\text{ième}} \text{ place}}}{1}, 0, \ldots),$$

alors $\Delta = \{e_n \,;\ n \geqslant 1\}$ est totale, puisque, pour tout $x = (\xi_1, \xi_2, \ldots) \in \ell_p$, on a :

$$\|x - (\xi_1 e_1 + \cdots + \xi_n e_n)\|^p = \sum_{k=n+1}^{\infty} |\xi_k|^p \xrightarrow[n \to \infty]{} 0\,;$$

et lorsque $x \in c_0$:

$$\|x - (\xi_1 e_1 + \cdots + \xi_n e_n)\|_\infty = \sup_{k \geqslant n+1} |\xi_k| \xrightarrow[n \to \infty]{} 0.$$

On peut montrer (*Exercice 19 du Chapitre I*) que ℓ_∞ n'est *pas* séparable.

Proposition II.3.3. *Tout sous-espace d'un espace* métrique *séparable est séparable.*

Preuve. Soit E un espace métrique séparable, $D = \{x_n\,;\ n \geqslant 1\}$ une partie de E dénombrable dense, et $F \subseteq E$. Pourtout couple d'entiers $n, k \geqslant$ tels que $F \cap B(x_n, 1/k)$ ne soit pas vide, choisissons un élément $y_{n,k} \in F \cap B(x_n, 1/k)$; sinon (pour des questions de notation), posons $y_{n,k} = y_0$, où y_0 est un élément fixe donné de F (on peut supposer F non vide). Alors $D_F = \{y_{n,k}\,;\ n, k \geqslant 1\}$ est une partie dénombrable de F, et elle est dense dans F : soit $y \in F$; il existe, pour tout $k \geqslant 1$, un entier $n \geqslant 1$ tel que $d(y, x_n) \leqslant 1/k$; on a donc $y \in B(x_n, 1/k)$; donc $F \cap B(x_n, 1/k) \neq \emptyset$, et $y_{n,k} \in F \cap B(x_n, 1/k)$; alors $d(y, y_{n,k}) \leqslant d(y, x_n) + d(x_n, y_{n,k}) \leqslant 2/k$. □

Remarque. Ce n'est pas vrai dans les espaces topologiques généraux. En effet, pour tout ensemble I, il existe un "gros" espace compact βI, appelé *compactifié de Stone-Čech* de I dans lequel I est dense (il a la propriété que toute fonction bornée sur I à valeurs scalaires se prolonge de façon unique en une fonction continue sur βI, avec les mêmes bornes). Le compactifié de Stone-Čech $\beta\mathbb{N}$ de $\mathbb{N}$ est donc séparable; mais on peut montrer que $\beta\mathbb{N} \setminus \mathbb{N}$ n'est pas séparable.

Notons que, d'après la propriété de prolongement, l'espace $\mathscr{C}(\beta\mathbb{N})$ des fonctions continues sur $\beta\mathbb{N}$ est isométrique à ℓ_∞. Alors c_0 est isométrique au sous-espace $\{f \in \mathscr{C}(\beta\mathbb{N})\,;\ f(x) = 0 \text{ pour } x \in \beta\mathbb{N} \setminus \mathbb{N}\}$. La non séparabilité de l'espace topologique $\beta\mathbb{N} \setminus \mathbb{N}$ correspond à la non séparabilité de l'espace de Banach quotient ℓ_∞/c_0.

II.3.2. Systèmes orthonormés

Nous supposerons dans la suite que H est un espace préhilbertien, de **dimension infinie**.

Définition II.3.4. *Soit $(u_i)_{i\in I}$ une famille d'éléments de H, indexée par un ensemble arbitraire I, non vide. On dit que c'est une* **famille orthonormée**, *ou un* **système orthonormé**, *si :*

1) $\|u_i\| = 1\,,\ \forall i \in I$;
2) $u_i \perp u_j\,,\ \forall i \neq j$.

Notons que tout sous-système $(u_i)_{i\in J}$ ($J \subseteq I$) d'un système orthonormé $(u_i)_{i\in I}$ est encore orthonormé.

Exemples.

1) Dans ℓ_2, la suite $(e_n)_{n\geqslant 1}$ est orthonormée.
2) Dans $L^2(0,1)$, on pose :

$$\boxed{e_n(t) = \mathrm{e}^{2\pi i n t}}\,, \qquad n \in \mathbb{Z}\,;$$

le système $(e_n)_{n\in\mathbb{Z}}$ est orthonormé; on dit que c'est le **système trigonométrique**.

Proposition II.3.5. *Si le système fini* $(u_1, \dots, u_n)$ *est orthonormé, alors, pour tous* $a_1, \dots, a_n \in \mathbb{K}$ *:*

$$\Big\|\sum_{k=1}^{n} a_k u_k\Big\|^2 = \sum_{k=1}^{n} |a_k|^2.$$

Preuve. Il suffit de développer en utilisant la Proposition II.1.3 :

$$\Big\|\sum_{k=1}^{n} a_k u_k\Big\|^2 = \sum_{k=1}^{n} \|a_k u_k\|^2 + \sum_{k\neq j} (a_k u_k \mid a_j u_j),$$

et d'utiliser que $\|a_k u_k\| = |a_k|\,\|u_k\| = |a_k|$ et que, pour $k \neq j$, $(a_k u_k \mid a_j u_j) = a_k \overline{a}_j (u_k \mid u_j) = 0$. □

Corollaire II.3.6. *Toute famille orthonormée est* <u>*libre*</u> *(c'est-à-dire que les vecteurs la composant sont linéairement indépendants).*

Proposition II.3.7 (Inégalité de Bessel). *Soit* H *un espace préhilbertien. Pour toute famille orthonormée* $(u_i)_{i\in I}$ *dans* H*, on a, pour tout* $x \in H$ *:*

$$\boxed{\sum_{i\in I} |(x \mid u_i)|^2 \leqslant \|x\|^2}.$$

Dans l'inégalité ci-dessus, la somme au premier membre est définie de la façon suivante : si $(a_i)_{i\in I}$ est une famille de nombres réels **positifs**, alors :

$$\sum_{i\in I} a_i \overset{déf}{=} \sup_{J\subseteq I,\, J \text{ finie}} \sum_{i\in J} a_i .$$

Si $\ell_2(I) = \{(a_i)_{i\in I} \in \mathbb{K}^I \,;\, \sum_{i\in I} |a_i|^2 < +\infty\}$, l'inégalité de Bessel entraîne que l'on a une application :

$$\begin{array}{rccc} S: & H & \longrightarrow & \ell_2(I) \\ & x & \longmapsto & \big((x \mid u_i)\big)_{i\in I} \end{array} \quad ;$$

elle est linéaire, et l'inégalité de Bessel dit de plus qu'elle est continue, et de norme $\leqslant 1$.

Preuve. Si $\xi_i = (x \mid u_i)$, on a, puisque la famille est orthonormée, pour toute partie finie J de I :

$$0 \leqslant \Big\|x - \sum_{i\in J} \xi_i u_i\Big\|^2 = \|x\|^2 - 2\sum_{i\in J} \operatorname{Re}(x \mid \xi_i u_i) + \sum_{i\in J} |\xi_i|^2,$$

ce qui donne le résultat car $(x \mid \xi_i u_i) = \overline{\xi}_i (x \mid u_i) = \overline{\xi}_i \xi_i = |\xi_i|^2$. □

Proposition II.3.8. *Soit H un espace préhilbertien et soit $(u_n)_{n\geqslant 1}$ une* **suite** *orthonormée dans H. Si un vecteur $x \in H$ peut s'écrire $x = \sum_{n=1}^{\infty} \xi_n u_n$, alors on a forcément* $\boxed{\xi_n = (x \mid u_n)}$ *pour tout $n \geqslant 1$.*

Ici "*suite*" signifie "famille dénombrable".

Preuve. Pour chaque $k \geqslant 1$, la forme linéaire Φ_{u_k} est continue ; donc :

$$(x \mid u_k) = \Phi_{u_k}(x) = \sum_{n=1}^{\infty} \Phi_{u_k}(\xi_n u_n) = \sum_{n=1}^{\infty} \xi_n (u_n \mid u_k) = \xi_k . \qquad \square$$

Proposition II.3.9. *Soit $(u_n)_{n\geqslant 1}$ une* **suite** *orthonormée et $x = \sum_{n=1}^{\infty} \xi_n u_n$. Soit F_n le sous-espace vectoriel engendré par $u_1, \dots, u_n$. Alors :*

$$P_{F_n}(x) = \sum_{k=1}^{n} \xi_k u_k .$$

Preuve. Comme on a $\xi_k = (x \mid u_k)$, par la proposition précédente, on obtient que $\big(x - \sum_{k=1}^{n} \xi_k u_k \mid u_j\big) = 0$ pour tout $j \leqslant n$; donc si $y_n = \sum_{k=1}^{n} \xi_k u_k$, on a $x - y_n \in F_n^{\perp}$. Comme $y_n \in F_n$, la caractérisation du Théorème II.2.4 dit que $y_n = P_{F_n}(x)$. $\square$

Proposition II.3.10. *Si H est un espace de* **Hilbert**, *et $(u_n)_{n\geqslant 1}$ est une suite orthonormée dans H, alors, pour* **toute** *suite $(\xi_n)_{n\geqslant 1} \in \ell_2$, la série $\sum_{n=1}^{\infty} \xi_n u_n$ converge dans H.*

En d'autres termes (en utilisant la Proposition II.3.8), l'application linéaire continue :

$$\begin{array}{rccc} S\colon & H & \longrightarrow & \ell_2 \\ & x & \longmapsto & \big((x \mid u_n)\big)_{n\geqslant 1} \end{array}$$

est **surjective**.

Preuve. Il suffit de remarquer que la série vérifie le *critère de Cauchy*, car la Proposition II.3.5 donne :

$$\Big\| \sum_{k=n}^{n+p} \xi_k u_k \Big\|^2 = \sum_{k=n}^{n+p} |\xi_n|^2 \xrightarrow[n\to\infty]{} 0 ,$$

uniformément en p. $\square$

II.3.3. Bases orthonormées

Définition II.3.11. *On dit qu'une suite orthonormée $(u_n)_{n\geqslant 1}$ dans un espace préhilbertien H est une* **base orthonormée** *de H si l'ensemble $\{u_n \,;\, n \geqslant 1\}$ est <u>total</u> dans H. On dit aussi que $(u_n)_{n\geqslant 1}$ est une* **base hilbertienne**.

Notons que, comme on s'est restreint à prendre des familles dénombrables, l'espace H sera forcément séparable.

D'autre part, il faut noter que cette notion de *base orthonormée* est, en dimension infinie, différente de la notion de base, au sens algébrique du terme : une famille de vecteurs d'un espace vectoriel est une base si tout vecteur peut s'écrire, de façon unique, comme combinaison linéaire d'un nombre fini de termes de la famille ; or le théorème qui suit dit que, pour une base orthonormée, tout élément s'écrit comme la somme d'une série, qui fait intervenir *tous* les termes de la base orthonormée.

Théorème II.3.12. *Soit H un espace préhilbertien et soit $(u_n)_{n\geqslant 1}$ une base orthonormée de H. Alors, tout élément $x \in H$ s'écrit :*

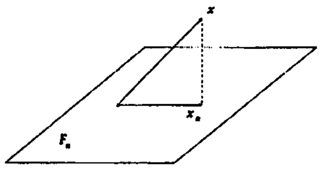
$$\boxed{x = \sum_{n=1}^{\infty} \xi_n u_n}, \qquad \textit{avec} \quad \boxed{\xi_n = (x \mid u_n)}.$$

De plus, pour tous $x, y \in H$, on a les **formules de Parseval** *:*

1) $\boxed{\|x\|^2 = \sum_{n=1}^{\infty} |(x \mid u_n)|^2}$;

2) $\boxed{(x \mid y) = \sum_{n=1}^{\infty} (x \mid u_n)\,\overline{(y \mid u_n)}}$, *la série convergeant absolument.*

Preuve. Notons F_n le sous-espace vectoriel engendré par $u_1, \ldots, u_n$, et posons $x_n = P_{F_n}(x)$.



L'ensemble $\{u_n\,;\ n \geqslant 1\}$ étant total, le sous-espace $\bigcup_{n\geqslant 1} F_n$ est dense dans H ; alors, la suite $(F_n)_{n\geqslant 1}$ étant croissante, on a :

$$\|x - x_n\| = \operatorname{dist}(x, F_n) \xrightarrow[n\to\infty]{} 0.$$

D'autre part, d'après le Corollaire II.3.6, $\{u_1, \ldots, u_n\}$ est une base, au sens usuel, de F_n ; et, par la Proposition II.3.8, on a donc :

$$x_n = \sum_{k=1}^{n} (x_n \mid u_k)\, u_k \,.$$

Mais $(x - x_n) \in F_n^{\perp}$; donc, pour $k \leqslant n$, $(x_n \mid u_k) = (x \mid u_k) = \xi_k$ ne dépend pas de n. On a donc bien :

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \xi_k u_k = \sum_{k=1}^{\infty} \xi_k u_k \,.$$

De même $y = \sum_{k=1}^{\infty} \zeta_k u_k$, avec $\zeta_k = (y \mid u_k)$. Alors, par continuité (Corollaire II.1.6) :

$$(x \mid y) = \Big(\sum_{k=1}^{\infty} \xi_k u_k \mid y \Big) = \sum_{k=1}^{\infty} \xi_k (u_k \mid y) = \sum_{k=1}^{\infty} \xi_k \overline{\zeta_k}\,,$$

qui donne l'autre identité lorsque $y = x$. □

Il résulte du Théorème II.3.12 et de la Proposition II.3.10 que l'on a :

Corollaire II.3.13. *Soit H un espace de* **Hilbert**, *séparable, et soit $(u_n)_{n\geqslant 1}$ une base orthonormée de H. Alors l'application linéaire :*

$$\begin{array}{cccl} S: & H & \longrightarrow & \ell_2 \\ & x & \longmapsto & \big((x \mid u_n)\big)_{n\geqslant 1} \end{array}$$

est un <u>**isomorphisme**</u> *d'espaces de Hilbert, c'est-à-dire un isomorphisme conservant le produit scalaire : $\big(S(\xi) \mid S(\zeta)\big) = (\xi \mid \zeta)$ pour tous $\xi, \zeta \in \ell_2$.*

C'est en particulier une isométrie : $\|S(x)\| = \|x\|$ pour tout $x \in H$. Lorsque H n'est pas complet, on a toujours une isométrie conservant le produit scalaire, mais elle n'est pas surjective.

L'isomorphisme récipoque est :

$$\begin{array}{cccl} S^{-1}: & \ell_2 & \longrightarrow & H \\ & (\xi_n)_{n\geqslant 1} & \longmapsto & \sum_{n=1}^{\infty} \xi_n u_n \end{array}$$

Nous allons voir qu'en fait *tout* espace de Hilbert séparable possède des bases orthonormées, et donc le corollaire précédent s'applique à tous les espaces de Hilbert séparables.

II.3.4. Existence des bases orthonormées

Théorème II.3.14. *Tout espace de Hilbert séparable possède des bases orthonormées.*

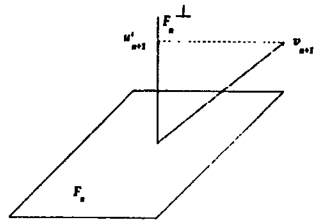
En fait la complétude ne sert pas ici (car à chaque étape, on ne travaille que dans des sous-espaces vectoriels de dimension finie, donc complets).

On obtient, comme conséquence du Théorème II.3.14 et du Corollaire II.3.13, le résultat essentiel suivant, dans lequel, cette fois-ci l'hypothèse de complétude ne peut être omise.

Théorème II.3.15. *Tous les espaces de* **Hilbert** *séparables, de dimension infinie, sont <u>isomorphes</u> entre-eux, et en particulier à ℓ_2.*

Preuve du Théorème II.3.14 . On utilise tout simplement le *procédé d'orthonormalisation de Gram-Schmidt.*

Prenons une partie dénombrable $\{v_n\,;\ n \geqslant 1\}$ totale. On peut supposer que les v_n, $n \geqslant 1$, sont linéairement indépendants (en supprimant ceux qui sont combinaison linéaire des précédents).



Soit F_n le sous-espace vectoriel engendré par $v_1, \dots, v_n$. On pose $u_1 = \dfrac{v_1}{\|v_1\|}$, et

$$u'_{n+1} = P_{F_n^\perp}(v_{n+1}), \quad u_{n+1} = \frac{u'_{n+1}}{\|u'_{n+1}\|}.$$

Alors la suite $(u_n)_{n\geqslant 1}$ est orthonormée, et l'ensemble $\{u_n\,;\ n \geqslant 1\}$ est total car le sous-espace vectoriel engendré par $u_1, \dots, u_n$ est F_n. En effet, par le Théorème II.2.4, pour $2 \leqslant k \leqslant n$, on a $u'_k - v_k \in F_{k-1}^{\perp\perp} = F_{k-1}$, et donc $u'_k \in F_k$ puisque $v_k \in F_k$ et $F_{k-1} \subseteq F_k$. □

II.4. Séparabilité de $L^2(0,1)$

II.4.1. Théorème de Stone-Weierstrass

C'est un théorème de densité dans l'espace $\mathscr{C}_\mathbb{R}(K)$ ou $\mathscr{C}_\mathbb{C}(K)$ des fonctions continues $f\colon K \to \mathbb{R}$ ou $\mathbb{C}$, où K est un espace compact. Selon que l'espace est réel ou complexe, il ne s'énonce pas de la même façon : il faut ajouter une hypothèse dans le cas complexe.

II.4.1.1. Cas réel

Théorème II.4.1 (Théorème de Stone-Weierstrass, cas réel). *Soit K un espace <u>compact</u> et A une* **sous-algèbre** *de l'algèbre de Banach* **réelle** $\mathscr{C}_\mathbb{R}(K)$.

On suppose de plus que :

a) A **sépare les points** *de K ;*

b) A **contient les constantes.**

Alors A est **dense** *dans $\mathscr{C}_\mathbb{R}(K)$.*

Remarques. 1) Une sous-algèbre de $\mathscr{C}(K)$ est un sous-espace vectoriel stable par multiplication.

2) Dire que A *sépare les points* de K signifie que si $x, y \in K$ sont distincts, alors il existe $f \in A$ telle que $f(x) \neq f(y)$.

3) L'hypothèse que A contienne les fonctions constantes n'est faite que pour éliminer le cas des sous-algèbres $A = \{f \in \mathscr{C}(K)\,;\ f(a) = 0\}$ pour un $a \in K$ donné.

Notons que, A étant un sous-espace vectoriel, A contient les constantes si et seulement si $\mathbb{1} \in A$.

On obtient la conséquence immédiate suivante.

Théorème II.4.2. *Soit K une partie compacte de $\mathbb{R}^d$; alors l'ensemble $\mathscr{P}_{\mathbb{R}}(K)$ de tous les* **polynômes réels** *à d variables, restreints à K, est* dense *dans $\mathscr{C}_{\mathbb{R}}(K)$.*

Théorème II.4.3. *L'espace réel $L^2_{\mathbb{R}}(0,1)$ est* **séparable**.

Preuve. Nous savons que $\mathscr{C}_{\mathbb{R}}([0,1])$ est dense dans $L^2_{\mathbb{R}}(0,1)$. D'autre part, le Théorème II.4.2 nous dit que $\mathscr{P}_{\mathbb{R}}([0,1])$ est dense dans $\mathscr{C}_{\mathbb{R}}([0,1])$. Donc $\mathscr{P}_{\mathbb{R}}([0,1])$ est dense dans $L^2_{\mathbb{R}}(0,1)$, *parce que la norme uniforme* sur $\mathscr{C}_{\mathbb{R}}([0,1])$ est *plus fine* que la norme de $L^2_{\mathbb{R}}(0,1)$: pour toute $f \in L^2_{\mathbb{R}}(0,1)$ et tout $\varepsilon > 0$, il existe $g \in \mathscr{C}_{\mathbb{R}}([0,1])$ telle que $\|f-g\|_2 \leqslant \varepsilon/2$; il existe ensuite $p \in \mathscr{P}_{\mathbb{R}}([0,1])$ tel que $\|g-p\|_\infty \leqslant \varepsilon/2$; mais alors $\|g-p\|_2 \leqslant \|g-p\|_\infty \leqslant \varepsilon/2$, et donc $\|f-p\|_2 \leqslant \varepsilon$.

Il ne reste plus qu'à remarquer que $\mathscr{P}_{\mathbb{R}}([0,1])$ est engendré par la suite définie par :

$$p_0(t) = 1\,, \quad p_1(t) = t\,, \quad p_2(t) = t^2\,, \qquad , \quad p_n(t) = t^n\,, \qquad ,$$

pour obtenir la séparabilité de $L^2_{\mathbb{R}}(0,1)$. □

Notons qu'au passage, nous avons prouvé la séparabilité de $\mathscr{C}_{\mathbb{R}}([0,1])$.

Corollaire II.4.4. $L^2_{\mathbb{R}}(0,1)$ *est isomorphe à l'espace réel ℓ_2.*

C'est le théorème démontré par Fisher et Riesz en 1907. Le point essentiel étant le fait que $L^2_{\mathbb{R}}(0,1)$ soit *complet.*

Preuve du Théorème de Stone-Weierstrass.
Elle se fait en plusieurs étapes.

Étape 1. Il existe une suite de polynômes *réels $(r_n)_{n\geqslant 0}$ qui converge* uniformément *sur $[0,1]$ vers la fonction racine carrée $r\colon t \mapsto \sqrt{t}$.*

Preuve. On définit $(r_n)_{n\geqslant 0}$ par récurrence, en partant de $r_0 = 0$ et en posant, pour tout $n \geqslant 0$:

$$r_{n+1}(t) = r_n(t) + \frac{1}{2}\left(t - [r_n(t)]^2\right).$$

Il est clair, par récurrence, que les r_n sont des polynômes.

De plus, pour tout $n \geqslant 0$, on a $0 \leqslant r_n(t) \leqslant \sqrt{t}$; en effet, par récurrence : on a, d'une part, $t - [r_n(t)]^2 \geqslant 0$ et donc $r_{n+1}(t) \geqslant r_n(t) \geqslant 0$, et d'autre part :

$$\sqrt{t} - r_{n+1}(t) = \left[\sqrt{t} - r_n(t)\right]\left[1 - \frac{1}{2}\left(\sqrt{t} + r_n(t)\right)\right] \geqslant 0\,,$$

car $\sqrt{t} + r_n(t) \leqslant \sqrt{t} + \sqrt{t} = 2\sqrt{t} \leqslant 2$. Notons qu'au passage, on a vu que la suite $(r_n)_{n\geqslant 0}$ est *croissante.*

Étant croissante et majorée, elle converge, vers une limite $r(t)$. La relation de récurrence montre que $r(t) = \sqrt{t}$.

Reste à voir qu'il y a convergence uniforme.

Première méthode : "à la main".

Posons $\varepsilon_n(t) = \sqrt{t} - r_n(t)$. On a vu ci-dessus, puisque $r_n(t) \geqslant 0$, que :

$$0 \leqslant \varepsilon_{n+1}(t) = \varepsilon_n(t)\Big[1 - \frac{1}{2}\big(\sqrt{t} + r_n(t)\big)\Big] \leqslant \varepsilon_n(t)\Big(1 - \frac{\sqrt{t}}{2}\Big)\,;$$

donc :

$$\begin{aligned} 0 \leqslant \varepsilon_n(t) &\leqslant \varepsilon_0(t)\Big(1 - \frac{\sqrt{t}}{2}\Big)^n = \sqrt{t}\Big(1 - \frac{\sqrt{t}}{2}\Big)^n \\ &\leqslant \sup_{0\leqslant x\leqslant 1/2} 2(1-x)x^n \qquad (\text{poser } x = 1 - \sqrt{t}/2) \\ &= 2x_n(1-x_n)x_n^n \qquad \text{avec } x_n = n/(n+1); \\ &= \frac{2}{n+1}x_n^n \leqslant \frac{2}{n+1}\,. \end{aligned}$$

□

Deuxième méthode. Il suffit d'utiliser le théorème suivant.

Théorème II.4.5 (Théorème de Dini). *Soit K un espace* **compact**.

Si $(u_n)_{n\geqslant 1}$ est une suite **croissante** *de fonctions* **continues** *$u_n\colon K \to \mathbb{R}$ qui converge simplement vers une fonction* **continue** *$u\colon K \to \mathbb{R}$, la convergence est* **uniforme**.

C'est bien sûr évidemment faux si l'on ne suppose pas la limite continue.

Preuve. Soit $\varepsilon > 0$.

Pour chaque $x \in K$, il existe un entier $N(x)$ tel que :

$$n \geqslant N(x) \quad\Longrightarrow\quad 0 \leqslant u(x) - u_n(x) \leqslant \varepsilon/3.$$

Comme u et $u_{N(x)}$ sont *continues*, il existe un voisinage de x, que l'on peut prendre ouvert, tel que :

$$x' \in V(x) \quad\Longrightarrow\quad \begin{cases} |u(x) - u(x')| \leqslant \varepsilon/3\,; \\ |u_{N(x)}(x') - u_{N(x)}(x)| \leqslant \varepsilon/3. \end{cases}$$

Comme K est compact, il existe $x_1, \dots, x_m \in K$ tels que :

$$K = \bigcup_{i=1}^{m} V(x_i)\,.$$

Si $N = \max\{N(x_1), \dots, N(x_m)\}$, on a, pour $n \geqslant N$:

$$0 \leqslant u(x) - u_n(x) \leqslant \varepsilon\,, \quad \forall x \in K\,,$$

car x appartient à l'un des $V(x_i)$ et $n \geqslant N(x_i)$; donc :

$$\begin{aligned} 0 \leqslant u(x) - u_n(x) &\leqslant u(x) - u_{N(x_i)}(x) \\ &\leqslant \big(u(x) - u(x_i)\big) + \big(u(x_i) - u_{N(x_i)}(x_i)\big) + \big(u_{N(x_i)}(x_i) - u_{N(x_i)}(x)\big) \\ &\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon . \end{aligned}$$ □

Étape 2. Si $f \in A$, alors $|f| \in \overline{A}$.

Preuve. En effet, on peut supposer $f \neq 0$. Soit $a = \|f\|_\infty$. On a $[f(x)]^2/a^2 \in [0,1]$ pour tout $x \in K$. Mais, comme r_n est un polynôme, et A est une algèbre, on a $r_n(f^2/a^2) \in A$ si $f \in A$. En passant à la limite, on obtient :

$$|f| = a \lim_{n\to\infty} r_n(f^2/a^2) \in \overline{A},$$

la limite étant uniforme, c'est-à-dire prise pour la norme de $\mathscr{C}_{\mathbb{R}}(K)$. □

Étape 3. Si $f, g \in A$, alors $\max\{f,g\}, \min\{f,g\} \in \overline{A}$.

Preuve. Il suffit de remarquer que :

$$\begin{cases} \max\{f,g\} = \dfrac{1}{2}\,(f+g+|f-g|) \\ \min\{f,g\} = \dfrac{1}{2}\,(f+g-|f-g|), \end{cases}$$

et d'utiliser l'Étape 2 (ainsi que le fait que $\overline{A}$ est un sous-espace vectoriel). □

Étape 3 bis. Si $f, g \in \overline{A}$, alors $\max\{f,g\}, \min\{f,g\} \in \overline{A}$.

Preuve. Cela résulte de ce que $\overline{A}$ vérifie les conditions demandées pour A : elle reste une sous-algèbre (rappelons que la convergence dans $\mathscr{C}(K)$ est la convergence uniforme), et, puisque A contient les constantes et sépare les points de K, il en est *a fortiori* de même pour $\overline{A}$. □

Bien sûr, par récurrence :

$$f_1, \dots, f_n \in \overline{A} \quad \Longrightarrow \quad \max\{f_1,\dots,f_n\} \in \overline{A} \quad \text{et} \quad \min\{f_1,\dots,f_n\} \in \overline{A}.$$

Étape 4. Si $x, y \in K$ et $x \neq y$, alors :

$$(\forall \alpha, \beta \in \mathbb{R}) \quad (\exists h \in A) \qquad h(x) = \alpha \quad \textit{et} \quad h(y) = \beta.$$

C'est la première étape dans l'approximation : on peut obtenir avec une fonction de A des valeurs données en deux points donnés distincts de K.

Preuve. Comme A ***sépare les points***, il existe $g \in A$ telle que $g(x) \neq g(y)$. Posons :

$$h = \alpha\,\mathbb{1} + \frac{\beta - \alpha}{g(y) - g(x)}\,\big(g - g(x)\,\mathbb{1}\big).$$

On a bien $h(x) = \alpha$, $h(y) = \beta$, et $h \in A$, car $g \in A$, $\mathbb{1} \in A$, et A est un sous-espace vectoriel. □

Étape 5. Pour toute $f \in \mathscr{C}(K)$, pour tout $x \in K$, et tout $\varepsilon > 0$, il existe $g \in \overline{A}$ telle que :

$$g(x) = f(x) \quad et \quad g(y) \leqslant f(y) + \varepsilon, \forall y \in K.$$

Preuve. Pour tout $z \in K$ tel que $z \neq x$, il existe, par l'Étape 4, en prenant $\alpha = f(x)$ et $\beta = f(z)$, une $h_z \in A$ telle que $h_z(x) = f(x)$ et $h_z(z) = f(z)$.

Notons h_x la fonction constante égale à $f(x)\,\mathbb{1}$. Alors :

$$(\forall z \in K) \qquad h_z(x) = f(x) \quad \text{et} \quad h_z(z) = f(z).$$

La continuité de f et celle de h_z donnent un voisinage, que l'on peut prendre ouvert, V_z de z tel que :

$$y \in V(z) \quad \Longrightarrow \quad h_z(y) \leqslant f(y) + \varepsilon.$$

Comme K est compact, il existe un nombre fini d'éléments $z_1, \dots, z_m \in K$ tels que :

$$K = V(z_1) \cup \dots \cup V(z_m).$$

Alors $g = \inf\{h_{z_1}, \dots, h_{z_m}\} \in \overline{A}$, par l'Étape 3 *bis*, et l'on a, pour tout $y \in K$: $g(y) \leqslant f(y) + \varepsilon$, puisque y appartient à l'un des $V(z_i)$. □

Étape 6. On a $\overline{A} = \mathscr{C}_{\mathbf{R}}(K)$.

Preuve. Soit $f \in \mathscr{C}_{\mathbf{R}}(K)$, et soit $\varepsilon > 0$.

Pour tout $x \in K$, il existe $g_x \in \overline{A}$ vérifiant les conditions données dans l'Étape 5.

La continuité de f et celle de g_x donnent un voisinage, que l'on peut choisir ouvert, $U(x)$ de x tel que :

$$y \in U(x) \quad \Longrightarrow \quad g_x(y) \geqslant f(y) - \varepsilon.$$

La compacité de K permet de trouver un nombre fini d'éléments $x_1, \dots, x_p \in K$ tels que :

$$K = U(x_1) \cup \dots \cup U(x_p).$$

Alors $\varphi = \max\{g_{x_1}, \dots, g_{x_p}\} \in \overline{A}$, grâce à l'Étape 3 bis ; et elle vérifie :

$$f(y) - \varepsilon \leqslant \varphi(y) \leqslant f(y) + \varepsilon, \qquad \forall y \in K,$$

car chaque $y \in K$ est dans l'un des $U(x_j)$.

Cela veut dire que $\|f - \varphi\|_\infty \leqslant \varepsilon$.

Comme $\varepsilon > 0$ était arbitraire, on a bien $f \in \overline{(\overline{A})} = \overline{A}$.

Cela achève la preuve du Théorème II.4.1. □

La preuve de Bernstein pour un intervalle compact de $\mathbb{R}$

La forme générale du Théorème de Stone-Weierstrass a été donnée par Stone en 1948. À l'origine, Weierstrass avait montré, en 1885, que toute fonction continue sur un intervalle fermé borné de $\mathbb{R}$ pouvait y être approchée uniformément par des polynômes. Il utilisait pour cela un produit de convolution (voir le chapitre suivant).

En 1913, Bernstein en a donné une belle preuve probabiliste, que l'on va exposer ci-dessous. Notons d'abord que, par un changement de variable, on peut supposer que l'intervalle en question est $[0,1]$.

L'idée de départ est la suivante : on fixe $t \in [0,1]$ (aussi bien, si on veut, on peut ne prendre que $0 < t < 1$), et on considère des *variables aléatoires indépendantes* $X_1, \ldots, X_n$ suivant toutes la *loi de Bernoulli* de paramètre t. Alors $S_n = X_1 + \cdots + X_n$ suit la loi binômiale $\mathscr{B}(n,t)$ de paramètres n et t. La loi faible des grands nombres dit que $\frac{S_n}{n} \xrightarrow[n\to\infty]{} t = \mathbb{E}(X_1)$ en probabilité. Alors, pour toute fonction f continue sur $[0,1]$, on a $\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] \xrightarrow[n\to\infty]{} f(t)$. En effet, si $\varepsilon > 0$ est donné, l'uniforme continuité de f sur $[0,1]$ permet de trouver $\delta > 0$ tel que $|f(x) - f(x')| \leqslant \varepsilon$ pour $|x - x'| \leqslant \delta$; la convergence en probabilité donne alors un $N \geqslant 1$ tel que $\mathbb{P}\left(\left|\frac{S_n}{n} - t\right| > \delta\right) \leqslant \varepsilon$ si $n \geqslant N$. Alors, pour $n \geqslant N$, on a :

$$\begin{aligned}\left|\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] - f(t)\right| &= \int_{\left\{\left|\frac{S_n}{n} - t\right| > \delta\right\}} \left|f\left[\frac{S_n(\omega)}{n}\right] - f(t)\right| d\mathbb{P}(\omega) \\ &\quad + \int_{\left\{\left|\frac{S_n}{n} - t\right| \leqslant \delta\right\}} \left|f\left[\frac{S_n(\omega)}{n}\right] - f(t)\right| d\mathbb{P}(\omega) \\ &\leqslant 2\|f\|_\infty \varepsilon + \varepsilon .\end{aligned}$$

Or $\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^n C_n^k t^k (1-t)^{n-k} f\left(\frac{k}{n}\right)$. On pose :

$$[B_n(f)](t) = \sum_{k=0}^n C_n^k t^k (1-t)^{n-k} f\left(\frac{k}{n}\right) ;$$

c'est un polynôme de degré n. On l'appelle le $n^{\text{ème}}$ *polynôme de Bernstein* de f.

On vient de voir que l'on a convergence simple de $B_n(f)$ vers f.

Nous allons voir que, grâce à une estimation uniforme de la variance des variables de Bernoulli, la preuve de la loi faible des grands nombres pour ces variables permet d'obtenir la convergence uniforme de $B_n(f)$ vers f.

Rappelons d'abord que si X est une variable aléatoire suivant la loi de Bernoulli de paramètre t, alors sa variance vaut $\mathrm{Var}(X) = t(1-t)$. On a, par l'inégalité de Bienaymé-Tchebychev, pour tout $\delta > 0$:

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n} - t\right| > \delta\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X)\right| > \delta\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| > \delta\right) \\ &\leqslant \frac{1}{\delta^2} \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2\delta^2} \mathrm{Var}(S_n) \\ &= \frac{1}{n^2\delta^2} \sum_{j=1}^n \mathrm{Var}(X_j) \quad \text{(par indépendance)} \\ &= \frac{\mathrm{Var(X)}}{n\delta^2} = \frac{t(1-t)}{n\delta^2} \leqslant \frac{1/4}{n\delta^2} .\end{aligned}$$

Considérons le *module de continuité* de f, défini par :

$$\omega_f(h) = \sup\{|f(t) - f(t')| \,;\ |t - t'| \leqslant h\} .$$

Dire que f est uniformément continue signifie que $\omega_f(h) \underset{h\to 0}{\longrightarrow} 0$. Fixons un $\delta > 0$, que l'on précisera après. On a, pour tout $t \in [0,1]$ (on prendra garde à différencier l'occurence $\omega \in \Omega$ du module de continuité ω_f ; on aurait pu modifier ces notations, mais ce sont celles habituellement utilisées !) :

$$\begin{aligned}
|f(t) - [B_n(f)](t)| &= \left|\mathbb{E}\left[f(t) - f\Big(\frac{S_n}{n}\Big)\right]\right| \\
&\leqslant \mathbb{E}\left(\Big|f(t) - f\Big(\frac{S_n}{n}\Big)\Big|\right) = \int_\Omega \Big|f(t) - f\Big(\frac{S_n(\omega)}{n}\Big)\Big|\, d\mathbb{P}(\omega) \\
&= \int_{\{|t-\frac{S_n(\omega)}{n}|\leqslant\delta\}} + \int_{\{|t-\frac{S_n(\omega)}{n}|>\delta\}} \\
&\leqslant \omega_f(\delta) + 2\,\|f\|_\infty \mathbb{P}\left(\Big|t - \frac{S_n}{n}\Big| > \delta\right) \\
&\leqslant \omega_f(\delta) + 2\,\|f\|_\infty \frac{1}{4n\delta^2}\,\cdot
\end{aligned}$$

Pour tout $\varepsilon > 0$ donné, choisissons maintenant δ de sorte que $\omega_f(\delta) \leqslant \varepsilon/2$, puis $N \geqslant 1$ tel que $\|f\|_\infty \frac{1}{2N\delta^2} \leqslant \varepsilon/2$. On aura, pour $n \geqslant N$, $|f(t) - [B_n(f)]t)| \leqslant \varepsilon$ pour tout $t \in [0,1]$, ce qui prouve que $B_n(f)$ tend uniformément vers f. □

II.4.1.2. Cas complexe

Tel quel, l'énoncé du Théorème II.4.1 est faux pour les espaces de fonctions à valeurs complexes. Par exemple, si K est le disque unité fermé $\overline{\mathbb{D}}$ du plan complexe, toute limite uniforme sur K de polynômes p_n est holomorphe dans le disque ouvert $\mathbb{D}$, grâce au Théorème de Weierstrass sur la convergence uniforme des suites de fonctions holomorphes. L'adhérence de l'algèbre des polynômes n'est donc pas $\mathscr{C}_\mathbb{C}(K)$ tout entier : par exemple, la fonction $z \mapsto \bar{z}$ n'est pas dedans. En fait, cet exemple est essentiellement le seul cas dont il faut tenir compte ; en effet, on a :

Théorème II.4.6 (Théorème de Stone-Weierstrass, cas complexe).
Soit K un espace compact et soit A une sous-algèbre, complexe, de l'espace de Banach complexe $\mathscr{C}_\mathbb{C}(K)$. Si :
- *a) A sépare les points de K ;*
- *b) A contient les fonctions constantes ;*
- *c) A est* **stable par conjugaison** *: $f \in A \Rightarrow \overline{f} \in A$,*

alors A est dense dans $\mathscr{C}_\mathbb{C}(K)$.

Notons qu'ici $\overline{f}$ désigne la fonction $t \in K \mapsto \overline{f(t)} \in \mathbb{C}$, où $\overline{f(t)}$ est le nombre complexe conjugué de $f(t)$.

Preuve. La condition *c)* permet de dire que :

$$f \in A \implies \operatorname{Re} f = \frac{f + \overline{f}}{2} \in A \quad \text{et} \quad \operatorname{Im} f = \frac{f - \overline{f}}{2i} \in A\,.$$

Soit :

$$A_{\mathbb{R}} = \{f \in A\,;\ f(t) \in \mathbb{R}\,,\ \forall t \in K\}\,.$$

La remarque ci-dessus permet de dire que :

$$A = A_{\mathbb{R}} + iA_{\mathbb{R}}\,.$$

De plus, $A_{\mathbb{R}}$ est une sous-algèbre de $\mathscr{C}_{\mathbb{R}}(K)$, qui contient les fonctions constantes (réelles), et sépare les points de K : si $u \neq v$, il existe $f \in A$ telle que $f(u) \neq f(v)$; mais alors $\operatorname{Re} f(u) \neq \operatorname{Re} f(v)$ ou $\operatorname{Im} f(u) \neq \operatorname{Im} f(v)$, et $\operatorname{Re} f, \operatorname{Im} f \in A_{\mathbb{R}}$. Il résulte du cas réel que $A_{\mathbb{R}}$ est dense dans $\mathscr{C}_{\mathbb{R}}(K)$. Mais alors, $A = A_{\mathbb{R}} + iA_{\mathbb{R}}$ est dense dans $\mathscr{C}_{\mathbb{C}}(K) = \mathscr{C}_{\mathbb{R}}(K) + i\mathscr{C}_{\mathbb{R}}(K)$. □

Exemple. Soit K une partie compacte de $\mathbb{C}$. L'ensemble des polynômes, à coefficients complexes, en les **deux variables z et $\bar{z}$** est dense dans $\mathscr{C}_{\mathbb{C}}(K)$.

On notera que c'est aussi, en identifiant $\mathbb{C}$ à $\mathbb{R}^2$, l'ensemble des polynômes, à coefficients complexes, en les deux variables réelles x et y, en identifiant $z = x+iy \in \mathbb{C}$ avec $(x,y) \in \mathbb{R}^2$.

II.4.2. Le système trigonométrique

Nous allons considérer ici des fonctions $f\colon \mathbb{R} \to \mathbb{C}$ **périodiques**, de période 1 sur $\mathbb{R}$.

L'application *surjective* :

$$\begin{array}{rccl} e_1\colon & \mathbb{R} & \longrightarrow & \mathbb{U} = \{z \in \mathbb{C}\,;\ |z| = 1\} \\ & t & \longmapsto & \mathrm{e}^{2\pi i t} = u \end{array}$$

permet des les identifier aux fonctions définies sur $\mathbb{U}$. On peut aussi les identifier aux fonctions définies sur le tore $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.

De plus, on sait que pour toute fonction f continue sur $\mathbb{R}$ de période 1, il existe une unique fonction continue $\tilde{f}\colon \mathbb{U} \to \mathbb{C}$ telle que $f = \tilde{f} \circ e_1$ (resp. $\ddot{f}\colon \mathbb{T} \to \mathbb{C}$ telle que $f(x) = \ddot{f}(x + \mathbb{Z})$). L'espace $\mathscr{C}_1(\mathbb{R})$ des fonctions continues sur $\mathbb{R}$ de période 1, muni de la norme $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$, s'identifie donc à l'espace $\mathscr{C}(\mathbb{U})$ des fonctions continues sur le compact $\mathbb{U}$.

Il s'identifie aussi au sous-espace $\widetilde{\mathscr{C}} = \{f \in \mathscr{C}([0,1])\,;\ f(0) = f(1)\}$.

Ces identifications sont isométriques puisque :

$$\sup_{x\in\mathbb{R}} |f(x)| = \sup_{x\in[0,1]} |f(x)| = \sup_{u\in\mathbb{U}} |\tilde{f}(u)| = \sup_{\xi\in\mathbb{T}} |\ddot{f}(\xi)|\,.$$

Définition II.4.7. *On appelle* **polynôme trigonométrique** *toute somme finie*

$$\sum_{n=N_1}^{N_2} a_n \mathrm{e}^{2\pi i n t}$$

avec $a_n \in \mathbb{C}$ et $N_1, N_2 \in \mathbb{Z}$, $N_1 \leqslant N_2$.

Notons qu'en ajoutant au besoin des coefficients nuls, on peut toujours écrire un polynôme trigonométrique sous la forme symétrique :

$$\sum_{n=-N}^{N} a_n e^{2\pi int},$$

où N est un entier positif.

On notera, pour tout $n \in \mathbb{Z}$:

$$\boxed{e_n(t) = e^{2\pi int}}, \qquad t \in \mathbb{R}.$$

L'ensemble $\{e_n\,;\ n \in \mathbb{Z}\}$ s'appelle le **système trigonométrique.**

Les polynômes trigonométriques s'identifient aux polynômes usuels en u et $\bar{u}$ sur $\mathbb{U}$, puisque tout $u \in \mathbb{U}$ s'écrit sous la forme $u = e_1(t) = e^{2\pi it}$, et qu'alors $u^n = e^{2\pi int} = e_n(t)$, et que $\bar{u} = e^{-2\pi int} = e_{-n}(t)$. Le théorème de Stone-Weierstrass complexe appliqué à $\mathscr{C}_{\mathbb{C}}(\mathbb{U})$ donne donc :

Théorème II.4.8. *L'ensemble des polynômes trigonométriques est dense dans l'espace des fonctions continues de période* 1 *sur* $\mathbb{R}$.

Considérons maintenant l'espace des fonctions $f\colon \mathbb{R} \to \mathbb{C}$ mesurables de période 1 telles que :

$$\int_0^1 |f(t)|^2\,dt < +\infty.$$

Lorsque l'on le quotiente par le sous-espace des fonctions négligeables, ce quotient s'identifie à $L^2(0,1) = L^2_{\mathbb{C}}(0,1)$; en effet, pour toute fonction mesurable $g\colon [0,1] \to \mathbb{C}$, la fonction mesurable :

$$\begin{array}{rccl} \widetilde{g}\colon & [0,1] & \longrightarrow & \mathbb{C} \\ & t & \longmapsto & \begin{cases} g(t) \text{ si } 0 \leqslant t < 1\,; \\ g(0) \text{ si } t = 1 \end{cases} \end{array}$$

se prolonge par périodicité en une fonction mesurable $f\colon \mathbb{R} \to \mathbb{C}$ de période 1, et $\int_0^1 |g(t)|^2\,dt = \int_0^1 |f(t)|^2\,dt$.

Ces identifications étant faites, on peut énoncer :

Théorème II.4.9. *Le système trigonométrique est une* base orthonormée *de* $L^2_{\mathbb{C}}(0,1)$.

Corollaire II.4.10. *L'espace* réel $L^2_{\mathbb{R}}(0,1)$ *possède une base orthonormée formée des fonctions :*

$$1, \sqrt{2}\cos(2\pi t), \sqrt{2}\cos(4\pi t), \dots, \sqrt{2}\cos(2\pi n t), \dots,$$
$$\sqrt{2}\sin(2\pi t), \sqrt{2}\sin(4\pi t), \dots, \sqrt{2}\sin(2\pi n t), \dots.$$

Remarques. 1) $\mathbb{Z}$ étant dénombrable, on pourrait ré-indexer le système trigonométrique avec les entiers positifs.

2) Le théorème signifie que, pour toute $f \in L^2_{\mathbb{C}}(0,1)$, on a :

$$\lim_{N\to\infty} \left\| f - \sum_{n=-N}^{N} \widehat{f}(n)\, e_n \right\|_2 = 0,$$

où les produits scalaires :

$$\boxed{\widehat{f}(n) = (f \mid e_n) = \int_0^1 f(t)\,\overline{e_n(t)}\,dt = \int_0^1 f(t)\,\mathrm{e}^{-2\pi i n t}\,dt},$$

pour $n \in \mathbb{Z}$, sont appelés les **coefficients de Fourier** de f. La formule de Parseval s'écrit alors $\boxed{\int_0^1 |f(x)|^2\,dx = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2}$.

Nous savons qu'il existe alors une suite strictement croissante d'entiers $(l_n)_{n\geqslant 1}$ telle que :

$$\lim_{n\to\infty} \sum_{k=-l_n}^{l_n} \widehat{f}(k)\,\mathrm{e}^{2\pi i k t} = f(t)$$

pour presque tout $t \in [0,1]$.

Répondant à une question posée par Lusin en 1913, L. Carleson a montré en 1966 qu'en fait, **sans prendre de sous-suite**, on a, pour toute $f \in L^2(0,1)$:

$$\lim_{n\to\infty} \sum_{k=-n}^{n} \widehat{f}(k)\,\mathrm{e}^{2\pi i k t} = f(t)$$

pour presque tout $t \in [0,1]$. C'est un résultat d'une **extrême** difficulté.

Preuve du théorème II.4.9. Il est d'abord facile de voir que $\{e_n\,;\ n \in \mathbb{Z}\}$ est orthonormé :

$$(e_n \mid e_p) = \int_0^1 \mathrm{e}^{2\pi i n t}\,\mathrm{e}^{-2\pi i p t}\,dt = \int_0^1 \mathrm{e}^{2\pi i (n-p) t}\,dt = \begin{cases} 1 \text{ si } n = p\,; \\ 0 \text{ si } n \neq p\,. \end{cases}$$

Il est total car les polynômes trigonométriques sont denses dans $\mathscr{C}(\mathbb{U})$ et $\|\cdot\|_\infty \geqslant \|\cdot\|_2$, en utilisant le lemme suivant :

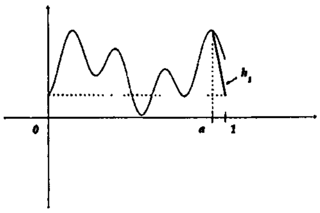
Lemme II.4.11. *L'ensemble $\mathscr{C}_1(\mathbb{R})$ des fonctions continues sur $\mathbb{R}$ de période 1, identifié à $\widetilde{\mathscr{C}} = \{f \in \mathscr{C}([0,1])\,;\ f(0) = f(1)\}$, est dense dans $L^2(0,1)$.*

En effet, si $f \in L^2(0,1)$, il existe alors, pour tout $\varepsilon > 0$, $g \in \widetilde{\mathscr{C}} \cong \mathscr{C}(\mathbb{U})$ telle que $\|f - g\|_2 \leqslant \varepsilon/2$; il existe ensuite un polynôme trigonométrique p tel que $\|g - p\|_\infty \leqslant \varepsilon/2$; mais $\|g - p\|_2 \leqslant \|g - p\|_\infty \leqslant \varepsilon/2$; donc $\|f - p\|_2 \leqslant \varepsilon$. □

Preuve du lemme. Soit $f \in L^2(0,1)$ et soit $\varepsilon > 0$. Nous savons (Corollaire II.2.9) qu'il existe $h \in \mathscr{C}([0,1])$ telle que $\|f - h\|_2 \leqslant \varepsilon/2$.

Soit $M > 0$ tel que $|h(t)| \leqslant M$ pour tout $t \in [0,1]$, et notons $a = 1 - \left(\frac{\varepsilon}{4M}\right)^2$.

Nous allons modifier h sur $[a,1]$ en posant $h_1(1) = h(0)$ et en prenant h_1 affine entre a et 1. Alors $h_1 \in \widetilde{\mathscr{C}}$, $\|h_1\|_\infty \leqslant M$, et :

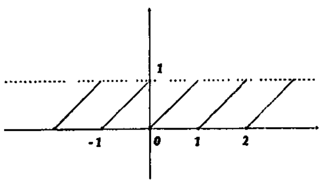


$$\|h - h_1\|_2 = \left(\int_a^1 |h(t) - h_1(t)|^2\,dt\right)^{1/2}$$
$$\leqslant (1-a)^{1/2} \sup_{a \leqslant t \leqslant 1} \big(|h(t)| + |h_1(t)|\big)$$
$$\leqslant \frac{\varepsilon}{4M} \times (M + M) = \frac{\varepsilon}{2}\cdot$$

On a donc $\|f - h_1\|_2 \leqslant \varepsilon$. □

Exemple d'application.

Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f(t) = t$ pour $0 \leqslant t < 1$, et prolongée par périodicité sur $\mathbb{R}$.



Alors $f \in L^2(0,1)$ et :

$$\|f\|_2^2 = \int_0^1 t^2\,dt = \frac{1}{3}\cdot$$

Les coefficients de Fourier de f sont :

$$\widehat{f}(n) = \int_0^1 t\,e^{-2\pi i n t}\,dt, \quad n \in \mathbb{Z}.$$

Pour $n = 0$: $\int_0^1 t\,dt = 1/2$; pour $n \neq 0$:

$$\widehat{f}(n) = \left[\frac{t\,e^{-2\pi i n t}}{-2\pi i n}\right]_0^1 - \int_0^1 \frac{e^{-2\pi i n \cancel{t}}}{-\cancel{2}\pi i n}\,dt = \frac{1}{-2\pi i n} = \frac{i}{2\pi n}\cdot$$

La formule de Parseval $\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2$ donne donc :

$$\frac{1}{3} = \frac{1}{4} + \sum_{n\neq 0} \frac{1}{4\pi^2 n^2} = \frac{1}{4} + \frac{1}{2\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\,;$$

d'où :

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = 2\pi^2\Big(\frac{1}{3}-\frac{1}{4}\Big) = \frac{\pi^2}{6}\cdot$$

II.4.2.1. Coefficient de Fourier des fonctions de $L^1(0,1)$

Pour toute $f\colon [0,1]\to\mathbb{C}$ mesurable, l'inégalité de Cauchy-Schwarz :

$$\int_0^1 |f(t)|\,dt \leqslant \Big(\int_0^1 1^2\,dt\Big)^{1/2}\Big(\int_0^1 |f(t)|^2\,dt\Big)^{1/2} = \Big(\int_0^1 |f(t)|^2\,dt\Big)^{1/2}$$

dit que $\mathscr{L}^2([0,1]) \subseteq \mathscr{L}^1([0,1])$. On a donc une injection naturelle de $L^2(0,1)$ dans $L^1(0,1)$. Par identification de $L^2(0,1)$ avec son image dans $L^1(0,1)$, on écrira : $\boxed{L^2(0,1)\subseteq L^1(0,1)}$.

Pour toute $f\in L^1(0,1)$, on peut définir les **coefficients de Fourier** :

$$\boxed{\widehat{f}(n) = \int_0^1 f(t)\,\mathrm{e}^{-2\pi int}\,dt}\,,\qquad n\in\mathbb{Z},$$

puisque $|\mathrm{e}^{-2\pi int}| = 1$. On a $\boxed{|\widehat{f}(n)| \leqslant \|f\|_1}$ pour tout $n\in\mathbb{Z}$. De plus :

> **Théorème II.4.12** (Lemme de Riemann-Lebesgue). *Pour toute fonction $f\in L^1(0,1)$, ses coefficients de Fourier tendent vers* 0 *quand $|n|$ tend vers l'infini :*
>
> $$\boxed{\widehat{f}(n)\xrightarrow[|n|\to\infty]{}0}\,.$$

Preuve. Si $g\in L^2(0,1)$, la formule de Parseval :

$$\|g\|_2^2 = \sum_{n\in\mathbb{Z}} |\widehat{g}(n)|^2$$

montre que l'on a, en particulier, $\widehat{g}(n)\xrightarrow[|n|\to\infty]{}0$.

Maintenant, si $f\in L^1(0,1)$, il existe, pour tout $\varepsilon>0$, une fonction $g\in L^2(0,1)$ (par exemple g étagée, ou bien g continue) telle que $\|f-g\|_1\leqslant\varepsilon$. Comme on a $|\widehat{f}(n)-\widehat{g}(n)|\leqslant\|f-g\|_1\leqslant\varepsilon$, on obtient $|\widehat{f}(n)|\leqslant|\widehat{g}(n)|+\varepsilon$; donc :

$$\limsup_{|n|\to\infty}|\widehat{f}(n)| \leqslant \limsup_{|n|\to\infty}|\widehat{g}(n)| + \varepsilon = \varepsilon.$$