Espaces préhilbertiens

Plan du chapitre

I - Produit scalaire . . . . **page 2**
1) Définition d'un produit scalaire . . . . page 2
2) Espaces préhilbertiens de référence . . . . page 3
3) Inégalité de Cauchy-Schwarz . . . . page 5
4) Norme hilbertienne . . . . page 5
4-a) Définition . . . . page 5
4-b) Carré scalaire . . . . page 6
4-c) Théorème de Pythagore . . . . page 6
4-d) Identité du parallélogramme . . . . page 7
4-e) Identités de polarisation . . . . page 7
II - Orthogonalité . . . . **page 7**
1) Vecteurs orthogonaux . . . . page 7
2) Orthogonal d'une partie de E . . . . page 7
3) Familles orthogonales, familles orthonormales . . . . page 10
3-a) Définitions . . . . page 10
3-b) Propriétés . . . . page 10
III - Le théorème de la projection orthogonale . . . . **page 12**
1) Le théorème de la projection orthogonale . . . . page 12
2) Familles totales . . . . page 16

I - Produit scalaire

1) Définition d'un produit scalaire

DÉFINITION 1. Soit E un $\mathbb{R}$-espace vectoriel. Soit φ une application de $E \times E$ dans $\mathbb{R}$.

φ est un produit scalaire sur E si et seulement si φ est une forme **bilinéaire**, **symétrique**, **définie**, **positive** sur E c'est-à-dire

Bilinéarité : $\forall (x_1, x_2, y) \in E^3$, $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2$, $\varphi(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \varphi(x_1, y) + \lambda_2 \varphi(x_2, y)$ (linéarité par rapport à la 1ère variable) et
$\forall (x, y_1, y_2) \in E^3$, $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2$, $\varphi(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 \varphi(x, y_1) + \lambda_2 \varphi(x, y_2)$ (linéarité par rapport à la 2ème variable).
Symétrie : $\forall (x, y) \in E^2$, $\varphi(x, y) = \varphi(y, x)$.
Positivité : $\forall x \in E$, $\varphi(x, x) \geqslant 0$.
Définition : $\forall x \in E$, $\varphi(x, x) = 0 \Rightarrow x = 0$.

Notation. Si φ est un produit scalaire sur E, $\varphi(x, y)$ peut se noter $x.y$. C'est cette notation qui est adoptée quand on fait de la géométrie en dimension 2 ou 3 avec le produit scalaire usuel. Quand E est un espace de suites, de polynômes, de fonctions ..., la notation avec un point devient trop ambigüe ($u.v$, $P.Q$, $f.g$). On préfèrera alors la notation $(x|y)$ ou $\langle x, y \rangle$. La notation $<\ ,\ >$ peut s'avérer pénible si on est amené à écrire des inégalités : $\forall x \in E \setminus \{0\}$, $< x, x >> 0$.

Commentaire 1. Dans la pratique, pour vérifier que φ est un produit scalaire, on commence par vérifier la symétrie puis on vérifie la linéarité par rapport à la première variable. La linéarité par rapport à la deuxième variable s'en déduit par symétrie

$$\varphi(x, \lambda_1 x y_1 + \lambda_2 y_2) = \varphi(\lambda_1 x y_1 + \lambda_2 y_2, x) = \lambda_1 \varphi(y_1, x) + \lambda_2 \varphi(y_2, x) = \lambda_1 \varphi(x, y_1) + \lambda_2 \varphi(x, y_2).$$

Commentaire 2. La positivité et la définition peuvent se traiter ensemble :

$$\varphi \text{ est définie, positive} \Leftrightarrow \forall x \in E \setminus \{0\},\ \varphi(x, x) > 0.$$

Exercice 1. Soit $E = \mathbb{R}[X]$. Pour $(P, Q) \in E^2$, on pose $(P|Q) = \displaystyle\int_{-1}^{1} \frac{P(t)Q(t)}{\sqrt{1-t^2}}\, dt$.

Montrer que $(\ |\ )$ est un produit scalaire sur E.

Solution 1. • Soit $(P, Q) \in E^2$. La fonction $t \mapsto \dfrac{P(t)Q(t)}{\sqrt{1-t^2}}$ est continue sur $]-1, 1[$. De plus, la fonction $t \mapsto \dfrac{P(t)Q(t)}{\sqrt{1+t}}$ est continue en 1 (resp. la fonction $t \mapsto \dfrac{P(t)Q(t)}{\sqrt{1-t}}$ est continue en -1) et en particulier bornée au voisinage de 1 (resp. -1). Par suite, $\dfrac{P(t)Q(t)}{\sqrt{1-t^2}} = \dfrac{P(t)Q(t)}{\sqrt{1+t}} \times \dfrac{1}{\sqrt{1-t}} \underset{t\to 1,\ t<1}{=} O\left(\dfrac{1}{\sqrt{1-t}}\right)$ (resp. $\dfrac{P(t)Q(t)}{\sqrt{1-t^2}} = \dfrac{P(t)Q(t)}{\sqrt{1-t}} \times \dfrac{1}{\sqrt{1+t}} \underset{t\to -1,\ t>-1}{=} O\left(\dfrac{1}{\sqrt{1+t}}\right)$). On en déduit que la fonction $t \mapsto \dfrac{P(t)Q(t)}{\sqrt{1-t^2}}$ est intégrable sur $]-1, 1[$ et donc que $(P|Q)$ existe dans $\mathbb{R}$.

• Soit $(P, Q) \in E^2$. $(Q|P) = \displaystyle\int_{-1}^{1} \frac{Q(t)P(t)}{\sqrt{1-t^2}}\, dt = \int_{-1}^{1} \frac{P(t)Q(t)}{\sqrt{1-t^2}}\, dt = (P|Q)$. Donc, $(\ |\ )$ est une forme symétrique.

• Soient $(P_1, P_2, Q) \in E^3$ et $(\lambda_1, \lambda_2) \in \mathbb{R}^2$.

$$\begin{aligned}
(\lambda_1 P_1 + \lambda_2 P_2 | Q) &= \int_{-1}^{1} \frac{(\lambda_1 P_1(t) + \lambda_2 P_2(t))\, Q(t)}{\sqrt{1-t^2}}\, dt \\
&= \lambda_1 \int_{-1}^{1} \frac{P_1(t)Q(t)}{\sqrt{1-t^2}}\, dt + \lambda_2 \int_{-1}^{1} \frac{P_2(t)Q(t)}{\sqrt{1-t^2}}\, dt \text{ (par linéarité de l'intégrale)} \\
&= \lambda_1 (P_1|Q) + \lambda_2 (P_2|Q).
\end{aligned}$$

Donc, $(\ |\ )$ est une forme linéaire par rapport à sa première variable puis bilinéaire par symétrie.

• Soit $P \in E$. $(P|P) = \displaystyle\int_{-1}^{1} \frac{P^2(t)}{\sqrt{1-t^2}}\, dt \geqslant 0$ par positivité de l'intégrale. Donc, $(\ |\ )$ est une forme positive.

• Soit $P \in E$.

$$\begin{aligned}(P|P) = 0 &\Rightarrow \int_{-1}^{1} \frac{P^2(t)}{\sqrt{1-t^2}}\,dt = 0\\
&\Rightarrow \forall t \in]-1,1[,\ \frac{P^2(t)}{\sqrt{1-t^2}} = 0 \text{ (fonction continue, positive, d'intégrale nulle)}\\
&\Rightarrow \forall t \in]-1,1[,\ P(t) = 0\\
&\Rightarrow P = 0 \text{ (polynôme ayant une infinité de racines).}\end{aligned}$$

Donc, $(\ |\ )$ est une forme définie.

En résumé, $(\ |\ )$ est une forme bilinéaire, symétrique, définie, positive sur E et donc $(\ |\ )$ est un produit scalaire sur E.

Commentaire. Dans l'exercice précédent, on peut traiter la positivité et la définition ensembles mais il faut être méticuleux : soit $P \in E \setminus \{0\}$. P admet donc un nombre fini de racines. Puisque $]-1,1[$ est infini, la fonction $t \mapsto \dfrac{P^2(t)}{\sqrt{1-t^2}}$ est continue, positive et non nulle sur $]-1,1[$. On en déduit que

$$(P|P) = \int_{-1}^{1} \frac{P^2(t)}{\sqrt{1-t^2}}\,dt > 0.$$

Définition 2. Un **espace préhilbertien réel** est un couple (E,φ) où E est un $\mathbb{R}$-espace vectoriel et φ un produit scalaire sur E.
Si de plus, E est un $\mathbb{R}$-espace vectoriel de dimension finie, (E,φ) est un **espace euclidien**.

Commentaire. Si φ et ψ sont deux produits scalaires différents sur un même espace E, les espaces préhilbertiens (E,φ) et (E,ψ) sont différents ou encore, quand on change de produit scalaire, on change d'espace préhilbertiens.

2) Espaces préhilbertiens de référence

• Soit $E = \mathbb{R}^n$. Le produit scalaire canonique sur E est défini par

$$\forall\left((x_1,\ldots,x_n),(y_1,\ldots,y_n)\right) \in (\mathbb{R}^n)^2,\ (x_1,\ldots,x_n)\,|\,(y_1,\ldots,y_n) = \sum_{i=1}^{n} x_i y_i.$$

• Soit $E = \mathscr{M}_{n,1}(\mathbb{R})$. Le produit scalaire canonique sur E est défini par

$$\text{Pour tout } X = (x_i)_{1\leqslant i\leqslant n} \text{ et tout } Y = (y_i)_{1\leqslant i\leqslant n} \text{ de } \mathscr{M}_{n,1}(\mathbb{R}),\ X|Y = {}^tX \times Y = \sum_{i=1}^{n} x_i y_i.$$

• Soit $E = \mathscr{M}_n(\mathbb{R})$. Le produit scalaire canonique sur E est défini par

$$\text{Pour tout } A = (a_{i,j})_{1\leqslant i,j\leqslant n} \text{ et tout } B = (b_{i,j})_{1\leqslant i,j\leqslant n} \text{ de } \mathscr{M}_n(\mathbb{R}),\ A|B = \mathrm{Tr}\left({}^tA \times B\right) = \sum_{1\leqslant i,j\leqslant n} a_{i,j} b_{i,j}.$$

• Soit $E = C^0([a,b],\mathbb{R})$. Un produit scalaire sur E est défini par

$$\forall (f,g) \in E^2,\ f|g = \int_a^b f(x)g(x)\,dx.$$

• Posons $\ell^2(\mathbb{R}) = \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} / \displaystyle\sum_{n=0}^{+\infty} u_n^2 < +\infty\right\}$. $\ell^2(\mathbb{R})$ est l'ensemble des suites réelles de carré sommable.

Vérifions tout d'abord que $\ell^2(\mathbb{R})$ est un sous-espace vectoriel de $\left(\mathbb{R}^{\mathbb{N}},+,.\right)$.

- La suite nulle est dans $\ell^2(\mathbb{R})$.
- Soit $(u,v) \in \left(\ell^2(\mathbb{R})\right)^2$ et soit $(\lambda,\mu) \in \mathbb{R}^2$. Puisque $(|u|-|v|)^2 \geqslant 0$, on a $|uv| \leqslant \dfrac{1}{2}\left(u^2+v^2\right)$. On en déduit que

$$\begin{aligned}0 \leqslant (\lambda u + \mu v)^2 &= \lambda^2 u^2 + 2\lambda\mu uv + \mu^2 v^2\\
&\leqslant \lambda^2 u^2 + 2|\lambda||\mu||uv| + \mu^2 v^2\\
&\leqslant \lambda^2 u^2 + |\lambda||\mu|\left(u^2+v^2\right) + \mu^2 v^2 = \left(\lambda^2 + |\lambda||\mu|\right)u^2 + \left(|\lambda||\mu| + \mu^2\right)v^2.\end{aligned}$$

La série de terme général $\left(\lambda^2+|\lambda||\mu|\right)u_n^2+\left(|\lambda||\mu|+\mu^2\right)v_n^2$, $n\in\mathbb{N}$, est convergente et donc la série de terme général $(\lambda u_n+\mu v_n)^2$, $n\in\mathbb{N}$, est convergente ou encore $\lambda u+\mu v\in\ell^2(\mathbb{R})$.

On a montré que $\ell^2(\mathbb{R})$ est un sous-espace vectoriel de $\left(\mathbb{R}^{\mathbb{N}},+,.\right)$. Pour $(u,v)\in\left(\ell^2(\mathbb{R})\right)^2$, on pose alors

$$u|v=\sum_{n=0}^{+\infty}u_nv_n.$$

Vérifions que $(\ |\ )$ est un produit scalaire sur $\ell^2(\mathbb{R})$.

- Pour $(u,v)\in\left(\ell^2(\mathbb{R})\right)^2$, $|uv|\leqslant\dfrac{1}{2}\left(u^2+v^2\right)$. Ceci montre que la série de terme général u_nv_n est absolument convergente et donc convergente. On en déduit que $u|v$ existe dans $\mathbb{R}$.
- La symétrie et la bilinéarité de $(\ |\ )$ sont claires.
- Soit $u\in\ell^2(\mathbb{R})\setminus\{0\}$. Alors, tous les u_n^2 sont positifs ou nuls et l'un des u_n^2 est strictement positif. On en déduit que

$$u|u=\sum_{n=0}^{+\infty}u_n^2>0.$$

Finalement, $(\ |\ )$ est un produit scalaire sur $\ell^2(\mathbb{R})$ ou encore $\left(\ell^2(\mathbb{R}),(\ |\ )\right)$ est un espace préhilbertien réel.

• Soit I un intervalle de $\mathbb{R}$. Posons $L_c^2(I,\mathbb{R})=\left\{f\in C^0(I,\mathbb{R})/\ \displaystyle\int_I f^2(x)\ dx<+\infty\right\}$. $L_c^2(\mathbb{R})$ est l'ensemble des fonctions réelles, continues sur I et de carré intégrable sur I.

Vérifions tout d'abord que $L_c^2(I,\mathbb{R})$ est un sous-espace vectoriel de $\left(C^0(I,\mathbb{R}),+,.\right)$. Déjà, $L_c^2(I,\mathbb{R})\subset C^0(I,\mathbb{R})$. Ensuite,

- la fonction nulle est dans $L_c^2(I,\mathbb{R})$.
- Soit $(f,g)\in\left(L_c^2(I,\mathbb{R})\right)^2$ et soit $(\lambda,\mu)\in\mathbb{R}^2$.

$$\begin{aligned}0\leqslant(\lambda f+\mu g)^2&=\lambda^2f^2+2\lambda\mu fg+\mu^2g^2\\&\leqslant\lambda^2f^2+2|\lambda||\mu||fg|+\mu^2g^2\\&\leqslant\lambda^2f^2+|\lambda||\mu|\left(f^2+g^2\right)+\mu^2g^2=\left(\lambda^2+|\lambda||\mu|\right)f^2+\left(|\lambda||\mu|+\mu^2\right)g^2.\end{aligned}$$

La fonction $\left(\lambda^2+|\lambda||\mu|\right)f^2+\left(|\lambda||\mu|+\mu^2\right)g^2$ est intégrable sur I et donc la fonction $(\lambda f+\mu g)^2$ est intégrable sur I ou encore $\lambda f+\mu g\in L_c^2(I,\mathbb{R})$.

On a montré que $L_c^2(I,\mathbb{R})$ est un sous-espace vectoriel de $\left(C^0(I,\mathbb{R}),+,.\right)$. Pour $(f,g)\in\left(L_c^2(I,\mathbb{R})\right)^2$, on pose alors

$$f|g=\int_I f(x)g(x)\ dx.$$

Vérifions que $(\ |\ )$ est un produit scalaire sur $L_c^2(I,\mathbb{R})$.

- Pour $(f,g)\in\left(L_c^2(I,\mathbb{R})\right)^2$, $|fg|\leqslant\dfrac{1}{2}\left(f^2+g^2\right)$. Ceci montre que la fonction fg est intégrable sur I. On en déduit que $f|g$ existe dans $\mathbb{R}$.
- La symétrie et la bilinéarité de $(\ |\ )$ sont claires.
- Soit $f\in L_c^2(I,\mathbb{R})\setminus\{0\}$. Alors, la fonction f^2 est continue, positive et non nulle sur I. On en déduit que

$$f|f=\int_I f^2(x)\ dx>0.$$

Finalement, $(\ |\ )$ est un produit scalaire sur $L_c^2(I,\mathbb{R})$ ou encore $\left(L_c^2(I,\mathbb{R}),(\ |\ )\right)$ est un espace préhilbertien réel.

3) *Inégalité de* Cauchy-Schwarz

Théorème 1. Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Alors

$$\forall (x, y) \in E^2,\ |(x|y)| \leqslant \sqrt{(x|x)}\sqrt{(y|y)}.$$

De plus, $|(x|y)| = \sqrt{(x|x)}\sqrt{(y|y)}$ si et seulement si la famille (x, y) est liée (cas d'égalité de l'inégalité de Cauchy-Schwarz).

Démonstration. Soit $(x, y) \in E^2$. Si $x = 0$, l'inégalité de Cauchy-Schwarz est vérifiée et de plus, il s'agit d'un cas d'égalité.
Supposons maintenant $x \neq 0$. Pour tout $\lambda \in \mathbb{R}$, on pose $P(\lambda) = (\lambda x - y|\lambda x - y)$. Pour tout $\lambda \in \mathbb{R}$, $P(\lambda) \geqslant 0$ et d'autre part

$$P(\lambda) = (x|x)\lambda^2 - 2(x|y)\lambda + (y|y).$$

Puisque le coefficient $(x|x)$ de λ^2 n'est pas nul, P est un trinôme du second degré de signe constant sur $\mathbb{R}$. Son discriminant réduit est donc négatif ou nul. On en déduit que

$$0 \geqslant \Delta' = (x|y)^2 - (x|x)(y|y),$$

puis que $(x|y)^2 \leqslant (x|x)(y|y)$ puis que $\sqrt{(x|y)^2} \leqslant \sqrt{(x|x)(y|y)}$ et finalement que $|(x|y)| \leqslant \sqrt{(x|x)}\sqrt{(y|y)}$ $(*)$. De plus,

$$\begin{aligned}
(*) \text{ est une égalité} &\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \Delta' = 0)\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda_0 \in \mathbb{R}/\ P(\lambda_0) = 0)\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda_0 \in \mathbb{R}/\ (\lambda_0 x - y|\lambda_0 x - y) = 0)\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda_0 \in \mathbb{R}/\ y = \lambda_0 x)\\
&\Leftrightarrow (x, y) \text{ est liée.}
\end{aligned}$$

---

4) *Norme hilbertienne*

4-a) *Définition*

Théorème 2. (Inégalité de Minkowski)

Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Pour tout $(x, y) \in E^2$,

$$\sqrt{(x+y)|(x+y)} \leqslant \sqrt{(x|x)} + \sqrt{(y|y)},$$

avec égalité si et seulement si $x = 0$ ou $(x \neq 0$ et $\exists \lambda \in \mathbb{R}^+/\ y = \lambda x)$.

Démonstration. Soit $(x, y) \in E^2$.

$$\begin{aligned}
(x+y)|(x+y) &= (x|x) + 2(x|y) + (y|y)\\
&\leqslant (x|x) + 2|(x|y)| + (y|y)\\
&\leqslant (x|x) + 2\sqrt{(x|x)}\sqrt{(y|y)} + (y|y) = \left(\sqrt{(x|x)} + \sqrt{(y|y)}\right)^2 \quad (*),
\end{aligned}$$

et donc $\sqrt{(x+y)|(x+y)} \leqslant \sqrt{(x|x)} + \sqrt{(y|y)}$. De plus, on a l'égalité si et seulement si chaque inégalité écrite est une égalité. La deuxième inégalité est une égalité si et seulement si $(x = 0)$ ou $(x \neq 0$ et $\exists \lambda \in \mathbb{R}/\ y = \lambda x)$ (cas d'égalité de l'inégalité de Cauchy-Schwarz). La première inégalité est une égalité si et seulement si $(x|y) = |(x|y)|$ ou encore si et seulement si $(x|y) \geqslant 0$. Donc,

$$\begin{aligned}
(*) \text{ est une égalité} &\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda \in \mathbb{R}/\ y = \lambda x \text{ et } (x|y) \geqslant 0)\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda \in \mathbb{R}/\ y = \lambda x \text{ et } \lambda(x|x) \geqslant 0)\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda \in \mathbb{R}/\ y = \lambda x \text{ et } \lambda \geqslant 0 \text{ (car } (x|x) > 0))\\
&\Leftrightarrow (x = 0) \text{ ou } (x \neq 0 \text{ et } \exists \lambda \in \mathbb{R}^+/\ y = \lambda x).
\end{aligned}$$

Théorème 3. Soit $(E, (\ |\ ))$ un espace préhilbertien réel. L'application $x \mapsto \sqrt{(x|x)}$ est une norme sur E.

Démonstration. • Par définition d'un produit scalaire, $\forall x \in E$, $(x|x) \geqslant 0$. Donc, $x \mapsto \sqrt{(x|x)}$ est bien une application de E dans $\mathbb{R}$.

• $\forall x \in E$, $\sqrt{(x|x)} \geqslant 0$.
• Pour $x \in E$, $\sqrt{(x|x)} = 0 \Rightarrow (x|x) = 0 \Rightarrow x = 0$.
• Pour $x \in E$ et $\lambda \in \mathbb{R}$, $\sqrt{(\lambda x|\lambda x)} = \sqrt{\lambda^2}\sqrt{(x|x)} = |\lambda|\sqrt{(x|x)}$.
• L'inégalité triangulaire est l'inégalité de MINKOWSKI.

On a montré que l'application $x \mapsto \sqrt{(x|x)}$ est une norme sur E.

DÉFINITION 3. Pour $x \in E$, on pose $\|x\| = \sqrt{(x|x)}$. $\|\ \|$ est la **norme hilbertienne** sur E ou encore $\|\ \|$ est la **norme associée au produit scalaire** $(\ |\ )$.

Commentaire. L'inégalité de CAUCHY-SCHWARZ s'écrit dorénavant : $\forall (x, y) \in E^2$, $|(x|y)| \leqslant \|x\|\|y\|$. En particulier,

$$\forall x \neq 0,\ \forall y \neq 0,\ -1 \leqslant \frac{(x|y)}{\|x\|\|y\|} \leqslant 1.$$

On décide alors de poser

$$\forall x \neq 0,\ \forall y \neq 0,\ \cos(x, y) = \frac{(x|y)}{\|x\|\|y\|}.$$

4-b) Carré scalaire

Le produit scalaire d'un vecteur x par lui-même est encore appelé le **carré scalaire** de x. On le note x^2. On a donc

$$\forall x \in E,\ x^2 = (x|x) = \|x\|^2.$$

La bilinéarité du produit scalaire fournit immédiatement les identités remarquables suivantes :

Théorème 4. Pour tout $(x, y) \in E^2$,

• $\|x + y\|^2 = (x + y)^2 = x^2 + 2(x|y) + y^2 = \|x\|^2 + 2(x|y) + \|y\|^2$.
• $\|x - y\|^2 = (x + y)^2 = x^2 - 2(x|y) + y^2 = \|x\|^2 - 2(x|y) + \|y\|^2$.
• $(x + y)|(x - y) = x^2 - y^2 = \|x\|^2 - \|y\|^2$.

4-c) Théorème de PYTHAGORE

Le théorème 4 fournit immédiatement

Théorème 5. (Théorème de PYTHAGORE)

$$\forall (x, y) \in E^2,\ \|x + y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow (x|y) = 0.$$

Commentaire. On peut noter que dorénavant, le théorème de PYTHAGORE est directement une équivalence contrairement à ce qui se passe au collège puis au lycée où on parle de théorème de PYTHAGORE et de réciproque du théorème de PYTHAGORE. ❑

Si on cherche à généraliser à plus de deux vecteurs le théorème précédent, on n'a plus une équivalence :

Théorème 6. Soient $x_1, \ldots, x_n$ $(n \geqslant 2)$ n vecteurs de E.
Si pour tout $(i, j) \in [\![1, n]\!]^2$ tel que $i \neq j$, on a $(x_i|x_j) = 0$, **alors** $\|x_1 + \ldots + x_n\|^2 = \|x_1\|^2 + \ldots + \|x_n\|^2$.

Démonstration. Soit $(x_1, \ldots, x_n) \in E^n$ tel que pour tout $(i, j) \in [\![1, n]\!]^2$ tel que $i \neq j$, on a $(x_i|x_j) = 0$. Alors, par bilinéarité du produit scalaire,

$$\begin{aligned}\|x_1 + \ldots + x_n\|^2 = (x_1 + \ldots + x_n)^2 &= \sum_{1 \leqslant i,j \leqslant n} (x_i|x_j) \\ &= \sum_{i=1}^{n} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} (x_i|x_j) \\ &= \sum_{i=1}^{n} \|x_i\|^2.\end{aligned}$$

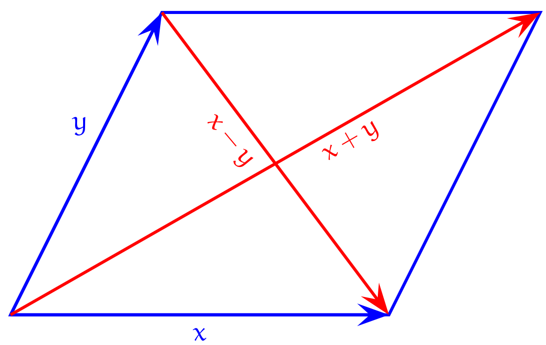
4-d) Identité du parallélogramme

En additionnant les deux premières égalités du théorème 4, on obtient

Théorème 7. (Identité du parallélogramme)

$$\forall (x, y) \in E^2,\ \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Commentaire. L'égalité ci-dessus doit être comprise sous la forme $\|x+y\|^2 + \|x-y\|^2 = \|x\|^2 + \|y\|^2 + \|x\|^2 + \|y\|^2$. Cette identité signifie que la somme des carrés des côtés d'un parallélogramme est égale à la somme des carrés des diagonales de ce parallélogramme.



❑

4-e) Identités de polarisation

La norme hilbertienne est exprimée en fonction du produit scalaire $(\ |\ )$:

$$\forall x \in E,\ \|x\| = \sqrt{(x|x)}.$$

On fait maintenant le contraire à savoir on exprime le produit scalaire en fonction de la norme. A partir du théorème 4, on obtient

Théorème 8. (Identités de polarisation)

- $\forall x \in E$, $x^2 = (x|x) = \|x\|^2$.
- $\forall (x, y) \in E^2$, $(x|y) = \dfrac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right) = \dfrac{1}{2}\left(\|x\|^2 + \|y\|^2 - \|x-y\|^2\right) = \dfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$.

Commentaire. Le produit scalaire de deux vecteurs non nuls est

$$(x|y) = \|x\|\|y\| \cos(x, y).$$

Sa valeur est définie à partir de longueurs et d'angles. Le théorème précédent montre qu'un produit scalaire peut s'exprimer à partir de longueurs uniquement.

❑

II -Orthogonalité

1) Vecteurs orthogonaux

DÉFINITION 4. Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soient x et y deux éléments de E.
x et y sont **orthogonaux** si et seulement si $(x|y) = 0$.

2) Orthogonal d'une partie de E

DÉFINITION 5. Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit A une partie non vide de E.
L'orthogonal de A, noté $A^\perp$, est l'ensemble des éléments de E qui sont orthogonaux à tous les éléments de A :

$$A^\perp = \{x \in E / \ \forall y \in A,\ (x|y) = 0\}.$$

Convention. $\varnothing^\perp = E$.

Notation. Si x est un élément de E, $\{x\}^\perp$ se note plus simplement $x^\perp$.

Théorème 9. Soit $(E, (\ |\ ))$ un espace préhilbertien réel.
Pour toute partie A de E, $A^\perp$ est un sous-espace vectoriel de E.

Démonstration. Si $A = \varnothing$, alors conventionnellement $A^\perp = E$. Dans ce cas, $A^\perp$ est un sous-espace vectoriel de E. Supposons maintenant $A \neq \varnothing$.

- Pour tout y de A, $(0|y) = 0$ et donc $0 \in A^\perp$.
- Soient $(x_1, x_2) \in \left(A^\perp\right)^2$ et $(\lambda_1, \lambda_2) \in \mathbb{R}^2$. Alors, pour tout y de A,

$$(\lambda_1 x_1 + \lambda_2 x_2)\,|y = \lambda_1\,(x_1|y) + \lambda_2\,(x_2|y) = 0 + 0 = 0,$$

et donc $\lambda_1 x_1 + \lambda_2 x_2 \in A^\perp$.

Ceci montre que $A^\perp$ est un sous-espace vectoriel de E.

Théorème 10. Soit $(E, (\ |\ ))$ un espace préhilbertien réel. $0^\perp = E$ et $E^\perp = 0$.

Démonstration.
- Tout vecteur est orthogonal à 0 et donc $0^\perp = E$.
- Soit $x \in E^\perp$. x est orthogonal à tout vecteur de E et est en particulier orthogonal à lui-même. Par suite, $(x|x) = 0$ puis $x = 0$. Ainsi, $E^\perp \subset \{0\}$ puis $E^\perp = \{0\}$ car $E^\perp$ est un sous-espace vectoriel de E.

Théorème 11. Soit $(E, (\ |\ ))$ un espace préhilbertien réel.
1) $\forall (A, B) \in (\mathscr{P}(E))^2,\ \left(A \subset B \Rightarrow B^\perp \subset A^\perp\right)$.
2) $\forall A \in \mathscr{P}(E),\ A^\perp = (\mathrm{Vect}(A))^\perp$.

Démonstration. • Soient A et B deux parties de E telles que $A \subset B$. Si $A = \varnothing$, alors $A^\perp = E$ puis $B^\perp \subset A^\perp$. Si $A \neq \varnothing$, alors $B \neq \varnothing$. Soit $x \in E$.

$$x \in B^\perp \Rightarrow \forall y \in B,\ (x|y) = 0 \Rightarrow \forall y \in A,\ (x|y) = 0 \Rightarrow x \in A^\perp.$$

Ceci montre que $B^\perp \subset A^\perp$.

• Soit A une partie de E. Si $A = \varnothing$, alors $\mathrm{Vect}(A) = \{0\}$ puis $A^\perp = E = (\mathrm{Vect}(A))^\perp$. Supposons maintenant $A \neq \varnothing$. Puisque $A \subset \mathrm{Vect}(A)$, on a $(\mathrm{Vect}(A))^\perp \subset A^\perp$.
Inversement, un élément de l'orthogonal de A est encore orthogonal à toute combinaison linéaire d'éléments de A par bilinéarité du produit scalaire et donc est dans l'orthogonal de $\mathrm{Vect}(A)$. Ceci montre que $A^\perp \subset (\mathrm{Vect}(A))^\perp$ et finalement que $(\mathrm{Vect}(A))^\perp = A^\perp$.

Commentaire. Le résultat précédent est très utilisé dans la pratique : pour qu'un vecteur soit orthogonal à un sous-espace vectoriel F, il faut et il suffit que ce vecteur soit orthogonal à une famille génératrice de F. Un cas particulier de ce résultat est utilisé en géométrie en dimension 3 au lycée : une droite affine D est perpendiculaire à un plan affine P si et seulement si D est orthogonale à deux droites sécantes du plan P. ❏

Théorème 12. Soit $(E, (\ |\ ))$ un espace préhilbertien réel.
1) Pour tout sous-espace vectoriel F de E, $F \subset \left(F^\perp\right)^\perp$.
2) Pour tout sous-espace vectoriel F de E, $F \cap F^\perp = \{0\}$.

Démonstration. Soit F un sous-espace vectoriel de E.

• Un vecteur de F est orthogonal à tout vecteur orthogonal à F et donc est dans $\left(F^\perp\right)^\perp$. Ceci montre que $F \subset \left(F^\perp\right)^\perp$.

• Soit $x \in F \cap F^\perp$. Alors $(x|x) = 0$ puis $x = 0$. Ceci montre que $F \cap F^\perp \subset \{0\}$ puis que $F \cap F^\perp = \{0\}$ car $F \cap F^\perp$ est un sous-espace vectoriel de E.

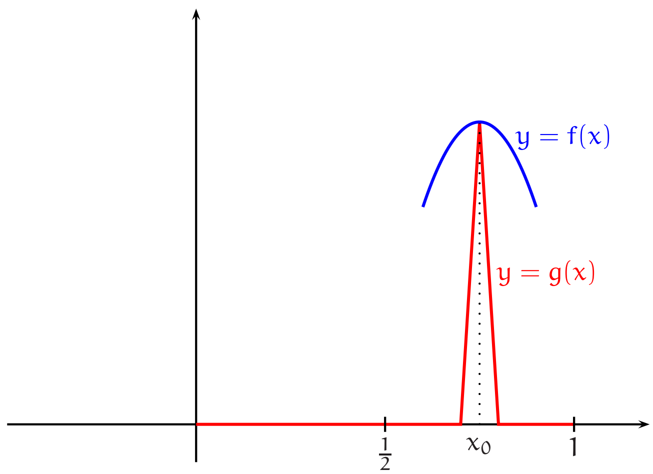
Commentaire. Le théorème précédent peut surprendre. On a toujours $F \subset \left(F^\perp\right)^\perp$ mais on ne précise pas que $F = \left(F^\perp\right)^\perp$. De même, on a toujours $F \cap F^\perp = \{0\}$ mais on ne précise pas que $E = F \oplus F^\perp$. Il est vrai que ces égalités sont exactes quand E est de dimension finie mais on va voir avec l'exemple qui suit qu'il est possible d'avoir $F \underset{\neq}{\subset} \left(F^\perp\right)^\perp$ et $F + F^\perp \underset{\neq}{\subset} E$ quand

E est de dimension infinie. Une représentation mentale d'une situation en dimension infinie par un dessin en dimension 3 peut donc être une représentation fausse. ❑

Exemple. Soit $E = C^0([0,1], \mathbb{R})$. Pour $(f,g) \in E^2$, on pose $(f|g) = \displaystyle\int_0^1 f(x)g(x)\, dx$ (on sait que $(\ |\ )$ est un produit scalaire sur E).

Soit $H = \left\{ g \in E / \displaystyle\int_0^{1/2} g(x)\, dx = 0 \right\}$. H est un sous-espace vectoriel de E (car H est le noyau de la forme linéaire $g \mapsto \displaystyle\int_0^{1/2} g(x)\, dx$) et $H \neq E$ car la fonction constante $f\ :\ x \mapsto 1$ est un élément de E qui n'est pas dans H. On va montrer que $H^\perp = \{0\}$ ce qui aura pour conséquence le fait que $\left(H^\perp\right)^\perp = \{0\}^\perp = E \neq H$ et $H + H^\perp = H + \{0\} = H \neq E$.

Soit $f \in H^\perp$. Supposons par l'absurde qu'il existe $x_0 \in \left]\frac{1}{2}, 1\right[$ tel que $f(x_0) \neq 0$. Quite à remplacer f par $-f$, on peut supposer que $f(x_0) > 0$. Par continuité de f en x_0, il existe $\alpha > 0$ tel que $[x_0 - \alpha, x_0 + \alpha] \subset \left]\frac{1}{2}, 1\right[$ et $\forall x \in [x_0 - \alpha, x_0 + \alpha]$, $f(x) \geqslant 0$. Soit alors g la fonction continue sur $[0,1]$, nulle sur $[0, x_0 - \alpha] \cup [x_0 + \alpha, 1]$, affine sur $[x_0 - \alpha, x_0]$ et sur $[x_0, x_0 + \alpha]$ et prenant la même valeur que f en x_0.



g est nulle sur $\left[0, \frac{1}{2}\right]$ et en particulier, $g \in H$. On en déduit que $\displaystyle\int_0^1 f(x)g(x)\, dx = 0$ puis que $\displaystyle\int_{x_0-\alpha}^{x_0+\alpha} f(x)g(x)\, dx = 0$. Mais la fonction fg est continue, positive et non nulle sur $[x_0 - \alpha, x_0 + \alpha]$ (car $f(x_0)g(x_0) = (f(x_0))^2 > 0$) et donc $\displaystyle\int_{x_0-\alpha}^{x_0+\alpha} f(x)g(x)\, dx > 0$. Ceci est une contradiction et donc $f_{/]\frac{1}{2},1[} = 0$ puis $f_{/[\frac{1}{2},1]} = 0$ par continuité de f en $\frac{1}{2}$ et en 1.

Soit alors $g = f - 2\displaystyle\int_0^{1/2} f(x)\, dx$ a une intégrale nulle sur $\left[0, \frac{1}{2}\right]$ et donc est dans H. Puisque $f \in H^\perp$, on a $\displaystyle\int_0^1 f(x)g(x)\, dx = 0$ puis $\displaystyle\int_0^{1/2} f(x)g(x)\, dx = 0$ car f est nulle sur $\left[\frac{1}{2}, 1\right]$ et donc $\displaystyle\int_0^{1/2} f(t)\left(f(t) - 2\int_0^{1/2} f(x)\, dx\right)\, dt = 0$ puis

$$\left(\int_0^{1/2} f(x)\, dx\right)^2 = \frac{1}{2}\int_0^{1/2} f^2(x)\, dx.$$

Cette dernière égalité s'écrit encore $\left(\displaystyle\int_0^{1/2} 1 \times f(x)\, dx\right)^2 = \left(\displaystyle\int_0^{1/2} 1^2\, dx\right)\left(\displaystyle\int_0^{1/2} f^2(x)\, dx\right)$. On est en présence d'un cas d'égalité de l'inégalité de Cauchy-Schwarz. On en déduit que la famille $(1, f)$ est liée puis que f est constante sur $\left[0, \frac{1}{2}\right]$ et donc nulle sur $\left[0, \frac{1}{2}\right]$ puisque $f\left(\frac{1}{2}\right) = 0$.

Finalement, si $f \in H^\perp$ alors $f = 0$ et donc $H^\perp = \{0\}$. ❑

Théorème 13. Soit $(E, (\ |\ ))$ un espace préhilbertien réel.
1) Pour tous sous-espaces vectoriels F et G de E, $F^\perp \cap G^\perp = (F + G)^\perp$.
2) Pour tous sous-espaces vectoriels F et G de E, $F^\perp + G^\perp \subset (F \cap G)^\perp$.

Démonstration. Soient F et G deux sous-espaces vectoriels de E.

• $F \subset F+G$ et $G \subset F+G$. Donc $(F+G)^{\perp} \subset F^{\perp}$ et $(F+G)^{\perp} \subset G^{\perp}$ puis $(F+G)^{\perp} \subset F^{\perp} \cap G^{\perp}$. Inversement, un élément x de $F^{\perp} \cap G^{\perp}$ est orthogonal à tout élément de F et de G et donc à toute somme d'un élément de F et d'un élément de G, par bilinéarité du produit scalaire. Ceci montre que $F^{\perp} \cap G^{\perp} \subset (F+G)^{\perp}$ et finalement que $(F+G)^{\perp} = F^{\perp} \cap G^{\perp}$.

• $F \cap G \subset F$ et $F \cap G \subset G$. Donc, $F^{\perp} \subset (F \cap G)^{\perp}$ et $G^{\perp} \subset (F \cap G)^{\perp}$ puis $F^{\perp} \cup G^{\perp} \subset (F \cap G)^{\perp}$ puis $\mathrm{Vect}\left(F^{\perp} \cup G^{\perp}\right) \subset (F \cap G)^{\perp}$ (car $(F \cap G)^{\perp}$ est un sous-espace vectoriel de E) et finalement $F^{\perp} + G^{\perp} \subset (F \cap G)^{\perp}$.

---

3) Familles orthogonales, familles orthonormales

3-a) Définitions

Définition 6. Soit $(E, (\,|\,))$ un espace préhilbertien réel. Soit I un ensemble non vide d'indices. Soit $(e_i)_{i \in I}$ une famille de vecteurs de E indexée par I.

• La famille $(e_i)_{i \in I}$ est **orthogonale** si et seulement si $\forall (i,j) \in I^2$, $i \neq j \Rightarrow (e_i|e_j) = 0$.

• La famille $(e_i)_{i \in I}$ est **orthonormale** si et seulement si $\forall (i,j) \in I^2$, $(e_i|e_j) = \left\{\begin{array}{l} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{array}\right. = \delta_{i,j}$.

Exemple. On munit $\mathbb{R}^n$ du produit scalaire usuel : pour $x = (x_i)_{1 \leqslant i \leqslant n} \in \mathbb{R}^n$ et $y = (y_i)_{1 \leqslant i \leqslant n} \in \mathbb{R}^n$, on pose

$$(x|y) = \sum_{i=1}^{n} x_i y_i.$$

Pour ce produit scalaire, la base canonique $(e_i)_{1 \leqslant i \leqslant n}$ est orthonormée car pour $(i,j) \in [\![1,n]\!]^2$,

$$(e_i|e_j) = \sum_{k=1}^{n} \delta_{k,i}\delta_{k,j} = \delta_{i,j}.$$

C'est pour cette raison que le produit scalaire usuel sur $\mathbb{R}^n$ s'appelle aussi le **produit scalaire canonique** sur $\mathbb{R}^n$. ❑

3-b) Propriétés

Théorème 14. Soit $(E, (\,|\,))$ un espace préhilbertien réel.

• Une famille orthogonale de vecteurs tous non nuls est libre.
• Une famille orthonormale est libre.

Démonstration. Il suffit de prouver le résultat pour une famille finie.

• Soit $(e_i)_{1 \leqslant i \leqslant n}$ une famille orthogonale de vecteurs tous non nuls. Soit $(\lambda_i)_{1 \leqslant i \leqslant n} \in \mathbb{R}^n$.

$$\begin{aligned}
\sum_{j=1}^{n} \lambda_j e_j = 0 &\Rightarrow \forall i \in [\![1,n]\!],\ \left(e_i \middle| \sum_{j=1}^{n} \lambda_j e_j\right) = 0 \\
&\Rightarrow \forall i \in [\![1,n]\!],\ \sum_{j=1}^{n} \lambda_j\,(e_i|e_j) = 0 \\
&\Rightarrow \forall i \in [\![1,n]\!],\ \lambda_i\,(e_i|e_i) = 0 \\
&\Rightarrow \forall i \in [\![1,n]\!],\ \lambda_i = 0 \text{ (car } e_i \neq 0).
\end{aligned}$$

On a montré que la famille $(e_i)_{1 \leqslant i \leqslant n}$ est libre.

• Soit $(e_i)_{1 \leqslant i \leqslant n}$ une famille orthonormale. $(e_i)_{1 \leqslant i \leqslant n}$ est en particulier une famille orthogonale de vecteurs tous nuls et donc $(e_i)_{1 \leqslant i \leqslant n}$ est une famille libre.

3-c) Orthonormalisation de Schmidt

En maths sup, on décrit le procédé d'orthonormalisation de Schmidt dans le cas d'une famille finie. On généralise maintenant le procédé à une famille infinie (dénombrable).

Théorème 15. Soit $(E, (\,|\,))$ un espace préhilbertien réel. Soit $(u_n)_{n\in\mathbb{N}}$ une famille libre.

Il existe une famille orthonormale $(e_n)_{n\in\mathbb{N}}$ et une seule telle que

- $\forall n \in \mathbb{N}$, $\text{Vect}\,(e_k)_{0\leqslant k\leqslant n} = \text{Vect}\,(u_k)_{0\leqslant k\leqslant n}$;
- $\forall n \in \mathbb{N}$, $(e_n|u_n) > 0$.

Démonstration.

On montre l'unicité de la famille $(e_n)_{n\in\mathbb{N}}$ par récurrence sur n.

- Il existe nécessairement $\lambda \in \mathbb{R}$ tel que $e_0 = \lambda u_0$. $\|e_0\| = 1$ impose $|\lambda|\,\|u_0\| = 1$ et donc nécessairement $\lambda = \pm\dfrac{1}{\|u_0\|}$.
Enfin, $(e_0|u_0) = \lambda\,(u_0|u_0)$ et donc nécessairement $\lambda = \dfrac{(e_0|u_0)}{(u_0|u_0)} \geqslant 0$. Ainsi, nécessairement $e_0 = \dfrac{1}{\|u_0\|}u_0$ ce qui montre l'unicité de u_0.
- Soit $n \in \mathbb{N}$. Supposons avoir démontré l'unicité de e_0, ..., e_n.
Nécessairement, $e_{n+1} \in \text{Vect}\,(u_0, \ldots, u_n, u_{n+1}) = \text{Vect}\,(e_0, \ldots, e_n, u_{n+1})$ et donc, il existe des réels λ, μ_0, ..., μ_n tels que $e_{n+1} = \lambda u_{n+1} + \displaystyle\sum_{k=0}^{n}\mu_k e_k$. Soit $k \in [\![0, n]\!]$. Nécessairement
$$0 = (e_{n+1}|e_k) = \lambda\,(u_{n+1}|e_k) + \sum_{l=0}^{n}\mu_l\,(e_l|e_k) = \lambda\,(u_{n+1}|e_k) + \mu_k$$
et donc nécessairement $e_{n+1} = \lambda e'_{n+1}$ où $e'_{n+1} = u_{n+1} - \displaystyle\sum_{k=0}^{n}(u_{n+1}|e_k)\,e_k$. Ensuite, $\|e_{n+1}\| = 1$ impose
$e_{n+1} = \pm\dfrac{1}{\|e'_{n+1}\|}e'_{n+1}$. Enfin,
$$0 < (e_{n+1}, u_{n+1}) = \left(e_{n+1}, e'_{n+1} + \sum_{k=0}^{n}(u_{n+1}|e_k)\,e_k\right) = \left(e_{n+1}, e'_{n+1}\right) = \lambda\,(e_{n+1}, e_{n+1}) = \lambda$$
impose $e_{n+1} = \dfrac{1}{\|e'_{n+1}\|}e'_{n+1}$.

L'unicité de la famille $(e_n)_{n\in\mathbb{N}}$ est démontrée par récurrence.

On montre l'existence de la famille $(e_n)_{n\in\mathbb{N}}$ par récurrence sur n

- La famille $(u_n)_{n\in\mathbb{N}}$ est libre donc $u_0 \neq 0$. On peut poser $e_0 = \dfrac{1}{\|u_0\|}u_0$. La famille (e_0) est orthonormale. Le vecteur e_0 est un vecteur non nul de la droite $\text{Vect}\,(u_0)$ et donc $\text{Vect}\,(e_0) = \text{Vect}\,(u_0)$. Enfin,
$$(u_0|e_0) = \frac{1}{\|u_0\|}\,(u_0|u_0) = \|u_0\| > 0.$$
- Soit $n \geqslant 0$. Supposons qu'il existe une famille orthonormale $(e_k)_{0\leqslant k\leqslant n}$ telle que $\forall k \in [\![0, n]\!]$,
$\text{Vect}\,(e_i)_{0\leqslant i\leqslant k} = \text{Vect}\,(u_i)_{0\leqslant i\leqslant k}$ et $\forall k \in [\![0, n]\!]$, $(e_k|u_k) > 0$.
Soit $e'_{n+1} = u_{n+1} - \displaystyle\sum_{k=0}^{n}(u_{n+1}|e_k)\,e_k$. Puisque $u_{n+1} \notin \text{Vect}\,(u_k)_{0\leqslant n} = \text{Vect}\,(e_k)_{0\leqslant n}$, le vecteur e'_{n+1} n'est pas nul.
On peut donc poser $e_{n+1} = \dfrac{1}{\|e'_{n+1}\|}e'_{n+1}$ de sorte que e_{n+1} est unitaire. Pour $0 \leqslant k \leqslant n$,
$$\begin{aligned}(e_{n+1}|e_k) &= \frac{1}{\|e'_{n+1}\|}\left((u_{n+1}|e_k) - \sum_{l=0}^{n}(u_{n+1}|e_l)\,(e_k|e_l)\right)\\ &= \frac{1}{\|e'_{n+1}\|}\left((u_{n+1}|e_k) - \sum_{l=0}^{n}(u_{n+1}|e_l)\,\delta_{k,l}\right) \text{ (par hypothèse de récurrence)}\\ &= \frac{1}{\|e'_{n+1}\|}\left((u_{n+1}|e_k) - (u_{n+1}|e_k)\right) = 0.\end{aligned}$$
Donc, la famille $(e_k)_{0\leqslant k\leqslant n+1}$ est orthonormale.
Par hypothèse de récurrence, $\text{Vect}\,(e_0, \ldots, e_n, e_{n+1}) = \text{Vect}\,\left(e_0, \ldots, e_n, e'_{n+1}\right) = \text{Vect}\,\left(u_0, \ldots, u_n, e'_{n+1}\right)$ et donc

$$\mathrm{Vect}\left(e_0,\dots,e_n,e_{n+1}\right)=\mathrm{Vect}\left(u_0,\dots,u_n,u_{n+1}-\underbrace{\sum_{k=0}^{n}\left(u_{n+1}|e_k\right)e_k}_{\in\mathrm{Vect}(u_0,\dots,u_n)}\right)=\mathrm{Vect}\left(u_0,\dots,u_n,u_{n+1}\right).$$

Enfin,

$$\left(e_{n+1}|u_{n+1}\right)=\left(e_{n+1}|e'_{n+1}+\sum_{k=0}^{n}\left(u_{n+1}|e_k\right)e_k\right)=\left(e_{n+1}|e'_{n+1}\right)=\left(e_{n+1}|\left\|e'_{n+1}\right\|e_{n+1}\right)=\left\|e'_{n+1}\right\|>0.$$

L'existence de la famille $\left(e_n\right)_{n\in\mathbb{N}}$ est démontrée par récurrence.

Définition 7. La famille $\left(e_n\right)_{n\in\mathbb{N}}$ s'appelle l'**orthonormalisée** de la famille libre $\left(u_n\right)_{n\in\mathbb{N}}$.

Le procédé permettant d'obtenir les vecteurs e_n, $n\in\mathbb{N}$, de proche en proche est le **procédé d'orthonormalisation de Schmidt**.

Le procédé d'orthonormalisation de Schmidt est :

- $e_0=\dfrac{1}{\|u_0\|}u_0$;
- $\forall n\in\mathbb{N},\ e_{n+1}=\dfrac{1}{\left\|e'_{n+1}\right\|}e'_{n+1}$ où $e'_{n+1}=u_{n+1}-\displaystyle\sum_{k=0}^{n}\left(u_{n+1}|e_k\right)e_k$.

Commentaire. On rappelle que le procédé d'orthonormalisation de Schmidt permet d'établir en maths sup l'existence de bases orthonormées quand E est de dimension finie non nulle. Il suffit pour cela d'orthonormaliser une base de E. ❏

Exercice 2. Soit $E=\mathbb{R}[X]$. Pour $(P,Q)\in E^2$, on pose $(P|Q)=\displaystyle\int_0^1 P(t)Q(t)\,dt$.

Déterminer l'orthonormalisée de la famille $\left(1,X,X^2\right)$.

Solution 2. On pose $P_0=1$, $P_1=X$ et $P_2=X^2$ et on note (E_0,E_1,E_2) l'orthonormalisée de la famille libre (P_0,P_1,P_2).

- $\|P_0\|^2=\displaystyle\int_0^1 1\,dt=1$ et donc $E_0=\dfrac{1}{\|P_0\|}P_0=1$.
- $(P_1|E_0)=\displaystyle\int_0^1 t\,dt=\frac{1}{2}$ puis $P_1-(P_1|E_0)\,E_0=X-\dfrac{1}{2}$. Ensuite

$$\|P_1-(P_1|E_0)\,E_0\|^2=\int_0^1\left(t-\frac{1}{2}\right)^2dt=\int_0^1\left(t^2-t+\frac{1}{4}\right)dt=\frac{1}{3}-\frac{1}{2}+\frac{1}{4}=\frac{1}{12},$$

et donc $E_1=\dfrac{1}{1/2\sqrt{3}}\left(X-\dfrac{1}{2}\right)=\sqrt{3}(2X-1)$.

- $(P_2|E_0)=\displaystyle\int_0^1 t^2\,dt=\frac{1}{3}$ et $(P_2|E_1)=\sqrt{3}\displaystyle\int_0^1\left(2t^3-t^2\right)dt=\sqrt{3}\left(\frac{2}{4}-\frac{1}{3}\right)=\frac{\sqrt{3}}{6}$ puis

$$P_2-(P_2|E_0)\,E_0-(P_2|E_1)\,E_1=X^2-\frac{1}{3}-\frac{\sqrt{3}}{6}\times\sqrt{3}(2X-1)=X^2-X+\frac{1}{6}.$$

Ensuite

$$\begin{aligned}\|P_2-(P_2|E_0)\,E_0-(P_2|E_1)\,E_1\|^2&=\int_0^1\left(t^2-t+\frac{1}{6}\right)^2dt=\int_0^1\left(t^4-2t^3+\frac{4t^2}{3}-\frac{t}{3}+\frac{1}{36}\right)dt\\&=\frac{1}{5}-\frac{2}{4}+\frac{4}{9}-\frac{1}{6}+\frac{1}{36}=\frac{36-90+80-30+5}{180}=\frac{1}{180}\end{aligned}$$

et donc $E_2=\dfrac{1}{1/6\sqrt{5}}\left(X^2-X+\dfrac{1}{6}\right)=\sqrt{5}\left(6X^2-6X+1\right)$.

III - Le théorème de la projection orthogonale

1) Le théorème de la projection orthogonale

On a vu (page 9) qu'il est possible, pour un sous-espace donné F de E, que $F^\perp$ ne soit pas un supplémentaire de F. On va voir que si de plus F est de dimension finie, F et $F^\perp$ sont supplémentaires et donc la projection orthogonale sur F est bien définie.

Théorème 16. (Le théorème de la projection orthogonale)

Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit F un sous-espace vectoriel de E **de dimension finie** n. Alors,

- $E = F \oplus F^\perp$.
- Si $n \geqslant 1$ et $(e_i)_{1 \leqslant i \leqslant n}$ est une base orthonormée de F alors

$$\forall x \in E,\ p_F(x) = \sum_{k=1}^{n} (x|e_k)\, e_k$$

où $p_F(x)$ est le projeté orthogonal de x sur F.

Démonstration. Soient $x \in E$ et $y \in F$.

$$\begin{aligned}
x - y \in F^\perp &\Leftrightarrow x - y \in (\text{Vect}\,(e_1, \ldots, e_n))^\perp \Leftrightarrow x - y \in (e_1, \ldots, e_n)^\perp \\
&\Leftrightarrow \forall k \in [\![1, n]\!],\ (x - y|e_k) = 0 \Leftrightarrow \forall k \in [\![1, n]\!],\ (y|e_k) = (x|e_k) \\
&\Leftrightarrow y = \sum_{k=1}^{n} (x|e_k)\, e_k \text{ (car } (e_k)_{1 \leqslant k \leqslant n} \text{ est une base orthonormée de } F).
\end{aligned}$$

On a montré que $\forall x \in E,\ \exists! y \in F /\ x - y \in F^\perp$ et donc $E = F \oplus F^\perp$ et on a montré au passage que $p_F(x) = \sum_{k=1}^{n} (x|e_k)\, e_k$.

Théorème 17. (Distance à un sous-espace de dimension finie)

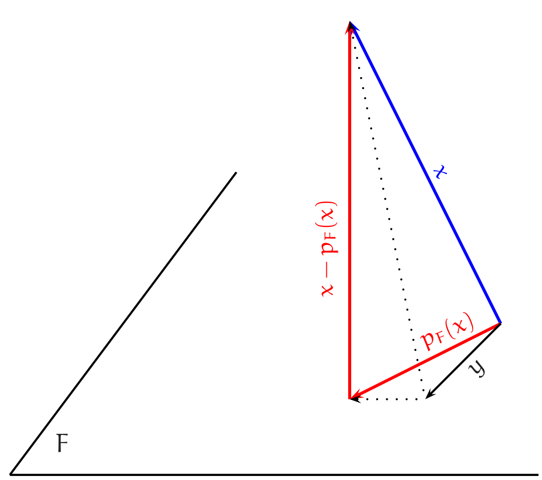
Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit F un sous-espace vectoriel de E de dimension finie. Soit $x \in E$.

$\forall y \in F,\ \|x - y\| \geqslant \|x - p_F(x)\|$ avec égalité si et seulement si $y = p_F(x)$.

Démonstration. Soient $x \in E$ et $y \in F$. D'après le théorème de Pythagore

$$\begin{aligned}
\|x - y\|^2 &= \left\| \left(\underbrace{x - p_F(x)}_{\in F^\perp} \right) + \left(\underbrace{p_F(x) - y}_{\in F} \right) \right\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \\
&\geqslant \|x - p_F(x)\|^2
\end{aligned}$$

avec égalité si et seulement si $y = p_F(x)$.



Le travail précédent montre que pour $x \in E$ donné

$$\mathrm{Inf}\{\|x-y\|,\ y \in F\} = \mathrm{Min}\{\|x-y\|,\ y \in F\} = \|x - p_F(x)\|\,.$$

Donc,

DÉFINITION 8. Soient $(E, (\ |\ ))$ un espace préhilbertien puis F un sous-espace vectoriel de E de dimension finie. Soit $x \in E$.

La **distance de** x **à** F est

$$d(x, F) = \mathrm{Inf}\{\|x-y\|,\ y \in F\} = \mathrm{Min}\{\|x-y\|,\ y \in F\} = \|x - p_F(x)\|\,.$$

Exercice 3. Soit $E = \mathbb{R}[X]$. Pour $(P, Q) \in E^2$, on pose $(P|Q) = \displaystyle\int_0^1 P(t)Q(t)\ dt$.

Déterminer la distance de X^3 à $\mathbb{R}_1[X]$ (voir exercice n° 2, page 11).

Solution 3. Puisque $\mathbb{R}_1[X]$ est un sous-espace de E de dimension finie, la projection orthogonale $P_{\mathbb{R}_1[X]}$ sur $\mathbb{R}_1[X]$ est définie.

D'après l'exercice n° 2, page 12, une base orthonormée de $\mathbb{R}_1[X]$ est (E_0, E_1) où $E_0 = 1$, $E_1 = \sqrt{3}(2X-1)$. Le projeté orthogonal de X^3 sur $\mathbb{R}_1[X]$ est

$$p_{\mathbb{R}_1[X]}\left(X^3\right) = \left(X^3|E_0\right)E_0 + \left(X^3|E_1\right)E_1$$

avec

- $\left(X^3|E_0\right) = \displaystyle\int_0^1 t^3\ dt = \frac{1}{4}$
- $\left(X^3|E_1\right) = \sqrt{3}\displaystyle\int_0^1 \left(2t^4 - t^3\right)\ dt = \sqrt{3}\left(\frac{2}{5} - \frac{1}{4}\right) = \frac{3\sqrt{3}}{20}$

puis

$$\begin{aligned} p_{\mathbb{R}_1[X]}\left(X^3\right) &= \frac{1}{4} + \frac{3\sqrt{3}}{20} \times \sqrt{3}(2X-1) = \frac{9X}{10} + \frac{1}{4} - \frac{9}{20} \\ &= \frac{9X}{10} - \frac{1}{5}. \end{aligned}$$

Par suite,

$$\begin{aligned} \left(d\left(X^3, \mathbb{R}_1[X]\right)\right)^2 &= \left\|X^3 - \frac{9X}{10} + \frac{1}{5}\right\|^2 = \int_0^1 \left(t^3 - \frac{9t}{10} + \frac{1}{5}\right)^2 dt = \int_0^1 \left(t^6 - \frac{9t^4}{5} + \frac{2t^3}{5} + \frac{81t^2}{100} - \frac{9t}{25} + \frac{1}{25}\right) dt \\ &= \frac{1}{7} - \frac{9}{25} + \frac{1}{10} + \frac{27}{100} - \frac{9}{50} + \frac{1}{25} = \frac{100 - 252 + 70 + 189 - 126 + 28}{700} \\ &= \frac{9}{700}. \end{aligned}$$

$$\boxed{d\left(X^3, \mathbb{R}_1[X]\right) = \frac{3}{10\sqrt{7}}.}$$

2) *Familles totales*

Théorème 18. (Inégalité de BESSEL)

Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit $(e_n)_{n \in \mathbb{N}}$ une famille orthonormale. Alors

$$\forall x \in E,\ \sum_{k=0}^{n} (x|e_k)^2 \leqslant \|x\|^2.$$

Démonstration. Soient $n \in \mathbb{N}$ puis $F = \mathrm{Vect}\,(e_k)_{0 \leqslant k \leqslant n}$. F est un sous-espace de E de dimension finie et $(e_k)_{0 \leqslant k \leqslant n}$ est une base orthonormée de F. Soit alors $x \in E$.

$$\|p_F(x)\|^2 = \left\|\sum_{k=0}^{n}(x|e_k)\,e_k\right\|^2 = \sum_{k=0}^{n}(x|e_k)^2\,.$$

Par suite, d'après le théorème de PYTHAGORE,

$$\|x\|^2 - \sum_{k=0}^{n}(x|e_k)^2 = \|x\|^2 - \|p_F(x)\|^2 = \|x - p_F(x)\|^2 = (d(x,F))^2 \geqslant 0.$$

Une conséquence immédiate du théorème 18 est

Théorème 19. (Inégalité de BESSEL)

Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit $(e_i)_{i\in I}$ une famille orthonormale. Soit $x \in E$. Alors, la famille $((x|e_i))_{i\in I}$ est de carré sommable et

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

On va maintenant donner une situation concernant la famille $(e_i)_{i\in I}$ où l'inégalité de BESSEL est en fait une égalité.

DÉFINITION 9. Soient $(E, (\ |\ ))$ un espace préhilbertien puis $(e_n)_{n\in\mathbb{N}}$ une famille orthonormale.

La famille $(e_n)_{n\in\mathbb{N}}$ une famille totale si et seulement si

$$\forall x \in E,\ \forall \varepsilon > 0,\ \exists y \in \text{Vect}\,(e_n)_{n\in\mathbb{N}} /\ \|x - y\| \leqslant \varepsilon.$$

Commentaire. La propriété précédente se réénoncera dans le cours de topologie en disant que $\text{Vect}\,(e_n)_{n\in\mathbb{N}}$ est dense dans l'espace vectoriel normé $(E, \|\ \|)$. ❑

On a alors le résultat suivant :

Théorème 20. (Egalité de PARSEVAL)

Soit $(E, (\ |\ ))$ un espace préhilbertien réel. Soit $(e_n)_{n\in\mathbb{N}}$ une famille totale. Soit $x \in E$.

$$\sum_{n=0}^{+\infty}(x|e_n)^2 = \|x\|^2.$$

Démonstration. Soit $x \in E$. On sait déjà que pour tout n de $\mathbb{N}$, $\displaystyle\sum_{k=0}^{n}(x|e_k)^2 \leqslant \|x\|^2$.

Pour $n \in \mathbb{N}$, posons $F_n = \text{Vect}\,(e_k)_{0\leqslant k\leqslant n}$.

Soit $\varepsilon > 0$. Il existe $y_0 \in \text{Vect}\,(e_n)_{n\in\mathbb{N}} /\ \|x - y_0\| \leqslant \sqrt{\varepsilon}$. y_0 est une combinaison linéaire d'un nombre fini des e_k et donc, il existe $n_0 \in \mathbb{N}$ tel que $y_0 \in F_{n_0}$. Mais alors, $\forall n \geqslant n_0$, $y_0 \in F_n$. Pour $n \geqslant n_0$, on a

$$\begin{aligned}
\varepsilon &\geqslant \|x - y_0\|^2 \geqslant (d\,(x, F_n))^2 = \|x - p_{F_n}(x)\|^2 \\
&= \|x\|^2 - \|p_{F_n}(x)\|^2 \ \text{(d'après le théorème de PYTHAGORE, } \|x\|^2 = \|p_{F_n}(x)\|^2 + \|x - p_{F_n}(x)\|^2\text{)} \\
&= \|x\|^2 - \sum_{k=0}^{n}(x|e_k)^2\,.
\end{aligned}$$

On a montré que $\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N}/\ \|x\|^2 - \varepsilon \leqslant \displaystyle\sum_{k=0}^{n}(x|e_k)^2 \leqslant \|x\|^2$ et donc que

$$\sum_{n=0}^{+\infty}(x|e_n)^2 = \lim_{n\to+\infty}\left(\sum_{k=0}^{n}(x|e_k)^2\right) = \|x\|^2.$$
