Chapitre 9

Espaces de Hilbert ; bases hilbertiennes

Dans tout ce chapitre, E désignera un $\mathbb{R}$-espace vectoriel.

9.1 Vocabulaire et premiers résultats sur les produits scalaires

Définition 9.1. Un *produit scalaire* sur E est une application $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ qui est :
- *symétrique*, c'est-à-dire que pour tout $x, y \in E$ on a $\langle x, y \rangle = \langle y, x \rangle$;
- *bilinéaire*, c'est-à-dire que pour tout $x, y, z \in E$ et tout $\lambda \in \mathbb{R}$ on a $\langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle$ (et donc aussi $\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle y, z \rangle$ par symétrie) ;
- *définie positive*, c'est-à-dire que pour tout $x \in E$ on a $\langle x, x \rangle \geq 0$ et $\langle x, x \rangle = 0 \Leftrightarrow x = 0_E$.

Exercice 9.2. Montrer que, si $(X, \mathcal{A}, \mu)$ est un espace mesuré, alors $\langle \cdot, \cdot \rangle$ défini par

$$\langle f, g \rangle = \int_X f g d\mu$$

est un produit scalaire sur $L^2(X)$.

Dans la suite, la notation $\langle \cdot, \cdot \rangle$ désignera toujours un produit scalaire sur E.

Notons qu'ici on travaillera avec des espaces vectoriels réels ; on peut aussi définir un produit scalaire pour des espaces vectoriels complexes mais dans ce cas-là on doit demander qu'il soit *sesquilinéaire* plutôt que bilinéaire.

Pour nous habituer aux calculs avec les produits scalaires, notons deux conséquences faciles. La première est qu'on connaît les valeurs de $\langle x, y \rangle$ pour tout x, y dès qu'on connaît la valeur de $\langle x, x \rangle$ pour tout $x \in E$.

Proposition 9.3 (Identité de polarisation)**.** *Pour tout $x, y \in E$ on a*

$$4\langle x, y \rangle = \langle x + y, x + y \rangle - \langle x - y, x - y \rangle \ .$$

Démonstration. Par bilinéarité et symétrie, on a

$$\begin{aligned} \langle x + y, x + y \rangle &= \langle x, x + y \rangle + \langle y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle \ . \end{aligned}$$

Un calcul similaire (ou le résultat précédent appliqué à $y' = -y$) donne

$$\langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle - 2\langle x, y \rangle$$

d'où le résultat. □

Proposition 9.4 (Identité du parallélogramme)**.** *Pour tout $x, y \in E$ on a*

$$\langle x + y, x + y \rangle + \langle x - y, x - y \rangle = 2\langle x, x \rangle + 2\langle y, y \rangle \ .$$

La preuve est laissée en exercice.

Théorème 9.5 (Inégalité de Cauchy-Schwarz)**.** *Pour tout $x, y \in E$ on a*

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle \ .$$

Avant de donner la preuve de l'inégalité de Cauchy-Schwarz, notons tout de suite une conséquence fondamentale.

Corollaire 9.6. L'application $x \mapsto \sqrt{\langle x, x \rangle}$ est une norme sur E.

On notera dans la suite

$$\|x\| = \sqrt{\langle x, x \rangle} \ .$$

Preuve du corollaire. L'axiome de séparation se déduit immédiatement du fait que le produit scalaire est défini positif. Celui d'homogénéité suit de sa bilinéarité. Il reste à vérifier l'inégalité triangulaire : soit $x, y \in E$. On a

$$\begin{aligned} \langle x+y, x+y \rangle &= \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle \\ &\leq \langle x, x \rangle + \langle y, y \rangle + 2\sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle} && \text{(Cauchy-Schwarz)} \\ &= \left(\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}\right)^2 \end{aligned}$$

Comme toutes les quantités en jeu sont positives, on peut passer à la racine carrée de chaque côté de l'inégalité et obtenir comme espéré

$$\sqrt{\langle x+y, x+y \rangle} \leq \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle} \ .$$

□

Preuve de l'inégalité de Cauchy-Schwarz. Fixons $x, y \in E$, et considérons l'application définie sur $\mathbb{R}$ par

$$t \mapsto \langle x + ty, x + ty \rangle = \langle x, x \rangle + 2t\langle x, y \rangle + t^2 \langle y, y \rangle.$$

Il s'agit d'une application polynomiale, de degré 2, à valeurs positives. Le polynôme en question ne peut avoir qu'une racine réelle au plus (sans quoi il changerait de signe) : son discriminant doit donc être négatif ou nul. Autrement dit, on doit avoir

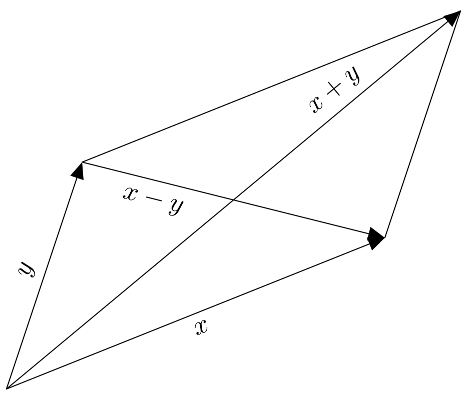
$$4\langle x, y \rangle^2 - 4\langle x, x \rangle \langle y, y \rangle \leq 0 \ .$$

C'est exactement l'inégalité qu'on cherchait à démontrer. □

Jusqu'à nouvel ordre, E désigne un espace vectoriel réel muni du produit scalaire $\langle \cdot, \cdot \rangle$ et de la norme $\| \cdot \|$ associée au produit scalaire.

En utilisant la norme, les identités de polarisation et du parallélogramme s'écrivent ainsi :

$$\begin{aligned} &\forall x, y \in E \quad 4\langle x, y \rangle = \|x+y\|^2 - \|x-y\|^2 && \text{(polarisation)} \\ &\forall x, y \in E \quad 2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 && \text{(parallélogramme)} \end{aligned}$$



Exercice 9.7. Montrer que le produit scalaire définit une application continue de $E \times E$ vers $\mathbb{R}$.

Définition 9.8. Soit A un sous-ensemble de E. On définit

$$A^\perp = \{x \in E\colon \forall a \in A\ \langle a, x\rangle = 0\}\ .$$

On appelle $A^\perp$ l'*orthogonal* de A.

Exercice 9.9. Montrer que, pour $A \subseteq E$:

1. $A^\perp$ est un sous-espace vectoriel de E.
2. $A^\perp$ est fermé dans E.
3. $A^\perp = (\overline{A})^\perp$.

Définition 9.10. Une famille $(a_i)_{i\in I}$ de vecteurs non nuls de E est dite *orthogonale* si pour tout $i \neq j$ on a $\langle a_i, a_j\rangle = 0$.

La famille est dite *orthonormale* si elle est orthogonale et de plus $\|a_i\| = 1$ pour tout $i \in I$.

Proposition 9.11. *Toute famille orthogonale est une famille libre.*

Démonstration. Soit $(a_i)_{i\in I}$ une famille orthogonale et $a_{i_1}, \ldots, a_{i_n}$, $\lambda_1, \ldots, \lambda_n$ tels que $\displaystyle\sum_{k=1}^n \lambda_k a_{i_k} = 0$.

Observons que pour tout $j \in \{1, \ldots, n\}$ on a

$$\begin{aligned}
\left\langle \sum_{k=1}^n \lambda_k a_{i_k}, a_{i_j} \right\rangle &= \sum_{k=1}^n \lambda_k \langle a_{ik}, a_{i_j}\rangle \\
&= \sum_{k\neq j} \lambda_k \langle a_{ik}, a_{i_j}\rangle + \lambda_j \|a_{i_j}\|^2 \\
&= \lambda_j \|a_{i_j}\|^2
\end{aligned}$$

On en déduit que $\lambda_j\|a_{i_j}\|^2 = 0$ pour tout j puis, puisque tous les a_{i_j} sont non nuls, que $\lambda_j = 0$ pour tout j : la famille est bien libre. □

Exercice 9.12. Montrer que, si $(a_i)_{i\in I}$ est une famille orthogonale de vecteurs de E on a pour tout $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ et tout $i_1, \ldots, i_n \in I$

$$\|\sum_{k=1}^n \lambda_k a_{i_k}\|^2 = \sum_{k=1}^n \lambda_k^2 \|a_{i_k}\|^2\ .$$

Exercice 9.13. On se place dans l'espace $L^2([0,1], \mathcal{B}([0,1]), \lambda)$ muni du produit scalaire $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)dx$. On considère la famille $(u_k)_{k\geq 1}$ définie par $u_k(x) = \sin(k\pi x)$ pour tout $x \in [0,1]$ et tout $k \geq 1$. Montrer que c'est une famille orthogonale.

9.2 Espaces métriques complets

Définition 9.14. Soit (X, d) un espace métrique, et (x_n) une suite d'éléments de X. On dit que (x_n) est une *suite de Cauchy* si la propriété suivante est vérifiée :

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N\ d(x_n, x_m) \leq \varepsilon\ .$$

Autrement dit : à partir d'un certain rang, les termes de la suite deviennent arbitrairement proches les uns des autres.

Exercice 9.15. Soit (X, d) un espace métrique. Montrer que toute suite convergente est de Cauchy, et que la réciproque est fausse en général (on pourra par exemple considérer la suite définie par $x_n = 2^{-n}$ dans l'espace $X =]0, +\infty[$ muni de sa distance usuelle).

Exercice 9.16. Soit (X, d) un espace métrique, et (x_n) une suite de Cauchy d'éléments de X. Montrer que (x_n) est bornée.

Définition 9.17. Soit (X,d) un espace métrique. On dit que (X,d) est *complet* si toute suite de Cauchy d'éléments de X est convergente.

Proposition 9.18. *Soit (X,d) un espace métrique complet, et $F \subseteq X$ un sous-ensemble fermé. Alors (F,d) est complet.*

Démonstration. Soit (x_n) une suite de Cauchy d'éléments de F. C'est en particulier une suite de Cauchy d'éléments de X, elle converge donc vers un certain $x \in X$ puisque (X,d) est complet. Comme F est fermé dans X, on doit avoir $x \in F$. On vient de montrer que toute suite de Cauchy d'éléments de F converge dans F : (F,d) est complet. $\square$

Cette proposition admet une forme de réciproque.

Proposition 9.19. *Soit (X,d) un espace métrique et $F \subseteq X$ un sous-ensemble tel que (F,d) est complet. Alors F est fermé dans X.*

Démonstration. Soit (x_n) une suite d'éléments de F qui converge vers $x \in X$. Alors (x_n) est de Cauchy puisque toute suite convergente est de Cauchy ; comme F est complet (x_n) converge dans F, ce qui montre que $x \in F$. $\square$

Proposition 9.20. *Soit (X,d) un espace métrique, et (x_n) une suite de Cauchy d'éléments de X telle que (x_n) admette une sous-suite (x_{n_k}) convergente. Alors (x_n) converge.*

Démonstration. Soit (x_n) une suite de Cauchy d'éléments de X, et $x \in X$ tel qu'il existe une sous-suite (x_{n_k}) qui converge vers x. Fixons $\varepsilon > 0$. Alors on sait qu'il existe N tel que $d(x_n, x_m) \leq \varepsilon$ pour tout $n, m \geq N$. De plus, il existe aussi K tel que $d(x_{n_k}, x) \leq \varepsilon$ pour tout $k \geq K$. Comme n_k tend vers $+\infty$, on peut trouver k_0 tel qu'on ait à la fois $k_0 \geq K$ et $n_{k_0} \geq N$. Par conséquent, pour tout $n \geq N$ on a à la fois $d(x_n, x_{n_{k_0}}) \leq \varepsilon$ et $d(x_{n_{k_0}}, x) \leq \varepsilon$, donc aussi $d(x_n, x) \leq 2\varepsilon$. Ceci prouve que (x_n) converge vers x. $\square$

Proposition 9.21. *Tout espace métrique compact est complet.*

Démonstration. Soit (X,d) un espace métrique compact, et (x_n) une suite de Cauchy d'éléments de X. Par définition de la compacité, on peut extraire une sous-suite de (x_n) qui converge vers $x \in X$. La proposition précédente nous permet donc de conclure que (x_n) converge vers x. $\square$

Ceci nous fournit nos premiers exemples d'espaces complets. Mais on en connaît beaucoup d'autres !

Théorème 9.22. *Soit $n \in \mathbb{N}^*$ et $\|\cdot\|$ une norme sur $\mathbb{R}^n$. Alors $(\mathbb{R}^n, \|\cdot\|)$ est complet.*

Démonstration. Soit (x_n) une suite de Cauchy d'élements de $(\mathbb{R}^n, \|\cdot\|)$. Alors (x_n) doit être bornée, et le théorème de Bolzano–Weierstrass nous permet donc d'en extraire une sous-suite convergente. La proposition 9.20 nous permet de conclure que (x_n) converge. $\square$

Exercice 9.23. Soit $(E, \|\cdot\|)$ un $\mathbb{R}$-espace vectoriel, et $F \subseteq E$ un sous-espace de dimension finie. Montrer que F est fermé dans E.

Théorème 9.24 (Théorème de Riesz–Fisher)**.** *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, et $p \in [1, +\infty]$. Alors l'espace $L^p(X)$ est un espace complet.*

On admettra ce résultat dont la preuve est difficile !

Définition 9.25. On dit qu'un espace vectoriel E muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ est un *espace de Hilbert* si E, muni de la norme induite par son produit scalaire, est un espace métrique complet.

Un exemple fondamental d'espace de Hilbert : l'espace $L^2(X)$ pour tout espace mesuré $(X, \mathcal{A}, \mu)$.

9.3 Projection sur un convexe fermé

Maintenant on suppose que E est un espace de Hilbert.

Définition 9.26. Soit A une partie non vide de E et $x \in E$. On dit que $a_0 \in A$ est une *projection* de x sur A si on a
$$\|x - a_0\| = \inf\{\|x - a\| \colon a \in A\} \ .$$

Remarque 9.27. Il n'existe pas toujours une projection ; et même quand elle existe, elle n'est pas nécessairement unique. Les deux exercices suivants donnent des exemples de ce phénomène.

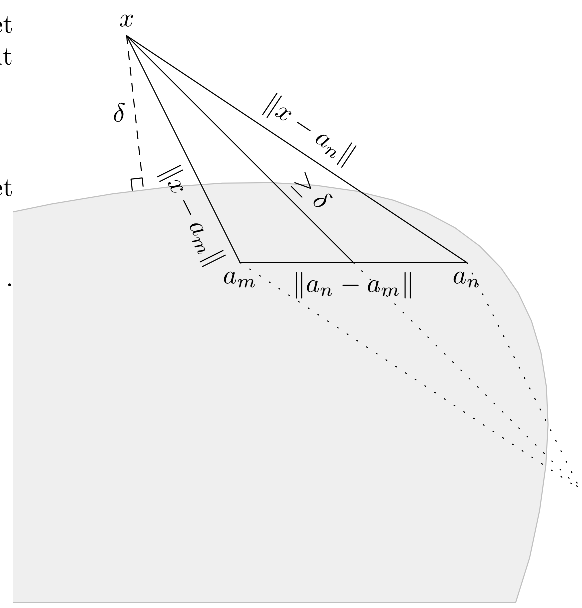
Exercice 9.28. Soit E un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire possédant une famille orthonormale $(e_n)_{n>0}$. On considère $F = \left\{(1 + \frac{1}{n})e_n \colon n > 0\right\}$.

1. Montrer que 0_E n'a pas de projection sur F.
2. Montrer que pour tout $n \neq m$, $\|(1 + \frac{1}{n})e_n - (1 + \frac{1}{m})e_m\| \geq 1$. En déduire que F est fermé.
3. Expliciter un tel exemple pour l'espace d'Hilbert $L^2([0,1])$.

Exercice 9.29. On considère un triangle équilatéral dans $\mathbb{R}^2$. Montrer que son centre de gravité a trois projections.

Théorème 9.30. *Soit E un espace de Hilbert, $A \subseteq E$ un sous-ensemble convexe et fermé. Alors tout point $x \in E$ admet une unique projection $p_A(x)$ sur A.*

Démonstration. Fixons $x \in E$, et notons $\delta = \inf_{a \in A} \|x - a\|$. On peut trouver une suite (a_n) d'éléments de A tels que $\|x - a_n\|$ converge vers δ. On va prouver que (a_n) est une suite de Cauchy ; alors on pourra conclure (comme E est complet) que (a_n) converge vers a, qui appartient à A puisque A est fermé. Par continuité de la norme on aura alors $\|x - a\| = \lim_{n \to +\infty} \|x - a_n\| = \delta$. Ceci montrera l'existence d'une projection (mais pas encore son unicité).



Pour justifier que (a_n) est de Cauchy, on va utiliser la convexité de A et utiliser le fait que $\frac{a+b}{2} \in A$ pour tout $a, b \in A$. En particulier, pour tout $a, b \in A$ on doit avoir
$$\|x - \frac{a+b}{2}\| \geq \delta \ .$$
Fixons $n, m \in \mathbb{N}$ et appliquons l'identité du parallélogramme à $x - a_n$ et $x - a_m$:
$$\|(x - a_n) + (x - a_m)\|^2 + \|(x - a_n) - (x - a_m)\|^2 = 2\|x - a_n\|^2 + 2\|x - a_m\|^2 \ .$$
On en déduit
$$\begin{aligned} \|a_n - a_m\|^2 &= 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - \|2x - a_n - a_m)\|^2 \\ &= 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - 4\|x - \frac{a_n + a_m}{2}\|^2 \\ &\leq 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - 4\delta^2 \end{aligned}$$

Fixons $\varepsilon > 0$. On sait qu'il existe N tel que pour tout $n, m \geq N$ on a $\|x - a_n\|^2 \leq \delta^2 + \varepsilon$ et $\|x - a_m\|^2 \leq \delta^2 + \varepsilon$, et donc aussi $\|a_n - a_m\|^2 \leq 4\varepsilon$. On vient de montrer comme promis que (a_n) est de Cauchy.

Pour montrer que le projeté est unique, considérons $a, b \in A$ tels que $\|x - a\| = \|x - b\| = \delta$. Le même calcul que ci-dessus nous donne
$$\|a - b\|^2 \leq 2\|x - a\|^2 + 2\|x - b\|^2 - 4\delta^2 \leq 0 \ .$$
Ceci n'est possible que si $a = b$. $\square$

Théorème 9.31. *Soit E un espace de Hilbert, A un convexe fermé de E et $x \in E$. Le projeté $p_A(x)$ de x sur A est l'unique $a \in A$ tel que*

$$\forall b \in A \quad \langle x-a, b-a \rangle \leq 0 \ .$$

On en déduit que l'application $x \mapsto p_A(x)$ est 1-lipschitzienne.

Démonstration. Notons $a = p_A(x)$. Pour tout $b \in A$, on a $\|x-a\|^2 \leq \|x-b\|^2$ par définition du projeté. En utilisant la relation

$$\begin{aligned}
\|x-b\|^2 &= \|(x-a)+(a-b)\|^2 \\
&= \|x-a\|^2 + \|a-b\|^2 + 2\langle x-a, a-b\rangle
\end{aligned}$$

On en déduit que pour tout $b \in A$

$$\begin{aligned}
2\langle x-a, b-a\rangle &= \|x-a\|^2 - \|x-b\|^2 + \|a-b\|^2 \\
&\leq \|a-b\|^2 \ .
\end{aligned}$$

Fixons maintenant $b \in A$, différent de a ; pour tout $t \in [0,1]$ on peut appliquer l'inégalité précédente à $b_t = tb + (1-t)a$, qui appartient à A, et obtenir en utilisant le fait que $b_t - a = t(b-a)$

$$2\langle x-a, t(b-a)\rangle \leq t^2\|b-a\| \ .$$

En divisant des deux côtés par t et en faisant tendre t vers 0, on voit que ceci n'est possible que si $\langle x-a, b-a\rangle \leq 0$, ce qu'on voulait démontrer.

Supposons maintenant qu'on a $a' \in A$ tel que $\langle x-a', b-a'\rangle \leq 0$ pour tout $b \in A$. Avec un calcul similaire au précédent, on a

$$\|x-a'\|^2 = \|x-a\|^2 + 2\langle x-a', a-a'\rangle - \|a-a'\|^2 \ .$$

Puisque $\langle x-a', a-a'\rangle \leq 0$ par hypothèse sur a', on a donc que $\|x-a'\|^2 \leq \|x-a\|^2$, ce qui n'est possible que si $a = a'$ par définition d'un projeté.

Enfin, il nous reste à démontrer que $x \mapsto p_A(x)$ est 1-lipschitzienne. Soit $x, x' \in X$. D'après la propriété qu'on vient d'établir, on a à la fois

$$\langle x - p_A(x), p_A(x') - p_A(x)\rangle \leq 0 \quad \text{ et } \quad \langle x' - p_A(x'), p_A(x) - p_A(x')\rangle \leq 0 \ .$$

En notant $r = x - x' + p_A(x') - p_A(x)$ on obtient en sommant les deux inégalités précédentes que

$$\langle p_A(x) - p_A(x'), x' - x + p_A(x) - p_A(x')\rangle \leq 0 \ ,$$

autrement dit $\langle p_A(x) - p_A(x'), r\rangle \geq 0$. Finalement,

$$\begin{aligned}
\|x-x'\|^2 &= \|p_A(x) - p_A(x') + r\|^2 \\
&= \|p_A(x) - p_A(x')\|^2 + \|r\|^2 + 2\langle p_A(x) - p_A(x'), r\rangle \\
&\geq \|p_A(x) - p_A(x')\|^2 \ .
\end{aligned}$$

On a bien montré que p_A est 1-lipschitzienne. □

Remarque 9.32. Reprenons les notations du théorème précédent, et supposons de plus que $x \notin A$. L'ensemble $H = \{z \in E \colon \langle z-a, x-a\rangle = 0\}$ est un sous-espace affine de E (égal à $a + (\mathrm{Vect}(x-a))^{\perp}$), qui admet un supplémentaire de dimension 1 (la droite de vecteur directeur $x-a$) ; on dit que H est un *hyperplan*. L'hyperplan H "coupe" E en deux demi-espaces : l'ensemble H^-des z tels que $\langle z-a, x-a\rangle \leq 0$ et l'ensemble H^+ des z tels que $\langle z-a, x-a\rangle > 0$. Le théorème ci-dessus nous dit que A est entièrement contenu dans H^-, tandis que x appartient à H^+. On dit alors que H *sépare* x et A. L'existence d'un hyperplan séparant un point d'un convexe fermé est une conséquence importante du théorème de projection. Par exemple, elle nous permet de voir que tout convexe fermé dans un espace de Hilbert est une intersection de demi-espaces fermés.

Un cas particulier très important est celui de la projection sur un sous-espace vectoriel fermé de E (ce qui est le cas, rappelons-le, de tout sous-espace vectoriel de dimension finie).

Théorème 9.33. *Soit E un espace de Hilbert et $F \subseteq E$ un sous-espace vectoriel fermé. Alors, pour tout $x \in E$, le projeté $p_F(x)$ est l'unique élément de F tel que $x - p_F(x)$ soit orthogonal à F. On le nomme* projeté orthogonal *de x sur F.*

La projection orthogonale sur F est une application linéaire continue.

Démonstration. Soit $x \in E$. Par définition, $p_F(x)$ appartient à F, et on a vu que $p_F(x)$ est l'unique élément de F tel que $\langle x - p_F(x), f - p_F(x)\rangle \leq 0$ pour tout $f \in F$. Mais comme $f + p_F(x) \in F$, on en déduit que $\langle x - p_F(x), f\rangle = 0$ pour tout $f \in F$. Réciproquement, si $a \in F$ est tel que $\langle x - a, f\rangle = 0$ pour tout $f \in F$, alors on a en particulier $\langle x - a, a\rangle = 0$ et donc $\langle x - a, f - a\rangle = 0$ pour tout $f \in F$, et d'après le théorème précédent cette propriété impose que $a = p_F(x)$.

Reste à montrer que p_F est linéaire (on sait déjà qu'elle est continue puisqu'elle est 1-lipschitzienne). Considérons $x, y \in E$ et $\lambda \in \mathbb{R}$. Alors pour tout $f \in F$ on a

$$\begin{aligned}\langle x + \lambda y - p_F(x) - \lambda p_F(y), f\rangle &= \langle x - p_F(x), f\rangle + \langle \lambda y - \lambda p_F(y), f\rangle \\ &= \langle x - p_F(x), f\rangle + \lambda\langle y - p_F(y), f\rangle \\ &= 0\end{aligned}$$

Par conséquent, $x+\lambda y-(p_F(x)+\lambda p_F(y))$ est orthogonal à tous les éléments de F, ce qui montre que $p_F(x+\lambda y) = p_F(x) + \lambda p_F(y)$. □

Corollaire 9.34. Soit E un espace de Hilbert, et F un sous-espace vectoriel fermé. Alors on a $E = F \oplus F^\perp$. De plus, $F^\perp$ est l'unique supplémentaire de F orthogonal à F : si $E = F \oplus G$ et G est orthogonal à F alors $G = F^\perp$.

Démonstration. On sait déjà que $F^\perp$ est un sous-espace vectoriel (fermé) de E ; de plus, si $x \in F \cap F^\perp$ alors $\|x\|^2 = \langle x, x\rangle = 0$. Enfin, pour tout $x \in E$ on a $x = (x - p_F(x)) + p_F(x)$, ce qui prouve que $x \in F^\perp + F$.

Soit G un supplémentaire de F orthogonal à F, c'est-à-dire $E = F \oplus G$ et $G \subset F^\perp$. Prenons $x \in F^\perp$. Alors, $x = f + g$ avec $f \in F$ et $g \in G$, d'où $f = x - g \in F \cap F^\perp = \{0_E\}$ et donc $x \in G$. □

Exercice 9.35. Montrer qu'un sous-espace vectoriel F de E est dense dans E si, et seulement si, $F^\perp = \{0_E\}$.

Exercice 9.36. Soit $(a_i)_{i\in I}$ une famille de vecteurs de E. Montrer que le sous-espace vectoriel engendré par les $(a_i)_{i\in I}$ est dense dans E si, et seulement si,

$$\forall x \in E \quad (\forall i \in I\ \langle x, a_i\rangle = 0) \Leftrightarrow x = 0_E\ .$$

Exercice 9.37. Soit F un sous-espace vectoriel de E. Montrer que $(F^\perp)^\perp = \overline{F}$.

Définition 9.38. Soit $(a_i)_{i\in I}$ une famille de vecteurs de E. On dit que $(a_i)_{i\in I}$ est une *base hilbertienne* de vecteurs de E si, et seulement si, la famille $(a_i)_{i\in I}$ est orthonormale et engendre un sous-espace vectoriel dense.

On rappelle que la famille engendre un sous-espace vectoriel dense si, et seulement si le seul vecteur qui est orthogonal à chacun des a_i est le vecteur nul. Attention, si E est de dimension infinie, une base hilbertienne de E n'est pas nécessairement une base algébrique ! Par contre, si E est de dimension finie, c'est-à-dire si E est un espace euclidien, alors toute base hilbertienne est une base de E, c'est-à-dire engendre E. Autrement dit, en dimension finie, aucun sous-espace vectoriel propre n'est dense dans E (ceci suit du fait qu'en dimension finie, tout sous-espace vectoriel est fermé).

9.4 Procédé de Gram-Schmidt, inégalité de Bessel, identités de Parseval

Notons tout d'abord que la projection d'un point sur un sous-espace vectoriel de dimension finie se calcule facilement à l'aide d'une base orthonormale de F (preuve laissée en exercice) :

Proposition 9.39. *Soit E un espace de Hilbert et F un sous-espace vectoriel de dimension finie avec $(e_1, \ldots, e_n)$ une base orthonormale de F. Alors pour tout $x \in E$, on a*

$$p_F(x) = \sum_{i=1}^{n} \langle x, e_i\rangle e_i.$$

Rappelons que le procédé de Gram-Schmidt permet de calculer une base orthonormale d'un espace euclidien à partir d'une base donnée :

Proposition 9.40 (Procédé de Gram-Schmidt)**.** *Soit E un espace euclidien et $(e_1, \ldots, e_n)$ une base de E. Pour chaque $0 < i < n$, notons F_i le sous-espace vectoriel* $\mathrm{Vec}(e_1, \ldots, e_i)$ *engendré par $e_1, \ldots, e_i$. Alors, la famille $(e'_1, \ldots, e'_n)$ définie de la manière suivante est une base orthonormale de E :*

$$\begin{aligned} e'_1 &= \frac{e'_1}{\|e'_1\|} \\ e'_i &= \frac{e_i - p_{F_{i-1}}(x)}{\|e_i - p_{F_{i-1}}(x)\|} = \frac{e_i - \sum_{k=1}^{i-1} \langle e_i, e'_k \rangle e'_k}{\|e_i - \sum_{k=1}^{i-1} \langle e_i, e'_k \rangle e'_k\|} \quad \textit{pour } 1 < i \leq n \end{aligned}$$

Exercice 9.41. Vérifier que les vecteurs $e_1 = (1, 1, 1)$, $e_2 = (1, 1, -1)$ et $e_3 = (0, 1, 2)$ forment une base de $\mathbb{R}^3$. Utiliser le procédé de Gram-Schmidt sur cette base pour obtenir une base orthonormale.

Maintenant, étant donnée une base orthonormale d'un espace euclidien, on a les identités suivantes (à vérifier en exercice) :

Proposition 9.42. *Soit E un espace euclidien et $(e_1, \ldots, e_n)$ une base orthonormale de E. Alors pour tous x, y de E, on a :*

$$\begin{aligned} x &= \sum_{i=1}^{n} \langle x, e_i \rangle e_i \\ \|x\|^2 &= \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \\ \langle x, y \rangle &= \sum_{i=1}^{n} \langle x, e_i \rangle \langle y, e_i \rangle \end{aligned}$$

Dans cette partie, nous allons voir que ces identités se généralisent aux espaces de Hilbert ayant une base hilbertienne dénombrable.

Définition 9.43. Un espace métrique est dit *séparable* s'il contient une partie dénombrable dense.

Exercice 9.44. Un théorème d'approximation de Weierstrass montre que toute fonction réelle définie et continue sur $[-1, 1]$ est limite uniforme de fonctions polynomiales à coefficients rationnels. Notons que l'ensemble des fonctions polynomiales à coefficients rationnels est dénombrable. En déduire que l'espace de Hilbert $(C([-1, 1], \mathbb{R}), \| \cdot \|_2)$ est séparable.

Proposition 9.45. *Un espace de Hilbert séparable de dimension infinie possède une base hilbertienne dénombrable $(e_i)_{i \in \mathbb{N}}$.*

Démonstration. La preuve est laissée en exercice. Indication : on considère une énumération $(a_n)_{n \in \mathbb{N}}$ d'une partie dénombrable A dense, puis on en extrait une base du sous-espace vectoriel engendré par A, et enfin, en utilisant le procédé de Gram-Schmidt, on construit une base orthonormale de ce sous-espace vectoriel. □

Exercice 9.46 (Polynômes de Legendre)**.** On se place dans l'espace de Hilbert $(C([-1, 1], \mathbb{R}), \| \cdot \|_2)$.

1. En utilisant l'exercice précédent et le procédé de Gram-Schmidt, montrer qu'il existe une base hilbertienne $(p_n)_{n \in \mathbb{N}}$, telle que pour chaque $n \in \mathbb{N}$, la fonction p_n est une fonction polynomiale de degré n.
2. Pour $n \in \mathbb{N}$, on considère la fonction polynomiale l_n définie par $l_n(t) = \dfrac{1}{2^n n!} \dfrac{d^n}{dx^n}(t^2 - 1)^n$ pour $t \in [-1, 1]$ (polynômes de Legendre).
   (a) Montrer que la famille $(l_n)_{n \in \mathbb{N}}$ est une famille orthogonale.
   (b) Déterminer le degré de l_n pour chaque n.

(c) En déduire que pour chaque n on a $p_n = \dfrac{l_n}{\|l_n\|}$.

Théorème 9.47 (Inégalité de Bessel)**.** *Soit E un espace de Hilbert et $(e_i)_{i\in\mathbb{N}}$ une famille orthonormale. Alors pour tout $x \in E$,*

$$\sum_{i=0}^{+\infty} |\langle x, e_i\rangle|^2 \leq \|x\|^2.$$

Démonstration. Soit $n \geq 0$. La projection de x sur l'espace vectoriel engendré par $e_0, \dots, e_n$ est égale à $\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ et est orthogonale à $x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i$.

Cette orthogonalité entraîne que :

$$\|x\|^2 = \left\|\sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 = \sum_{i=0}^{n} |\langle x, e_i\rangle|^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 \geq \sum_{i=0}^{n} |\langle x, e_i\rangle|^2.$$

On conclut en passant à la limite. □

Théorème 9.48 (Identités de Parseval)**.** *Soit E un espace de Hilbert possédant une base hilbertienne dénombrable $(e_i)_{i\in\mathbb{N}}$. Alors, pour tous x, y de E, on a :*

$$\begin{aligned} x &= \sum_{i=0}^{+\infty}\langle x, e_i\rangle e_i \\ \|x\|^2 &= \sum_{i=0}^{+\infty} |\langle x, e_i\rangle|^2 \\ \langle x, y\rangle &= \sum_{i=0}^{+\infty}\langle x, e_i\rangle\langle y, e_i\rangle \end{aligned}$$

Démonstration. Notons F le sous-espace vectoriel engendré par la famille $(e_i)_{i\in\mathbb{N}}$ et pour chaque n, notons F_n le sous-espace vectoriel engendré par $e_0, \dots, e_n$. Comme F est dense dans E, il existe une suite (x_n) d'éléments de F convergeant vers x. Quitte à ajouter au début de la suite suffisamment de termes nuls et à répéter des termes, on peut supposer que pour chaque n, le terme x_n appartient à F_n. Comme $\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ est la projection de x sur F_n, on obtient ainsi l'inégalité

$$\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\| \leq \|x - x_n\|,$$

ce qui entraîne que $\displaystyle\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ converge vers x.

En passant à la limite dans l'égalité

$$\|x\|^2 = \left\|\sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 = \sum_{i=0}^{n} |\langle x, e_i\rangle|^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2$$

on obtient la seconde identité du théorème.

Notons que par bilinéarité du produit scalaire on a,

$$\left\langle \sum_{i=0}^{n}\langle x, e_i\rangle e_i, \sum_{i=0}^{n}\langle y, e_i\rangle e_i \right\rangle = \sum_{i=0}^{n}\langle x, e_i\rangle\langle y, e_i\rangle.$$

Par continuité du produit scalaire, le terme de gauche converge vers $\langle x, y\rangle$ ce qui donne la dernière identité. □

Remarque 9.49. Le théorème ci-dessus est valide quel que soit l'ordre de l'énumération de la base de Hilbert et donc indépendamment de l'ordre avec lequel l'on somme les séries.

Les identités de Parseval permettent de montrer réciproquement que si un espace de Hilbert a une base hilbertienne dénombrable alors il est séparable :

Exercice 9.50. Soit E un espace de Hilbert possédant une base hilbertienne dénombrable $(e_i)_{i\in\mathbb{N}}$. On considère la partie A des combinaisons linéaires finies de $(e_i)_{i\in\mathbb{N}}$ à coefficients rationnels. Notons que A est dénombrable. Montrer que A est dense dans E

Terminons cette partie en montrant l'unicité du développement en série dans le théorème 9.48. On étudiera un exemple important dans la partie suivante, les séries de Fourier.

Proposition 9.51 (unicité du développement)**.** *Soit $(e_i)_{i\in\mathbb{N}}$ une base hilbertienne d'un espace de Hilbert E séparable. Soit $x \in E$ et une famille $(\lambda_i)_{i\in\mathbb{N}}$ de scalaires dans $\mathbb{R}$ tels que la série $\sum_{i=0}^{+\infty} \lambda_i e_i$ converge vers x. Alors pour tout $i \in \mathbb{N}$, on a $\lambda_i = \langle x, e_i\rangle$. On appellera ces scalaires, les coefficients de x dans la base $(e_i)_{i\in\mathbb{N}}$.*

Démonstration. Pour chaque n, posons $x_n = \sum_{i=0}^{n} \lambda_i e_i$. Par hypothèse, on a (x_n) qui converge vers x. Notons que pour $i \leq n$, on a $\langle x_n, e_i\rangle = \lambda_i$ et par continuité du produit scalaire, en faisant tendre n vers l'infini, on obtient $\langle x, e_i\rangle = \lambda_i$.

□

9.5 Une application : les séries de Fourier

On munit l'espace $L^2([-\pi,\pi])$ du produit scalaire défini par

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{[-\pi,\pi]} f g d\lambda \ .$$

Pour nous simplifier un peu les notations ci-dessous, notons $s_n(t) = \sin(nt)$ pour $n \geq 1$, et $c_n(t) = \cos(nt)$ pour $n \geq 0$ (en particulier c_0 est la fonction constante égale à 1). On a déjà vu que les fonctions (s_n) forment une famille orthogonale ; de plus le changement de variable $u = \frac{\pi}{2} - x$ (ou un dessin des graphes...) permet de voir que pour $n \geq 1$ on a $\|s_n\| = \|c_n\|$ (la norme étant celle associée à notre produit scalaire : la constante multiplicative est là pour que $\|1\| = 1$...), et comme $s_n^2 + c_n^2 = 1$, on déduit que pour $n \geq 1$ on a $\displaystyle\int_{-\pi}^{\pi} s_n^2 d\mu = \pi$, et donc $\|s_n\| = \frac{\sqrt{2}}{2}$, et de même pour c_n.

Le même type de calculs que ceux effectués pour la famille (s_n) permet de voir que la famille (c_m) est orthogonale, et aussi que pour tout n, m on a $\langle s_n, c_m\rangle = 0$. A titre d'exemple, montrons cette dernière égalité : pour $n \geq 1$ et $m \geq 0$ on a

$$\sin(nt)\cos(mt) = \frac{\sin((n+m)t) + \sin((n-m)t)}{2} \ .$$

De plus, $\displaystyle\int_{-\pi}^{\pi} \sin(kt) = 0$ pour tout k, ce qui prouve que $\displaystyle\int_{-\pi}^{\pi} \sin(nt)\cos(mt)dt = 0$ pour tout n, m.

Il ressort des calculs précédents que la famille de fonctions $((s_n)_{n\geq 1}, (c_m)_{m\geq 0}$ est une famille orthogonale ; cette famille n'est pas orthonormale mais le serait si pour $n \geq 1$ on remplaçait (s_n) par $\sqrt{2}s_n$ et c_n par $\sqrt{2}c_n$. L'espace vectoriel engendré par cette famille est appelé l'ensemble des *polynômes trigonométriques*. Les formules de linéarisation nous permettent de vérifier qu'un produit de polynômes trigonométriques est encore un polynôme trigonométrique.

Théorème 9.52. *L'espace des polynômes trigonométriques est dense dans $L^2([-\pi,\pi])$.*

Démonstration. On doit montrer que, si $f \in L^2([-\pi,\pi]$ est telle que $\langle f, c_n\rangle = 0$ pour tout $n \geq 0$ et $\langle f, s_n\rangle = 0$ pour tout $n \geq 1$ alors f est la fonction nulle.

Commençons par supposer que f est une fonction continue et non nulle ; on souhaite montrer que f n'est pas orthogonale à tous les polynômes trigonométriques. Pour simplifier les notations, on va se contenter de traiter le cas où il existe $h > 0$ tel que $f(x) > 0$ pour tout $x \in]-h, h[$ (la preuve ci-dessous s'adapte facilement au cas général). Considérons alors le polynôme trigonométrique

$$P_n \colon x \mapsto (1 + \cos(x) - \cos(h))^n \ .$$

Pour $x \in]-h, h[$ on a que $P_n(x)$ croît (à x fixé) vers $+\infty$ quand $n \to +\infty$, tandis que pour $x \in [-\pi, -h[\cup]h, \pi[$ on a $P_n(x) \to 0$ et $|P_n(x)| \leq 1$. Mais alors :

– Sur $]-h,h[$, on peut appliquer le théorème de convergence monotone pour conclure que $\displaystyle\int_{]-h,h[} fP_n d\lambda$ tend vers $+\infty$.
– Sur $]-\pi,-h[\cup]h,\pi[$, on peut utiliser le fait que $|f(x)P_n(x)| \leq |f(x)| \leq \|f\|_\infty$ pour conclure, à l'aide du théorème de convergence dominée, que $\displaystyle\int_{]-\pi,-h[\cup]h,\pi[} fP_n d\lambda$ tend vers 0.

Tout ceci nous dit que $\displaystyle\int_{]-\pi,\pi]} fP_n d\lambda$ tend vers $+\infty$; en particulier on ne peut avoir $\langle f, P_n\rangle = 0$ pour tout n, par conséquent f n'est pas orthogonale à l'espace des polynômes trigonométriques.

Si maintenant on ne suppose plus f continue, mais seulement L^2, et que f est orthogonale aux polynômes trigonométriques, remarquons d'abord que $x \mapsto F(x) = \int_{-\pi}^{x} f d\lambda$ est bien définie (d'après Hölder, f est intégrable puisque la fonction constante 1 appartient à $L^2([\pi,\pi])$ et continue : pour tout $x, y \in [-\pi,\pi]$ on a

$$|F(x)-F(y)| = \left|\int_y^x f d\lambda\right| \leq \sqrt{\int_y^x f^2 d\lambda}\sqrt{\int_y^x 1 d\lambda} \leq \|f\|_2\sqrt{y-x}\ .$$

Soit maintenant P un polynôme trigonométrique, et Q une primitive de P (qui est de nouveau un polynôme trigonométrique)

$$\begin{aligned}
\langle F, P\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)P(x)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{x} f(t)dt\right)P(x)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{t}^{\pi} P(x)dx\right)f(t)dt \text{ (par Fubini)} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}(Q(\pi)-Q(t))f(t)dt \\
&= \langle Q(\pi)-Q, f\rangle \\
&= 0\ .
\end{aligned}$$

Ci-dessus, l'application de Fubini est justifiée par le fait qu'on intègre une fonction continue sur un compact (l'ensemble $\{(x,t) \in [-\pi,\pi]^2 \colon t \leq x\}$) ; et la dernière égalité vient du fait que f est supposée orthogonale à tous les polynômes trigonométriques. Comme F est continue, la première partie de notre raisonnement nous permet de conclure que $F = 0$. On vient donc de conclure que, pour toute fonction $f \in L^2([-\pi,\pi])$ qui est orthogonale aux polynômes trigonométriques, on a pour tout $x, y \in [-\pi,\pi]$ que $\displaystyle\int_{[x,y]} f d\lambda = 0$. Ceci entraîne que $f = 0$ presque partout [i], ce qu'on souhaitait démontrer.

□

Une autre façon d'énoncer le théorème précédent est de dire que la famille $(c_0, (\sqrt{2}s_n)_{n\geq 1}, (\sqrt{2}c_n)_{n\geq 1})$ est une base hilbertienne de $L^2([-\pi,\pi])$. Pour $f \in L^2([-\pi,\pi])$, introduisons ses *coefficients de Fourier* :

$$a_0(f) = \langle f, c_0\rangle = \frac{1}{2\pi}\int_{[-\pi,\pi]} f d\lambda$$
$$\forall n \geq 1\ a_n(f) = 2\langle f, c_n\rangle = \frac{1}{\pi}\int_{[-\pi,\pi]} f(t)\cos(nt)d\lambda(t)$$
$$\forall n \geq 1\ b_n(f) = 2\langle f, s_n\rangle = \frac{1}{\pi}\int_{[-\pi,\pi]} f(t)\sin(nt)d\lambda(t)$$

i. Ici, on triche un peu : la preuve de ce fait n'est pas immédiate. On peut démontrer que deux mesures μ, ν sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telles que pour tout a, b on ait $\mu([a,b]) = \nu([a,b]) < +\infty$ doivent être égales, ce qui permet de conclure ici, et aussi de montrer l'unicité de la mesure de Lebesgue sur $\mathbb{R}$.

Puisque la famille $(c_0, (\sqrt{2}s_n)_{n\geq 1}, (\sqrt{2}c_n)_{n\geq 1})$ est une base hilbertienne, on sait que pour tout $f \in [-\pi, \pi]$ on peut écrire :

$$f = \langle f, c_0\rangle + \sum_{n\geq 1} \langle f, \sqrt{2}c_n\rangle \sqrt{2}c_n + \langle f, \sqrt{2}s_n\rangle \sqrt{2}s_n$$

$$\|f\|^2 = \langle f, c_0\rangle^2 + \sum_{n\geq 1} (\langle f, \sqrt{2}c_n\rangle)^2 + (\langle f, \sqrt{2}s_n\rangle)^2 \ .$$

En termes de coefficients de Fourier, on a donc, pour toute fonction $f \in L^2([-\pi,\pi])$ et presque tout $t \in [-\pi, \pi]$:

1. $f(t) = a_0(f) + \sum_{n\geq 1} a_n(f)\cos(nt) + b_n(f)\sin(nt)$;
2. $\frac{1}{2\pi}\int_{-\pi}^{\pi} f^2 d\lambda = a_0(f)^2 + \frac{1}{2}\sum_{n\geq 1}(a_n(f))^2 + (b_n(f))^2.$

La série apparaissant dans la première égalité ci-dessus est appelée *série de Fourier* de f. Remarquons que, même si f est continue, on n'a justifié la première égalité ci-dessus que presque partout (c'est une égalité au sens des fonctions L^2) : il peut a priori y avoir des points où on n'a pas égalité entre la fonction et sa série de Fourier ; effectivement, on a besoin d'hypothèses supplémentaires (par exemple, si f est C^1 par morceaux) pour conclure que la série de Fourier de f converge en tout point vers f.

La deuxième égalité ci-dessus est souvent appelée égalité de Parseval (et c'est bien une conséquence de ce que nous avons appelé égalité de Parseval plus tôt dans ce cours).

Exercice 9.53. Calculer les coefficients de Fourier de la fonction f définie sur $[-\pi, \pi]$ par $f(x) = x$. A l'aide de l'égalité de Parseval, en déduire l'égalité :

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \ .$$

9.6 Une autre application : le théorème du minimax

On va conclure ces notes en discutant un théorème qu'on peut montrer à l'aide du théorème de projection sur un convexe fermé (ou, dans la preuve qu'on va présenter, son corollaire sur la séparation d'un fermé convexe et d'un point) : le théorème du minimax de von Neumann.

Il s'agit d'un théorème de *théorie des jeux*, qu'on va ici essayer de présenter dans un contexte assez simple. Imaginons deux joueurs X et Y jouant à un jeu J ; X a un nombre fini de choix possibles $A = \{a_1, \ldots, a_n\}$ et Y un nombre fini de choix possibles $B = \{b_1, \ldots, b_m\}$. On a également une fonction de gain $g \colon A \times B \to \mathbb{R}$; si X joue a_i et Y joue b_j, alors X gagne $g(a_i, b_j)$ et Y perd $g(a_i, b_j)$ (on dit que le jeu est à somme nulle : ce que X gagne est perdu par Y). Attention : g peut prendre des valeurs négatives (et on considère que gagner -10 euros revient à perdre 10 euros...).

Par exemple, X et Y pourraient jouer à pierre-papier-ciseaux, ou à un jeu un peu plus subtil appelé *pair ou impair* : X et Y jouent chacun un entier entre 1 et 2 ; si la somme des deux est impaire X gagne un montant égal à la somme des deux chiffres joués, et si la somme est paire c'est Y qui gagne. Comment doivent jouer X et Y s'ils sont rationnels ?

Comme ni X ni Y n'a d'information sur ce que l'autre va faire, on va considérer qu'une stratégie pour un des deux joueurs consiste à jouer au hasard, suivant une loi de probabilité p qu'il s'est fixé à l'avance : X va jouer a_i avec la probabilité p_i, alors que Y va jouer b_j avec la probabilité q_j. Lorsque X suit p, et Y suit q, l'espérance de gain pour X (et donc, l'espérance de perte pour Y) est

$$g(p, q) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_i g(a_i, b_j) q_j \ .$$

Si X se fixe une stratégie p, le pire qui peut lui arriver est de gagner $\inf_q g(p, q)$ (si q choisit une stratégie aussi bien adaptée que possible pour contrer p) ; notons que cet inf est en fait un min, puisque $q \mapsto g(p, q)$ est une fonction continue sur le compact

$$\{(q_1, \ldots, q_m) \colon \forall i \ q_j \geq 0 \text{ et } \sum q_j = 1\} \ .$$

Ceci signifie que, au minimum, X peut trouver une stratégie lui garantissant un gain [ii] supérieur ou égal à

$$\underline{V}(J) = \max_p \min_q g(p,q) \ .$$

En suivant le même raisonnement en échangeant les rôles de X et Y, on voit que Y peut se garantir une perte inférieure ou égale à

$$\overline{V}(J) = \min_q \max_p g(p,q) \ .$$

La définition de $\underline{V}(J)$ et $\overline{V}(J)$ devrait rendre claire l'inégalité $\underline{V}(J) \le \overline{V}(J)$.

Théorème 9.54 (Théorème du minimax de von Neumann)**.** *Pour tout jeu J à somme nulle, on a $\underline{V}(J) = \overline{V}(J)$. Cette valeur commune est appelée la valeur du jeu J et notée $V(J)$.*

Avant de donner une preuve de ce résultat, observons une conséquence : on peut trouver des stratégies p_0, q_0 telles que $g(p_0,q_0) = V(J)$. Le fait que $g(p_0,q_0) = \underline{V}(J)$ nous dit que, pour toute stratégie q de Y, on a $g(p_0,q) \ge g(p_0,q_0) = \min g(p_0,q)$: autrement dit, si X joue p_0, toute autre stratégie que q_0 est moins bonne (au sens large) pour Y que q_0. Comme jouer q_0 garantit à Y sa perte minimale, Y doit jouer cette stratégie (ou une autre stratégie telle que $g(p_0,q) = V(J)$). De même, X doit jouer p_0 (ou une autre stratégie lui garantissant le même gain) s'il est rationnel. On voit ainsi apparaître un équilibre : si les deux joueurs sont rationnels, ils vont jouer des stratégies p, q telles que $g(p,q) = V(J)$.

Preuve du théorème du minimax. Une petite observation préliminaire :

$$\overline{V}(J) = \min_q \max_i g(a_i,q) \ .$$

En effet, X peut choisir la stratégie de toujours jouer a_i ; ceci montre que pour tout q, on a $\max_p g(p,q) \ge \max_i g(a_i,q)$; Réciproquement, si on choisit i_0 tel que $g(a_i,q) \le g(a_{i_0},q)$ pour tout i, alors pour toute stratégie p on a

$$g(p,q) = \sum_{i=1}^{n} p_i g(a_i,q) \le \sum_{i=1}^{n} p_i g(a_{i_0},q) = g(a_{i_0},q) \ .$$

Ceci nous donne l'inégalité $\max_p g(p,q) \le \max_i g(a_i,q)$, et donc l'égalité entre ces deux quantités, qui nous sera utile dans la suite de la preuve.

On va raisonner par l'absurde et supposer $\underline{V}(J) < \overline{V}(J)$; on peut alors trouver c entre ces deux valeurs et, quitte à remplacer la fonction g par $g - c$, se ramener au cas où $\underline{V}(J) < 0 < \overline{V}(J)$, dans lequel on se place dans la suite.

Introduisons

$$E = \{(e_1,\ldots,e_n) \in \mathbb{R}^n \colon \exists q \ \forall i \ e_i \ge g(a_i,q)\} \ .$$

Notons d'abord que E est convexe : si $(e_1,\ldots,e_n)$ et $(f_1,\ldots,f_n)$ appartiennent à E, et $t \in [0,1]$, on a des stratégies q_e, q_f pour Y telles que $e_i \ge g(a_i,q_e)$ et $f_i \ge g(a_i,q_f)$ pour tout i. Mais alors $q' = tq_e + (1-t)q_f$ est une stratégie pour Y, et on a bien pour tout i

$$te_i + (1-t)f_i \ge tg(a_i,q_e) + (1-t)g(a_i,q_f) = g(a_i, tq_e + (1-t)q_f) = g(a_i,q') \ .$$

Ensuite, montrons que E est fermé : soit (x_k) une suite d'éléments de E qui converge vers $x \in \mathbb{R}^n$; on veut prouver que $x \in E$. Pour tout k on peut trouver une stratégie q_k pour Y telle que $g(a_i,q_k) \le x_k(i)$. Comme on l'a déjà dit précédemment, l'ensemble des stratégies pour Y est $\{(q_1,\ldots,q_m) \colon \forall j \ q_j \ge 0 \text{ et } \sum q_j = 1\}$, qui est fermé et borné dans $\mathbb{R}^m$, et donc compact. Par conséquent, quitte à extraire on peut supposer que (q_k) converge vers une stratégie q. Mais alors, par continuité de g, on a pour tout i que $g(a_i,q_k)$ tend vers $g(a_i,q)$, et par définition de la convergence dans $\mathbb{R}^k$ la suite $x_k(i)$ converge vers x_i. Comme pour tout k on a $g(a_i,q_k) \le x_k(i)$ on conclut par passage à la limite que $g(a_i,q) \le x_i$, ce qui montre comme espéré que $x \in E$.

Ensuite, notons que $0 \notin E$: en effet, par l'observation faite au début de la preuve, pour toute stratégie q il existe i tel que $g(a_i,q) \ge \overline{V}(J) > 0$.

D'après la remarque 9.32, on peut trouver $u = (u_1,\ldots,u_n) \in \mathbb{R}^n$ tel que $\langle u, e\rangle > 0$ pour tout $e \in E$. En particulier, pour toute stratégie q et tout uplet $(r_1,\ldots,r_n)$ de réels positifs, on a $\sum_{i=1}^n u_i(g(a_i,q)+r_i) > 0$. Ceci n'est possible que si $u_i \ge 0$ pour tout i (sans quoi, faire tendre r_i vers $+\infty$ sans toucher aux autres coordonnées

ii. notons ici encore que ce "gain" peut-être négatif !

et regarder la limite...). De plus, on ne peut avoir $u_i = 0$ pour tout i ; par conséquent, on peut considérer la stratégie p où X joue a_i avec probabilité

$$p_i = \frac{u_i}{\sum_{i=1}^n u_i} \ .$$

Alors, le fait que $\langle u, e \rangle > 0$ pour tout $e \in E$ nous permet de conclure que pour toute stratégie q on a

$$\sum_{i=1}^n p_i g(a_i, q) = \frac{1}{\sum_{i=1}^n u_i} \sum_{i=1}^n u_i g(a_i, q) > 0 \ .$$

Autrement dit, la stratégie p que nous venons de produire est telle que $g(p, q) > 0$ pour toute stratégie q, ce qui démontre que $\underline{V}(J) > 0$, contredisant notre hypothèse de départ. □