Chapitre 3

Espaces Préhilbertiens réels

Ce chapitre généralise aux espaces vectoriels réels les notions de produit scalaire et d'orthogonalité bien connues en dimensions 2 et 3. Dans tout le chapitre, E est un $\mathbb{R}$-espace vectoriel.

3.1 Premières définitions

3.1.1 Produits scalaires

Définition 3.1.1 (Produit scalaire)
Soit $f : \mathrm{E}^2 \longrightarrow \mathbb{R}$ une application. On dit qu'elle est :

- *bilinéaire* si, et seulement si,

$$\forall x, y, z \in \mathrm{E} \quad \forall \lambda \in \mathbb{K} \qquad f(\lambda x + y, z) = \lambda f(x, z) + f(y, z)$$

et

$$\forall x, y, z \in \mathrm{E} \quad \forall \lambda \in \mathbb{K} \qquad f(z, \lambda x + y) = \lambda f(z, x) + f(z, y)$$

- *symétrique* si, et seulement si,

$$\forall x, y \in \mathrm{E} \qquad f(x, y) = f(y, x)$$

- *définie* si, et seulement si,

$$\forall x \in \mathrm{E} \qquad f(x, x) = 0 \implies x = 0$$

- *positive* si, et seulement si,

$$\forall x \in \mathrm{E} \qquad f(x, x) \geqslant 0$$

- *non dégénérée* si, et seulement si,

$$\forall x \in \mathrm{E} \qquad (\forall y \in \mathrm{E} \quad f(x, y) = 0) \implies x = 0$$

Enfin, on appelle *produit scalaire sur* E toute application bilinéaire, symétrique, définie, positive. Un *espace préhilbertien réel* est un couple $(\mathrm{E}, \langle\ |\ \rangle)$ où E est un $\mathbb{R}$-espace vectoriel et $\langle\ |\ \rangle$ est un produit scalaire. Si E est de dimension finie, on dit qu'il est *euclidien*.

Observons que, pour montrer qu'une application est un produit scalaire, il est plus rapide d'établir d'abord la symétrie : en effet, suffit alors de montrer la linéarité par rapport à la première variable pour avoir la bilinéarité.

Exemple 3.1.2

1. Soit n un entier non nul. Si x est un vecteur dans $\mathbb{R}^n$, on notera $x_1, \ldots, x_n$ ses coordonnées dans la base canonique. On pose
$$\forall x, y \in \mathbb{R}^n \qquad \langle x \mid y\rangle = \sum_{k=1}^{n} x_k y_k = {}^t x\, y$$
Il est évident que $\langle \mid \rangle$ est un produit scalaire. En particulier, on utilise le fait qu'un carré est toujours positif dans $\mathbb{R}$ et qu'une somme de réels positifs est nulle si, et seulement si, chaque réel est nul.
Ce produit scalaire fait de $\mathbb{R}^n$ un espace préhilbertien réel. On l'appelle le produit scalaire canonique sur $\mathbb{R}^n$.
Dans le cas où $n = 2$ ou $n = 3$, on retrouve les produits scalaires usuels dans $\mathbb{R}^2$ et $\mathbb{R}^3$.
2. n est toujours un entier non nul. On se donne un $\mathbb{R}$-espace vectoriel de dimension finie E, avec une base $\mathscr{B}$ quelconque. Si $x \in \mathrm{E}$, on notera $x_1, \ldots, x_n$ ses coordonnées dans la base $\mathscr{B}$. Si l'on pose
$$\forall x, y \in \mathrm{E} \qquad \langle x \mid y\rangle = \sum_{k=1}^{n} x_k y_k$$
on définit un produit scalaire sur E.
Observons que ce produit scalaire dépend évidemment de la base choisie ; en général, si l'on change de base, on n'obtient pas le même produit scalaire. Également, si $\mathrm{E} = \mathbb{R}^n$ et $\mathscr{B}$ est la base canonique, alors on a simplement le produit scalaire canonique défini juste au-dessus.
3. Prenons par exemple $\mathrm{E} = \mathbb{R}^3$ et la base $\mathscr{B}$ formée par les vecteurs
$$e_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad e_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$
Il s'agit bien d'une base puisque
$$\det(e_1, e_2, e_3) = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{vmatrix} = 1 \neq 0$$
et un petit calcul montre que si $x \in \mathbb{R}^n$, alors ses coordonnées dans la base $\mathscr{B}$ sont
$$[x]_{\mathscr{B}} = \begin{bmatrix} 2x_1 - 2x_2 + x_3 \\ -x_1 + 2x_2 - x_3 \\ x_1 - y_1 + z_1 \end{bmatrix}$$
Si on note $\langle \mid \rangle_c$ le produit scalaire canonique et $\langle \mid \rangle_{\mathscr{B}}$ le produit scalaire dans la base $\mathscr{B}$, on a pour tous x et y dans $\mathbb{R}^n$:
$$\langle x \mid y\rangle_c = x_1 y_1 + x_2 y_2 + x_3 y_3$$
et $$\langle x \mid y\rangle_{\mathscr{B}} = 6x_1 y_1 - 7x_1 y_2 + 4x_1 y_3 - 7x_2 y_1 + 9x_2 y_2 - 5x_2 y_3 + 4x_3 y_1 - 5x_3 y_2 + 3x_3 y_3$$
Ceci montre clairement que $\langle \mid \rangle_c$ et $\langle \mid \rangle_{\mathscr{B}}$ n'ont aucune raison d'être identiques.
4. On note $\mathrm{E} = \mathscr{C}([0;1])$, le $\mathbb{R}$-espace vectoriel des fonctions continues sur $[0;1]$ et on pose
$$\forall f, g \in \mathrm{E} \qquad \langle f \mid g\rangle = \int_{[0;1]} f g$$
D'après les propriétés de l'intégrale, $\langle \mid \rangle$ est bien un produit scalaire sur E. En particulier, pour montrer qu'il est défini, on utilise le fait qu'une fonction *continue* sur $[0;1]$, qui a une intégrale nulle, est nulle. La continuité est essentielle ici.

Théorème 3.1.3 (Inégalité de Cauchy-Schwarz)
Soit f une forme bilinéaire symétrique positive sur E. *Alors*

$$\forall x, y \in \mathrm{E} \qquad f(x, y)^2 \leqslant f(x, x)\, f(y, y)$$

Si, de plus, f est un produit scalaire, et si $x, y \in \mathrm{E}$, l'inégalité de Cauchy-Schwarz est une égalité si, et seulement si, (x, y) est liée.

Preuve : Soient x et y dans E. Comme f est positive, $f(x+\lambda y, x+\lambda y)$ est positif pour tout λ réel. Mais par bilinérité et symétrie, on a

$$f(x+\lambda y, x+\lambda y) = f(x, y) + 2\lambda f(x, y) + \lambda^2 f(y, y)$$

Cette quantité est positive pour tout λ donc le trinôme $\mathrm{X}^2 f(y, y) + 2f(x, y)\mathrm{X} + f(x, x)$ a un discriminant négatif :

$$4f(x, y)^2 - 4f(x, x)\, f(y, y) \leqslant 0$$

Supposons maintenant que f est un produit scalaire. Soient x et y dans E et on suppose que $f(x, y)^2 = f(x, x)\, f(y, y)$. Alors le trinôme $\mathrm{X}^2 f(y, y) + 2f(x, y)\mathrm{X} + f(x, x)$ a une racine (double) donc il existe $\lambda \in \mathbb{R}$ tel que $f(x+\lambda y, x+\lambda y) = 0$. Comme f est définie, on sait que $x+\lambda y = 0$ donc (x, y) est liée. La réciproque est triviale, par simple calcul. $\square$

Proposition 3.1.4
Soit $f : \mathrm{E}^2 \longrightarrow \mathbb{R}$ une application. Si elle est définie, alors elle est non dégénérée.

Si f est bilinéaire, symétrique, positive, non dégénérée, alors elle est définie (et c'est donc un produit scalaire).

Preuve : Supposons f définie. Soit $x \in \mathrm{E}$ tel que

$$\forall y \in \mathrm{E} \qquad f(x, y) = 0$$

Alors en particulier $f(x, x) = 0$ donc $x = 0$.

Supposons maintenant que f est bilinéaire, symétrique, positive, non dégénérée. On se donne un $x \in \mathrm{E}$ tel que $f(x, x) = 0$. D'après l'inégalité de Cauchy-Schwarz, pour tout $y \in \mathrm{E}$,

$$f(x, y)^2 \leqslant f(x, x)\, f(y, y) = 0$$

Comme $f(x, y)$ est réel, il est nul. Mais y était quelconque ; comme f est non dégénérée, $x = 0$: f est définie. $\square$

3.1.2 Normes

Définition 3.1.5
Soit E un $\mathbb{R}$-espace vectoriel. Soit $\|\ \| : \mathrm{E} \longrightarrow \mathbb{R}_+$ une application. On dit que $\|\ \|$ est une *norme* si, et seulement si,

$$\forall x \in \mathrm{E} \qquad \|x\| = 0 \implies x = 0$$

$$\forall x \in \mathrm{E} \quad \forall \lambda \in \mathbb{R} \qquad \|\lambda x\| = |\lambda|\, \|x\|$$

et
$$\forall x, y \in \mathrm{E} \qquad \|x+y\| \leqslant \|x\| + \|y\|$$

Définition 3.1.6 (Norme euclidienne)
Soit $(\mathrm{E}, \langle\ |\ \rangle)$ un espace préhilbertien réel. On appelle *norme euclidienne* sur E, associée au produit scalaire $\langle\ |\ \rangle$, l'application $\|\ \|$ définie par

$$\forall x \in \mathrm{E} \qquad \|x\| = \sqrt{\langle x \mid x\rangle}$$

Proposition 3.1.7 (Propriétés des normes euclidiennes)
Soit $(\mathrm{E}, \langle\ |\ \rangle)$ *un espace préhilbertien réel.*

1. ***Positivité stricte*** $\quad \forall x \in \mathrm{E} \setminus \{0\} \qquad \|x\| > 0$
2. ***Homogénéité*** $\quad \forall x \in \mathrm{E} \quad \forall \lambda \in \mathbb{R} \qquad \|\lambda x\| = |\lambda|\, \|x\|$
3. ***Inégalité de Cauchy-Schwarz*** $\quad \forall x, y \in \mathrm{E} \qquad |\langle x \mid y\rangle| \leqslant \|x\|\, \|y\|$
   avec égalité si, et seulement si, (x, y) *est liée.*
4. ***Relation du parallélogramme*** $\quad \forall x, y \in \mathrm{E} \qquad \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$
5. ***Identités de polarisation*** $\quad \forall x, y \in \mathrm{E} \qquad \langle x \mid y\rangle = \dfrac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2) = \dfrac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$
6. ***Inégalité de Minkowski*** $\quad \forall x, y \in \mathrm{E} \qquad \big|\|x\| - \|y\|\big| \leqslant \|x \pm y\| \leqslant \|x\| + \|y\|$

En particulier, la norme euclidienne est une norme.

Preuve : La première propriété est une conséquence du fait que le produit scalaire est positif, défini. La deuxième est une récriture de l'inégalité de Cauchy-Schwarz.

Donnons-nous x et y dans E. En utilisant la bilinéarité et la symétrie du produit scalaire, on a

$$\|x+y\|^2 = \langle x+y \mid x+y\rangle = \langle x \mid x\rangle + 2\langle x \mid y\rangle + \langle y \mid y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x \mid y\rangle$$

ce qui fournit la première identité de polarisation. On a également

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x \mid y\rangle$$

et l'on en déduit la deuxième identité de polarisation et la relation du parallélogramme. Ensuite, d'après Cauchy-Schwarz, on sait que

$$\big|\langle x \mid y\rangle\big| \leqslant \|x\|\, \|y\|$$

d'où
$$\|x \pm y\|^2 \leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\, \|y\| = (\|x\| + \|y\|)^2$$

et
$$\|x \pm y\| \leqslant \|x\| + \|y\|$$

Enfin, cette relation étant vraie pour tous x et y, on a

$$\|x\| = \|x \pm y \mp y\| \leqslant \|x \pm y\| + \|y\|$$

donc
$$\|x\| - \|y\| \leqslant \|x \pm y\|$$

et de même,
$$\|y\| - \|x\| \leqslant \|x \pm y\|$$
□

3.2 Orthogonalité

3.2.1 Définitions et premières propriétés

Définition 3.2.1 (Orthogonalité)
Soit $(\mathrm{E}, \langle\ |\ \rangle)$ un espace préhilbertien réel. Soient x et y deux vecteurs de E. On dit qu'ils sont *orthogonaux* si, et seulement si, $\langle x \mid y\rangle = 0$. On notera alors $x \perp y$.

Soit $\mathrm{A} \subset \mathrm{E}$ une partie formée d'au moins deux éléments. On dit que c'est une *famille orthogonale* si, et seulement si,

$$\forall x, y \in \mathrm{A} \qquad x \neq y \implies x \perp y$$

Enfin, on dira que A est *orthonormée* si, et seulement si, elle est orthogonale et tous ses éléments ont pour norme 1.

Il est évident, parce qu'un produit scalaire est symétrique, que $x \perp y$ équivaut à $y \perp x$. Également, le seul vecteur orthogonal à lui-même est 0 car le produit scalaire est défini; enfin, le seul vecteur orthogonal à tous les vecteurs de E est aussi le vecteur nul car le produit scalaire est non dégénéré.

Proposition 3.2.2 (Théorème de Pythagore)
Soient $(\mathrm{E}, \langle\ |\ \rangle)$ *un espace préhilbertien réel, n un entier non nul et* $(x_1, \dots, x_n)$ *une famille orthogonale. Alors pour tous réels* $\lambda_1, \dots, \lambda_n$,

$$\left\| \sum_{k=1}^{n} \lambda_k x_k \right\|^2 = \sum_{k=1}^{n} \lambda_k^2 \|x_k\|^2$$

Preuve : C'est une conséquence immédiate de la bilinéarité du produit scalaire. Si $\lambda_1, \dots, \lambda_n$ sont dans $\mathbb{R}$, on a

$$\left\| \sum_{k=1}^{n} \lambda_k x_k \right\|^2 = \left\langle \sum_{k=1}^{n} \lambda_k x_k \;\middle|\; \sum_{\ell=1}^{n} \lambda_\ell x_\ell \right\rangle = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \lambda_k \lambda_\ell \underbrace{\langle x_k \mid x_\ell\rangle}_{=0 \text{ si } k \neq \ell} = \sum_{k=1}^{n} \lambda_k^2 \|x_k\|^2$$ □

Définition 3.2.3
Soient $(\mathrm{E}, \langle\ |\ \rangle)$ un espace préhilbertien réel et $\mathrm{A} \subset \mathrm{E}$. On appelle *orthogonal de* A, noté A°, l'ensemble des vecteurs orthogonaux à tous les vecteurs de A :

$$\mathrm{A}^\circ = \{x \in \mathrm{E} \mid \forall a \in \mathrm{A} \quad x \perp a\}$$

Proposition 3.2.4
Soient $(\mathrm{E}, \langle\ |\ \rangle)$ *un espace préhilbertien réel. On se donne* A *et* B *des sous-ensembles de* E *non vides.*

1. $\{0\}^\circ = \mathrm{E}$ *et* $\mathrm{E}^\circ = \{0\}$.
2. *Soient* $e_1, \dots, e_n$ *dans* E, *non nuls. Si la famille* $(e_1, \dots, e_n)$ *est orthogonale, alors elle est libre.*
3. $\mathrm{A}^\circ = (\mathit{Vect}\mathrm{A})^\circ$ *; de plus,* A° *est un sous-espace vectoriel de* E *et* $\mathrm{A} \subset \mathrm{A}^{\circ\circ}$.
4. $(\mathrm{A} \cup \mathrm{B})^\circ = (\mathit{Vect}\mathrm{A} + \mathit{Vect}\mathrm{B})^\circ = \mathrm{A}^\circ \cap \mathrm{B}^\circ$.
5. $\mathrm{A}^\circ + \mathrm{B}^\circ \subset (\mathit{Vect}\mathrm{A} \cap \mathit{Vect}\mathrm{B})^\circ$.
6. *Si* $\mathrm{A} \subset \mathrm{B}$, *alors* $\mathrm{B}^\circ \subset \mathrm{A}^\circ$.
7. *Si* F *est un sous-espace vectoriel de* E, *alors* F *et* F° *sont en somme directe.*

Preuve : Tout vecteur x de E est orthogonal au vecteur nul, puisque

$$2\langle x \mid 0\rangle = \langle x \mid 2\times 0\rangle = \langle x \mid 0\rangle$$

Par suite, $\{0\}^\circ = \mathrm{E}$. Le fait que $\mathrm{E}^\circ = \{0\}$ a déjà été observé plus haut.

Soient $e_1,\dots,e_n$ dans E, non nuls, tels que $(e_1,\dots,e_n)$ est une famille orthogonale. On suppose la famille liée ; alors l'un de ces vecteurs, par exemple e_n, est combinaison linéaire des autres. Il existe $\lambda_1,\dots,\lambda_{n-1}\in\mathbb{R}$ tels que

$$e_n = \sum_{k=1}^{n-1}\lambda_k e_k$$

Alors
$$\|e_n\|^2 = \langle e_n \mid e_n\rangle = \left\langle \sum_{k=1}^{n-1}\lambda_k e_k \;\middle|\; e_n\right\rangle = \sum_{k=1}^{n-1}\lambda\langle e_k \mid e_n\rangle = 0$$

Mais e_n n'est pas nul donc $\|e_n\|\neq 0$ et on a une contradiction.

Pour tout élément x de E, on note φ_x l'application

$$\varphi_x :\ \begin{aligned} \mathrm{E} &\longrightarrow \mathbb{R}\\ y &\longmapsto \langle x \mid y\rangle \end{aligned}$$

Le produit scalaire est linéaire par rapport à la seconde variable donc φ_x est une forme linéaire. Si A est une partie de E, on a

$$\mathrm{A}^\circ = \{y\in\mathrm{E} \mid \forall x\in\mathrm{A}\quad \langle x \mid y\rangle = 0\} = \bigcap_{x\in\mathrm{A}}\operatorname{Ker}\varphi_x$$

donc A° est un sous-espace vectoriel de E, car c'est une intersection de sous-espaces.

Si $x\in\mathrm{A}$, il est, par définition, orthogonal à tous les éléments de A° donc $x\in\mathrm{A}^{\circ\circ}$.

Si B est une autre partie de E avec $\mathrm{A}\subset\mathrm{B}$, on a

$$\mathrm{B}^\circ = \bigcap_{x\in\mathrm{B}}\operatorname{Ker}\varphi_x \subset \bigcap_{x\in\mathrm{A}}\operatorname{Ker}\varphi_x = \mathrm{A}^\circ$$

En particulier, comme $\mathrm{A}\subset\operatorname{Vect}\mathrm{A}$, on a $(\operatorname{Vect}\mathrm{A})^\circ\subset\mathrm{A}^\circ$. Réciproquement, si $x\in\mathrm{A}^\circ$, il est orthogonal à tout élément de A ; donc orthogonal à toute combinaison linéaire finie d'éléments de A, par bilinéarité du produit scalaire. Donc $\mathrm{A}^\circ = (\operatorname{Vect}\mathrm{A})^\circ$.

Il s'ensuit que, pour toutes parties non vides A et B,

$$(\mathrm{A}\cup\mathrm{B})^\circ = (\operatorname{Vect}(\mathrm{A}\cup\mathrm{B}))^\circ = (\operatorname{Vect}\mathrm{A}+\operatorname{Vect}\mathrm{B})^\circ$$

puisque $\operatorname{Vect}\mathrm{A}+\operatorname{Vect}\mathrm{B} = \operatorname{Vect}(\mathrm{A}\cup\mathrm{B})$. Si $x\in\mathrm{A}^\circ\cap\mathrm{B}^\circ$, il est orthogonal à tout élément de A et à tout élément de B : il est donc orthogonal à tout élément de $\mathrm{A}\cup\mathrm{B}$. On a bien $\mathrm{A}^\circ\cap\mathrm{B}^\circ\subset(\mathrm{A}\cup\mathrm{B})^\circ$. La réciproque est tout aussi évidente.

Enfin, on a $(\operatorname{Vect}\mathrm{A}\cap\operatorname{Vect}\mathrm{B})$ inclus dans $\operatorname{Vect}\mathrm{A}$ et dans $\operatorname{Vect}\mathrm{B}$ donc

$$\mathrm{A}^\circ = (\operatorname{Vect}\mathrm{A})^\circ\subset(\operatorname{Vect}\mathrm{A}\cap\operatorname{Vect}\mathrm{B})^\circ \qquad\text{et}\qquad \mathrm{B}^\circ = (\operatorname{Vect}\mathrm{B})^\circ\subset(\operatorname{Vect}\mathrm{A}\cap\operatorname{Vect}\mathrm{B})\circ$$

d'où
$$(\mathrm{A}^\circ\cup\mathrm{B}^\circ)\subset(\operatorname{Vect}\mathrm{A}\cap\operatorname{Vect}\mathrm{B})\circ$$

Mais $\mathrm{A}^\circ+\mathrm{B}^\circ$ est le plus petit sous-espace de E qui contient $\mathrm{A}^\circ\cup\mathrm{B}^\circ$. En particulier, il est inclus dans $(\operatorname{Vect}\mathrm{A}\cap\operatorname{Vect}\mathrm{B})\circ$. □

3.2.2 L'algorithme de Schmidt

L'algorithme de Schmidt est un outil fondamental pour construire une famille orthonormée à partir d'une famille libre finie.

Théorème 3.2.5 (Théorème de Schmidt)
Soient $(\mathrm{E}, \langle\,|\,\rangle)$ *un espace préhilbertien réel,* n *un entier non nul et* $(e_1, \ldots, e_n)$ *une famille libre de* E. *Alors il existe une famille orthonormée* $(v_1, \ldots, v_n)$ *telle que*

$$\forall k \in [\![1\,;n]\!] \qquad Vect(e_1, \ldots, e_k) = Vect(v_1, \ldots, v_k)$$

Preuve : On démontre ceci par récurrence. On définit pour tout entier n non nul la proposition $\mathscr{P}(n)$: « si E est un espace préhilbertien réel et $(e_1, \ldots, e_n)$ est une famille libre, il existe $(v_1, \ldots, v_n)$ orthonormée telle que

$$\forall k \in [\![1\,;n]\!] \qquad \mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(v_1, \ldots, v_k)$$ »

- $\mathscr{P}$**(1) est vraie** : Soit (e_1) une famille libre dans un espace préhilbertien réel. Alors e_1 n'est pas nul, et $\|e_1\| \neq 0$. On pose $v_1 = \frac{e_1}{\|e_1\|}$ et on a gagné.
- $\boldsymbol{\mathscr{P}(n) \Longrightarrow \mathscr{P}(n+1)}$: Soit n un entier non nul tel que $\mathscr{P}(n)$ est vraie. On se donne un espace préhilbertien réel E et une famille libre $(e_1, \ldots, e_{n+1})$ dans E. Alors $(e_1, \ldots, e_n)$ est libre : d'après $\mathscr{P}(n)$, on peut trouver $(v_1, \ldots, v_n)$ orthonormée telle que

  $$\forall k \in [\![1\,;n]\!] \qquad \mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(v_1, \ldots, v_k)$$

  Il reste donc seulement à trouver v_{n+1}. On se donne des réels $\lambda_1, \ldots, \lambda_n$ et on pose

  $$u = e_{n+1} + \sum_{k=1}^{n} \lambda_k v_k$$

  On cherche les bonnes valeurs pour $\lambda_1, \ldots, \lambda_n$, de manière à ce que u soit orthogonal à $v_1, \ldots, v_n$. Si $i \in [\![1\,;n]\!]$, on a par bilinéarité du produit scalaire

  $$\langle u \mid v_i \rangle = \langle u \mid e_{n+1} \rangle + \sum_{k=1}^{n} \lambda_k \underbrace{\langle v_k \mid v_i \rangle}_{=\delta_{i,j}} = \langle e_{n+1} \mid v_i \rangle + \lambda_i$$

  On voit donc que $u \perp v_i$ si, et seulement si, $\lambda_i = -\langle e_{n+1} | v_i \rangle$. On pose donc

  $$\forall i \in [\![1\,;n]\!] \qquad \lambda_i = -\langle e_{n+1} \mid v_i \rangle$$

  De cette manière, $(v_1, \ldots, v_n, u)$ est orthogonale. Elle est donc libre et en particulier $u \neq 0$. Il suffit donc de poser $v_{n+1} = \frac{u}{\|u\|}$ pour avoir $(v_1, \ldots, v_{n+1})$ orthonormée.
  De plus, on sait déjà que

  $$\forall k \in [\![1\,;n]\!] \qquad \mathrm{Vect}(e_1, \ldots, e_k) = \mathrm{Vect}(v_1, \ldots, v_k)$$

  En outre, par construction, $v_{n+1} \in \mathrm{Vect}(v_1, \ldots, v_n, e_{n+1}) = \mathrm{Vect}(e_1, \ldots, e_n, e_{n+1})$ donc

  $$\mathrm{Vect}(v_1, \ldots, v_{n+1}) \subset \mathrm{Vect}(e_1, \ldots, e_{n+1})$$

  Mais ces espaces sont de même dimension, égale à $n+1$, car les familles $(e_1, \ldots, e_{n+1})$ et $(v_1, \ldots, v_{n+1})$ sont libres. Donc ils sont égaux, ce qui démontre $\mathscr{P}(n+1)$.
- **Conclusion** : $\mathscr{P}(n)$ est vraie pour tout entier n non nul. $\square$

Exemple 3.2.6
La preuve du théorème de Schmidt a pour avantage de montrer précisément comment construire une base orthonormée à partir d'une famille libre.

Prenons par exemple comme espace $\mathrm{E} = \mathbb{R}[\mathrm{X}]$, avec le produit scalaire

$$\forall \mathrm{P}, \mathrm{Q} \in \mathrm{E} \qquad \langle \mathrm{P} \mid \mathrm{Q} \rangle = \int_{-1}^{1} \mathrm{P}(x)\mathrm{Q}(x)\,\mathrm{d}x$$

Observons déjà que c'est bien un produit scalaire ; toutes les propriétés de la définition sont triviales, sauf peut-être le fait que $\langle\ \mid\ \rangle$ est définie. Si P est tel que $\langle \mathrm{P} \mid \mathrm{P} \rangle = 0$, alors P est la fonction nulle sur $[-1\,;1]$. Mais un polynôme qui a une infinité de racines est nul, donc $\mathrm{P} = 0$.

On considère la famille $(1, \mathrm{X}, \mathrm{X}^2)$, qui est libre dans E et on lui applique le procédé de Schmidt pour l'orthonormaliser. On a

$$\|1\|^2 = \int_{-1}^{1} \mathrm{d}x = 2$$

et on pose
$$v_1 = \frac{1}{\|1\|} = \frac{\sqrt{2}}{2}$$

On a calcule alors

$$\langle v_1 \mid \mathrm{X} \rangle = \frac{\sqrt{2}}{2} \int_{-1}^{1} x\,\mathrm{d}x = 0$$

et on pose
$$u_2 = \mathrm{X} - \langle v_1 \mid \mathrm{X} \rangle\, v_1 = \mathrm{X}$$

qui est orthogonal à v_1. Il reste à le rendre normé donc calcule

$$\|u_2\|^2 = \int_{-1}^{1} x^2\,\mathrm{d}x = \frac{2}{3}$$

et on pose
$$v_2 = \frac{u_2}{\|u_2\|} = \frac{\sqrt{3}}{\sqrt{2}}\mathrm{X}$$

À ce stade, on a une famille orthonormée (v_1, v_2). Pour calculer le troisième vecteur, on commence par déterminer

$$\langle v_1 \mid \mathrm{X}^2 \rangle = \frac{\sqrt{2}}{2} \int_{-1}^{1} x^2\,\mathrm{d}x = \frac{\sqrt{2}}{3} \qquad \langle v_2 \mid \mathrm{X}^2 \rangle = \frac{\sqrt{3}}{\sqrt{2}} \int_{-1}^{1} x^3\,\mathrm{d}x = 0$$

On définit
$$u_3 = \mathrm{X}^2 - \langle v_2 \mid \mathrm{X}^2 \rangle\, v_2 - \langle v_1 \mid \mathrm{X}^2 \rangle\, v_1 = \mathrm{X}^2 - \frac{1}{3}$$

Par construction, (v_1, v_2, u_3) est orthogonale. Il reste à normer u_3 :

$$\|u_3\|^2 = \int_{-1}^{1} \left(x^2 - \frac{\sqrt{1}}{3}\right)^2 \mathrm{d}x = \frac{8}{45}$$

En posant
$$v_3 = \frac{3\sqrt{5}}{2\sqrt{2}}\left(\mathrm{X}^2 - \frac{1}{3}\right)$$

on a construit une base orthonormée (v_1, v_2, v_3), à partir de (e_1, e_2, e_3), qui conserve les sous-espaces intermédiaires. La famille

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{\sqrt{2}}\mathrm{X}, \frac{3\sqrt{5}}{2\sqrt{2}}\left(\mathrm{X}^2 - \frac{1}{3}\right)\right)$$

est une base orthonormée de $\mathbb{R}_2[\mathrm{X}]$ pour le produit scalaire $\langle\ \mid\ \rangle$.

Corollaire 3.2.7
Soient $(E, \langle \mid \rangle)$ *un espace préhilbertien réel et* F *un sous-espace de dimension finie, non nulle, de* E. *Alors* F *admet des bases orthonormées.*

De plus, si $n = \dim F$ *et* $(e_1, \dots, e_n)$ *est une base orthonormée de* F, *on a*

$$\forall x \in F \qquad x = \sum_{k=1}^{n} \langle x \mid e_k \rangle e_k$$

Preuve : Le théorème de Schmidt assure l'existence de bases orthonormées pour F. On en construit une, notée $(e_1, \dots, e_n)$. On se donne $x \in F$ et on considère

$$y = \sum_{k=1}^{n} \langle x \mid e_k \rangle e_k$$

Comme les vecteurs $(e_1, \dots, e_n)$ sont orthogonaux deux-à-deux, on voit que

$$\forall k \in [\![1\,;n]\!] \qquad \langle y \mid e_k \rangle = \langle x \mid e_k \rangle$$

ou encore
$$\forall k \in [\![1\,;n]\!] \qquad \langle y - x \mid e_k \rangle = 0$$

Alors par linéarité du produit scalaire par rapport à la deuxième variable, et parce que la famille $(e_1, \dots, e_n)$ engendre F, on a

$$\forall z \in F \qquad \langle y - x \mid z \rangle = 0$$

En particulier, $y - x \in F$ puisque y et x sont dans F d'où $\|y - x\|^2 = 0$ et $y = x$. □

3.2.3 Projection orthogonale sur un sous-espace

Il s'agit maintenant de généraliser la notion de projection orthogonale dont on a l'habitude dans le plan ou dans l'espace. Intuitivement, le projeté orthogonal d'un vecteur $\|x\|$ sur un sous-espace F doit être un élément de F qui minimise la distance entre x et les vecteurs de F.

Proposition 3.2.8
Soient $(E, \langle \mid \rangle)$ *un espace préhilbertien réel et* F *un sous-espace de dimension finie. Si* $x \in E$, *il existe un unique élément* $p_F x$ *de* F, *tel que*

$$\|x - p_F x\| = Inf\{\|x - z\| \mid z \in F\}$$

De plus, $x - p_F x$ *est orthogonal à* F *et*

$$\|p_F x\| \leqslant \|x\|$$

avec égalité si, et seulement si, $x \in F$.
Enfin, si F *est de dimension finie* n *non nulle et rapporté à une base orthonormée* $(e_1, \dots, e_n)$, *on a*

$$p_F x = \sum_{k=1}^{n} \langle x \mid e_k \rangle e_k$$

$p_F x$ *est appelé la projection orthogonale de* x *sur* F.

Preuve : Notons n la dimension de F. Le théorème est trivial si $n = 0$ donc on suppose n non nul. Comme F est de dimension finie, on peut y trouver une base orthonormée $(e_1, \dots, e_n)$. Si x est donné dans E, on définit

$$p_F x = \sum_{k=1}^{n} \langle x \mid e_k \rangle e_k$$

Par définition, $p_F x$ est dans F et parce que $(e_1,\dots,e_n)$ est orthonormée, on a

$$\forall k \in [\![1\,;n]\!] \qquad \langle p_F x \mid e_k\rangle = \langle x \mid e_k\rangle$$

On se donne maintenant un $z \in F$, quelconque. Toujours du fait que $(e_1,\dots,e_n)$ est orthonormée, on sait que

$$z = \sum_{k=1}^{n} \langle z \mid e_k\rangle e_k$$

et
$$\langle p_F x \mid z\rangle = \sum_{k=1}^{n} \langle x \mid e_k\rangle \langle z \mid e_k\rangle = \left\langle x \,\middle|\, \sum_{k=1}^{n} \langle z \mid e_k\rangle\, e_k \right\rangle = \langle x \mid z\rangle$$

D'où
$$\forall z \in F \qquad \langle x - p_F x \mid z\rangle = 0$$

Ceci démontre que $x - p_F x$ est orthogonal à F. Par conséquent, d'après le théorème de Pythagore,

$$\forall z \in F \qquad \|x-z\|^2 = \|\underbrace{x - p_F x}_{\in F^\circ} + \underbrace{p_F x - z}_{\in F}\|^2 = \|x - p_F x\|^2 + \|p_F x - z\|^2 \geqslant \|x - p_F x\|^2 \qquad (\star)$$

et l'on a bien
$$\|x - p_F x\| = \operatorname{Min}\{\|x-z\| \mid z \in F\}$$

Mais la relation $(\star)$ montre aussi que si $z \in F$ est tel que $\|x - z\| = \|x - p_F x\|$, alors $\|p_F x - z\| = 0$ donc $z = p_F x$. Par suite, $p_F x$ est bien l'unique vecteur de F qui réalise ce minimum.

Enfin, comme $x - p_F x$ est orthogonal à $p_F x$, on a

$$\|p_F x\|^2 = \|x\|^2 - \|x - p_F x\|^2 \leqslant \|x\|^2$$

ce qui fournit l'inégalité de Bessel. Et on voit qu'on a une égalité si, et seulement si, $\|x - p_F x\|^2 = 0$, c'est-à-dire $x = p_F x$, ou encore $x \in F$. □

Exemple 3.2.9
Supposons qu'on cherche à calculer

$$m = \operatorname{Inf}\left\{ \int_{-1}^{1} \left(e^x + a + bx + cx^2\right)^2 dx \,\middle|\, a,b,c \in \mathbb{R}\right\}$$

Une manière consiste à étudier la fonction définie par

$$\forall a,b,c \in \mathbb{R}^3 \qquad f(a,b,c) = \int_{-1}^{1} \left(e^x + a + bx + cx^2\right)^2 dx$$

en étudiant ses dérivées partielles et en cherchant où celles-ci s'annulent. Ce n'est pas très drôle.

On peut aussi donner une structure euclidienne à l'espace $E = \mathscr{C}([-1\,;1])$ en posant

$$\forall f,g \in E \qquad \langle f \mid g\rangle = \int_{[-1\,;1]} fg$$

Si on note

$$e_1 : x \longmapsto 1 \qquad e_2 : x \longmapsto x \qquad e_3 : x \longmapsto x^2$$

et $F = \operatorname{Vect}(e_1,e_2,e_3)$, alors on a simplement

$$m = \operatorname{Inf}\{\|\exp - f\|^2 \mid f \in F\}$$

Le théorème de projection répond exactement à cette question puisque $m = \|\exp - p_F \exp\|^2$. Il suffit donc de trouver une base orthonormée (v_1,v_2,v_3) de F (déjà calculée dans l'exemple précédent), calculer

$$\langle \exp \mid v_1\rangle = \frac{\sqrt{2}}{2}\left(e - \frac{1}{e}\right) \qquad \langle \exp \mid v_2\rangle = \frac{\sqrt{6}}{e} \qquad \langle \exp \mid v_3\rangle = \frac{\sqrt{5}}{\sqrt{2}}\left(e - \frac{7}{e}\right)$$

$$\|\exp\|^2 = \frac{1}{2}\Big(\mathrm{e}^2 - \frac{1}{e^2}\Big) \qquad \|p_{\mathrm{F}}\exp\|^2 = \langle \exp \mid v_1\rangle^2 + \langle \exp \mid v_2\rangle^2 + \langle \exp \mid v_3\rangle^2 = 3\Big(\mathrm{e}^2 - 12 + \frac{43}{\mathrm{e}^2}\Big)$$

d'où $$m = \|\exp - p_{\mathrm{F}}\exp\|^2 = \|\exp\|^2 - \|p_{\mathrm{F}}\exp\|^2 = \frac{1}{2}\Big(5\mathrm{e}^2 + 259 - \frac{72}{\mathrm{e}^2}\Big)$$

Proposition 3.2.10
Soient $(\mathrm{E}, \langle\ \mid\ \rangle)$ *un espace préhilbertien réel et* F *un sous-espace de dimension finie. On note* p_{F} *l'application de projection orthogonale de* E *sur* F. p_{F} *est la projection sur* F, *parallèlement à* F°. *De plus,*

$$\forall x, y \in \mathrm{E} \qquad \langle p_{\mathrm{F}}x \mid y\rangle = \langle x \mid p_{\mathrm{F}}y\rangle$$

Preuve : Soit n la dimension de F. Si $n = 0$, le théorème est trivial donc on suppose n non nul. Si on note $(e_1, \dots, e_n)$ une base orthonormée de F, on a

$$\forall x \in \mathrm{E} \qquad p_{\mathrm{F}}x = \sum_{k=1}^{n} \langle x \mid e_k\rangle\, e_k$$

La linéarité du produit scalaire par rapport à la première variable assure la linéarité de p_{F}.

Si $x \in \mathrm{F}$, on a $\|x - x\| = 0 = \mathrm{Inf}\{\|x - z\| \mid z \in \mathrm{F}\}$ donc $p_{\mathrm{F}}x = x$. Ceci montre à la fois que $\mathrm{F} = \mathrm{Im}\, p_{\mathrm{F}}$ et que

$$\forall x \in \mathrm{E} \qquad p_{\mathrm{F}}(\underbrace{p_{\mathrm{F}}x}_{\in \mathrm{F}}) = p_{\mathrm{F}}x$$

donc p_{F} est une projection, d'image F. Enfin, si $x \in \mathrm{E}$, on a

$$\begin{aligned} x \in \mathrm{Ker}\, p_{\mathrm{F}} &\iff \sum_{k=1}^{n} \langle x \mid e_k\rangle\, e_k = 0 \\ &\iff \forall k \in [\![1\,;n]\!] \qquad \langle x \mid e_k = 0\rangle \\ &\iff x \in \{e_1, \dots, e_n\}^\circ = (\mathrm{Vect}\,(e_1, \dots, e_n))^\circ = \mathrm{F}^\circ \end{aligned}$$

d'où $$\mathrm{Ker}\, p_{\mathrm{F}} = \mathrm{F}^\circ$$

Enfin, donnons-nous x et y dans E. On a

$$p_{\mathrm{F}}x = \sum_{k=1}^{n} \langle x \mid e_k\rangle e_k \qquad \text{donc} \qquad \langle p_{\mathrm{F}}x \mid y\rangle = \sum_{k=1}^{n} \langle x \mid e_k\rangle\, \langle e_k \mid y\rangle$$

et de même $$\langle x \mid p_{\mathrm{F}}y\rangle = \sum_{k=1}^{n} \langle y \mid e_k\rangle\, \langle x \mid e_k\rangle$$

Comme le produit scalaire est symétrique, et la multiplication dans $\mathbb{R}$ est commutative, il vient $\langle p_{\mathrm{F}}x \mid y\rangle = \langle x \mid p_{\mathrm{F}}y\rangle$. □

Corollaire 3.2.11
Soient $(\mathrm{E}, \langle\ \mid\ \rangle)$ *préhilbertien réel et* F *un sous-espace de* E, *de dimension finie. Alors*

$$\mathrm{F}^{\circ\circ} = \mathrm{F} \qquad \textit{et} \qquad \mathrm{F}^{\circ\circ\circ} = \mathrm{F}^\circ$$

Preuve : On sait déjà que $\mathrm{F} \subset (\mathrm{F}^\circ)^\circ$. Réciproquement, soit $x \in \mathrm{F}^{\circ\circ}$. On a

$$x = p_{\mathrm{F}}x + (x - p_{\mathrm{F}}x)$$

On sait que $x - p_{\mathrm{F}}x \in \mathrm{F}^\circ$ donc $x \perp (x - p_{\mathrm{F}}x)$; et $p_{\mathrm{F}}x \in \mathrm{F}$ donc $p_{\mathrm{F}}x \perp (x - p_{\mathrm{F}}x)$. Par suite,

$$\langle x \mid x - p_F x\rangle = 0 = \underbrace{\langle p_F x \mid x - p_F x\rangle}_{=0} + \langle x - p_F x \mid x - p_F x\rangle$$

et $$\|x - p_F x\|^2 = 0 \qquad \text{d'où} \qquad x - p_F x = 0$$

Ceci montre que $p_F x = x$ donc $x \in F$: on a bien $F^{\circ\circ} = F$. Et du coup, $F^{\circ\circ\circ} = F^\circ$. □

3.3 Espaces euclidiens

3.3.1 Résumé

Un espace euclidien est, par définition, un espace préhilbertien réel $(E, \langle \mid \rangle)$ qui est de dimension finie. Comme tous les sous-espaces de E sont, eux-mêmes, de dimension finie, les résultats obtenus précédemment ont une formulation qui peut être simplifiée. La proposition suivante est une conséquence triviale de ce qui a été fait plus haut.

Proposition 3.3.1
Soit $(E, \langle \mid \rangle)$ *un espace euclidien de dimension n non nulle.*

- E *admet des bases orthonormées.*
- *Si* $(e_1, \dots, e_n)$ *est une base orthonormée, alors*

$$\forall x \in E \qquad x = \sum_{k=1}^{n} \langle x \mid e_k\rangle\, e_k$$

- *Si* F *est un sous-espace non nul de* E*, alors* $F^{\circ\circ} = F$ *et* F° *est un supplémentaire de* F*, appelé le supplémentaire orthogonal de* F *dans* E*. La projection orthogonale* p_F *sur* F *existe et il s'agit de la projection sur* F *parallèlement à* F°*. Enfin, si* $(u_1, \dots, u_k)$ *est une base de* F*, on a*

$$\forall x \in E \qquad p_F x = \sum_{i=1}^{k} \langle x \mid u_i\rangle\, u_i$$

3.3.2 Automorphismes orthogonaux

Définition 3.3.2
Soit $(E, \langle \mid \rangle$ un espace euclidien. Soit $f : E \longrightarrow E$ une application. On dira que

- f *préserve les distances* si, et seulement si,

$$\forall x, y \in E \qquad \|f(x) - f(y)\| = \|x - y\|$$

- f *préserve le produit scalaire* si, et seulement si,

$$\forall x, y \in E \qquad \langle f(x) \mid f(y)\rangle = \langle x \mid y\rangle$$

- f *préserve la norme* si, et seulement si,

$$\forall x \in E \qquad \|f(x)\| = \|x\|$$

Le très joli résultat qui suit dit que, essentiellement, ces trois propriétés sont équivalentes.

Lemme 3.3.3
Soit $(\mathrm{E}, \langle\ |\ \rangle)$ *un espace euclidien de dimension* n *non nulle. Soit* f *une application de* E *dans* E. *Les assertions suivantes sont équivalentes :*

1. f *préserve les distances et* $f(0) = 0$ *;*
2. f *préserve le produit scalaire ;*
3. f *est linéaire et il existe une base orthonormée* $(e_1, \dots, e_n)$ *de* E *telle que* $(f(e_1), \dots, f(e_n))$ *soit une base orthonormée ;*
4. f *est linéaire et préserve la norme.*

De plus, si f *satisfait une de ces propriétés, c'est un automorphisme de* E.

Ce théorème est particulièrement joli et puissant ; en particulier parce que le seul fait de préserver les distances et le vecteur nul (assertion 1) ou le produit scalaire (assertion 2) suffit à assurer la linéarité (assertions 3 et 4).
Preuve : Supposons que f préserve la distance. Comme $f(0) = 0$, on a immédiatement que

$$\forall x \in \mathrm{E} \qquad \|f(x)\| = \|f(x) - f(0)\| = \|x - 0\| = \|x\|$$

Par suite, si x et y sont dans E, on a

$$\begin{aligned} \|x-y\|^2 &= \|f(x) - f(y)\|^2 = \|f(x)\|^2 + \|f(y)\|^2 - 2\langle f(x) \mid f(y)\rangle \\ &= \|x\|^2 + \|y\|^2 - 2\langle f(x) \mid f(y)\rangle \\ \|x-y\|^2 &= \|x-y\|^2 + 2\langle x \mid y\rangle - 2\langle f(x) \mid f(y)\rangle \end{aligned}$$

d'où
$$\forall x, y \in \mathrm{E} \qquad \langle f(x) \mid f(y)\rangle = \langle x \mid y\rangle$$

La deuxième assertion est vraie.

Supposons maintenant que f préserve le produit scalaire. Soit $(e_1, \dots, e_n)$ une base orthonormée de E. On a

$$\forall i, j \in [\![1\,;n]\!] \qquad \langle f(e_i) \mid f(e_j)\rangle = \langle e_i \mid e_j\rangle$$

Ainsi, $(f(e_1), \dots, f(e_n))$ est une famille orthonormée de E. Elle est donc libre ; mais comme E est de dimension n, c'est une base orthonormée de E.

Reste à montrer que f est linéaire. C'est assez simple : prenons x et y dans E et λ dans $\mathbb{R}$. Alors

$$\begin{aligned} \|f(\lambda x + y) - \lambda f(x) - f(y)\|^2 = {}& \|f(\lambda x + y)\|^2 + \|f(x)\|^2 + \|f(y)\|^2 \\ & -2\lambda\langle f(\lambda x + y) \mid f(x)\rangle - 2\langle f(\lambda x + y) \mid f(y)\rangle - 2\langle f(x) \mid f(y)\rangle \end{aligned}$$

Comme f préserve le produit scalaire, il vient

$$\begin{aligned} \|f(\lambda x + y) - \lambda f(x) - f(y)\|^2 &= \|\lambda x + y\|^2 + \|x\|^2 + \|y\|^2 - 2\lambda\langle \lambda x + y \mid x\rangle - 2\langle \lambda x + y \mid y\rangle - 2\langle x \mid y\rangle \\ &= \|(\lambda x + y) - \lambda x - y\|^2 = 0 \end{aligned}$$

Par suite, $f(\lambda x + y) - \lambda f(x) - f(y) = 0$: f est bien linéaire. La proposition 3 est vraie. En outre, elle transforme une base en une base, donc c'est un automorphisme de E.

Supposons que f est linéaire et qu'il existe une base orthonormée $(e_1, \dots, e_n)$ de E telle que $(f(e_1), \dots, f(e_n))$ soit orthonormée. Soit x un vecteur dans E. Comme $(e_1, \dots, e_n)$ est orthonormée, on a

$$x = \sum_{k=1}^{n} \langle x \mid e_k\rangle\, e_k$$

donc $$\|x\|^2 = \sum_{k=1}^{n} |\langle x \mid e_k\rangle|^2 \qquad \text{et} \qquad f(x) = \sum_{k=1}^{n} \langle x \mid e_k\rangle f(e_k)$$

Mais $(f(e_1),\dots,f(e_n))$ est aussi une base orthonormée d'où

$$\|f(x)\|^2 = \sum_{k=1}^{n} |\langle x \mid e_k\rangle|^2 = \|x\|^2$$

Ainsi, f préserve la norme et 4 est vraie.

Enfin, supposons que f est linéaire et préserve la norme. Immédiatement, $\|f(0)\| = \|0\|$ donc $f(0) = 0$. Et comme f est linéaire,

$$\forall x, y \in \mathrm{E} \qquad \|f(x) - f(y)\| = \|f(x-y)\| = \|x-y\|$$

La proposition 1 est vraie. □

Définition 3.3.4 (Automorphismes orthogonaux)
Soit $(\mathrm{E}, \langle \mid \rangle)$ un espace euclidien. Toute application de E dans E, qui vérifie une des propriétés équivalentes du **lemme 3.3**, est appelée *automorphisme orthogonal de* E.

L'ensemble des automorphismes orthogonaux de E est noté $\mathscr{O}(\mathrm{E})$.

Définition 3.3.5 (Matrices orthogonales)
Soient $n \in \mathbb{N}^\star$ et $\mathrm{M} \in \mathrm{M}_n(\mathbb{R})$. On dit que M est une matrice orthogonale si, et seulement si, ${}^t\mathrm{M}\mathrm{M} = \mathrm{I}_n$.

L'ensemble des matrices $n \times n$ orthogonales est noté $\mathrm{O}_n(\mathbb{R})$.

Proposition 3.3.6
Soit n un entier non nul. $\mathrm{O}_n(\mathbb{R})$ *est un sous-groupe de* $\mathrm{GL}_n(\mathbb{R})$. *Toute matrice orthogonale a pour déterminant 1 ou* -1.

Si $\mathrm{M} \in \mathrm{M}_n(\mathbb{R})$*, on note* $\mathrm{C}_1,\dots,\mathrm{C}_n$ *ses colonnes et* $\mathrm{L}_1,\dots,\mathrm{L}_n$ *ses lignes. Les assertions suivantes sont équivalentes :*

1. $\mathrm{M} \in \mathrm{O}_n(\mathbb{R})$ *;*
2. ${}^t\mathrm{M} \in \mathrm{O}_n(\mathbb{R})$ *;*
3. $\mathrm{C}_1,\dots,\mathrm{C}_n$ *est une base orthonormée de* $\mathbb{R}^n$ *pour le produit scalaire canonique ;*
4. ${}^t\mathrm{L}_1,\dots,{}^t\mathrm{L}_n$ *est une base orthonormée de* $\mathbb{R}^n$ *pour le produit scalaire canonique.*

Preuve : Si $\mathrm{M} \in \mathrm{O}_n(\mathbb{R})$, alors ${}^t\mathrm{M}\mathrm{M} = \mathrm{I}_n$ donc M est inversible à gauche et son inverse à gauche est ${}^t\mathrm{M}$. Mais on sait qu'inversibilité à gauche équivaut à inversibilité à droite donc $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$ et $\mathrm{M}\,{}^t\mathrm{M} = \mathrm{I}_n$. Ceci démontre en même temps que les assertions *1* et *2* sont équivalentes, et que $\mathrm{O}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$. En outre, on a

$$\det({}^t\mathrm{M}\mathrm{M}) = (\det{}^t\mathrm{M})(\det\mathrm{M}) = (\det\mathrm{M})^2$$

donc si M est orthogonale, son déterminant vaut 1 ou -1.

Montrons que $\mathrm{O}_n(\mathbb{R})$ est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Soient M et N deux matrices orthogonales. Par définition,

$${}^t\mathrm{M}\mathrm{M} = \mathrm{I}_n \qquad \text{et} \qquad {}^t\mathrm{N}\mathrm{N} = \mathrm{I}_n$$

Alors $${}^t(\mathrm{M}\mathrm{N})\mathrm{M}\mathrm{N} = ({}^t\mathrm{N}\,{}^t\mathrm{M})\mathrm{M}\mathrm{N} = {}^t\mathrm{N}({}^t\mathrm{M}\mathrm{M})\mathrm{N} = {}^t\mathrm{N}\mathrm{N} = \mathrm{I}_n$$

donc $\mathrm{O}_n(\mathbb{R})$ est stable par produit. Enfin, si M est orthogonale, on a vu que $\mathrm{M}^{-1} = {}^t\mathrm{M}$ est aussi orthogonale. Ainsi, $\mathrm{O}_n(\mathbb{R})$ est bien un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

On note $\langle\,|\,\rangle_c$ le produit scalaire canonique dans $\mathbb{R}^n$. On se donne $M \in O_n(\mathbb{R})$; ses colonnes sont netées $(C_1,\ldots,C_n)$:

$$\forall i \in [\![1\,;n]\!] \qquad C_i = \begin{bmatrix} m_{1,i} \\ \vdots \\ m_{n,i} \end{bmatrix}$$

Si i et j sont dans $[\![1\,;n]\!]$, on remarque que

$$\langle C_i \mid C_j\rangle_c = \sum_{k=1}^{n} m_{k,i}\, m_{k,j} = \sum_{k=1}^{n} ({}^t M)_{j,k} M_{k,i} = ({}^t M M)_{j,i}$$

Compte-tenu de ce calcul, on voit facilement que

$$\begin{aligned} M \in O_n(\mathbb{R}) &\iff {}^t MM = I_n \\ &\iff \forall i,j \in [\![1\,;n]\!] \quad ({}^t MM)_{j,i} = \delta_{i,j} \\ &\iff \forall i,j \in [\![1\,;n]\!] \quad \langle C_i \mid C_j\rangle_c = \delta_{i,j} \\ M \in O_n(\mathbb{R}) &\iff (C_1,\ldots,C_n) \text{ est une base orthonormée de } \mathbb{R}^n \end{aligned}$$

Ceci montre que *1* et *3* sont équivalentes. On procède de même pour *2* et *4*. □

Définition 3.3.7
Soit $n \in \mathbb{N}^\star$. L'ensemble des matrices orthogonales $n \times n$, dont le déterminant vaut 1, est appelé *groupe spécial orthogonal d'ordre* n et on le note $SO_n(\mathbb{R})$. C'est un sous-groupe de $O_n(\mathbb{R})$ et ses éléments sont appelés des *rotations*.

Théorème 3.3.8
Soient $(E, \langle\,|\,\rangle)$ *euclidien,* $f \in \mathscr{L}(E)$ *et* $\mathscr{B}$ *une base orthonormée de* E. *Alors* f *est un automorphisme orthogonal si, et seulement si, sa matrice dans la base* $\mathscr{B}$ *est othogonale.*

> Ce théorème peut être reformulé d'une manière plus abstraite, mais aussi plus claire. Si l'on note $n \neq 0$ la dimension de E et $\mathscr{B}$ une base quelconque de E, on sait qu'on a un isomorphisme d'anneaux
>
> $$\mathrm{Mat}_{\mathscr{B}} : \begin{array}{rcl} \mathscr{L}(E) & \longrightarrow & M_n(\mathbb{R}) \\ f & \longmapsto & \mathrm{Mat}_{\mathscr{B}} f \end{array}$$
>
> On sait déjà que $\mathrm{Mat}_{\mathscr{B}}$ réalise un isomorphisme de groupes entre $\mathscr{GL}(E)$ et $GL_n(\mathbb{R})$. Le théorème précédent nous dit aussi que, si $\mathscr{B}$ est orthonormée, alors $\mathrm{Mat}_{\mathscr{B}}$ est aussi un isomorphisme de groupes entre $\mathscr{O}(E)$ et $O_n(\mathbb{R})$.

Preuve : Notons $\mathscr{B} = (e_1,\ldots,e_n)$ notre base orthonormée et $M = \mathrm{Mat}_{\mathscr{B}} f$:

$$M = \begin{bmatrix} m_{1,1} & \cdots & m_{1,n} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,n} \end{bmatrix}$$

Par définition, si $i \in [\![1\,;n]\!]$, la i-ème colonne de M donne les coordonnées de $f(e_i)$ dans la base $\mathscr{B}$:

$$\forall i \in [\![1\,;n]\!] \qquad f(e_i) = \sum_{k=1}^{n} m_{k,i}\, e_k$$

Comme la base $\mathscr{B}$ est orthonormée, si $i, j \in [\![1\,;n]\!]$, on a

$$\langle f(e_i) \mid f(e_j)\rangle = \sum_{k=1}^{n} m_{k,i} m_{k,j} = \langle \mathrm{C}_i \mid \mathrm{C}_j\rangle_c$$

On voit que $(f(e_1),\dots,f(e_n))$ est orthonormée si, et seulement si, $(\mathrm{C}_1,\dots,\mathrm{C}_n)$ est orthonormée dans $\mathbb{R}^n$, ce qui équivaut à dire (**proposition 3.6**) que M est une matrice orthogonale. $\square$

Proposition 3.3.9
Soient $(\mathrm{E}, \langle\ |\ \rangle)$ *un espace euclidien et* $\mathscr{B}$ *une base orthonormée. Soit* $\mathscr{B}'$ *une autre base de* E. *Alors* $\mathscr{B}'$ *est orthonormée si, et seulement si, la matrice de passage* $\mathrm{P}_{\mathscr{B}}^{\mathscr{B}'}$ *est orthogonale. Dans ce cas,* $\det_{\mathscr{B}}(\mathscr{B}')$ *vaut 1 ou* -1.

Preuve : Notons $\mathscr{B} = (e_1,\dots,e_n)$ et $\mathscr{B}' = (e'_1,\dots,e'_n)$. La matrice de passage $\mathrm{P}_{\mathscr{B}}^{\mathscr{B}'}$ est aussi la matrice, par rapport à la base $\mathscr{B}$, de l'application linéaire f qui transforme chaque e_i en e'_i. En utilisant le **lemme 3.3** et le **théorème 3.7**, on a

$$\begin{aligned}(e'_1,\dots,e'_n) \text{ orthonormée} &\iff (f(e_1),\dots,f(e_n)) \text{ orthonormée}\\ &\iff f \in \mathscr{O}(\mathrm{E})\\ &\iff \mathrm{P}_{\mathscr{B}}^{\mathscr{B}'} = \mathrm{Mat}_{\mathscr{B}} f \in \mathrm{O}_n(\mathbb{R})\end{aligned}$$

Si l'on suppose que $\mathscr{B}'$ est orthonormée, alors $\det_{\mathscr{B}}\mathscr{B}'$ est le déterminant de la matrice de passage $\mathrm{P}_{\mathscr{B}}^{\mathscr{B}'}$, qui vaut 1 ou -1 puisqu'il s'agit d'une matrice orthogonale. $\square$

On peut voir immédiatement un intérêt de travailler avec des bases orthonormées. En effet, lors d'un changement de base orthonormée, la matrice de passage P est orthogonale : son inverse est donc ${}^t\mathrm{P}$, ce qui rend son calcul absolument trivial.

Ainsi, supposons que f est un endomorphisme d'un espace euclidien, $\mathscr{B}$ et $\mathscr{B}'$ sont deux bases orthonormées. On note

$$\mathrm{M} = \mathrm{Mat}_{\mathscr{B}} f \qquad \mathrm{M}' = \mathrm{Mat}_{\mathscr{B}'} f \qquad \mathrm{P} = \mathrm{P}_{\mathscr{B}}^{\mathscr{B}'}$$

On a tout simplement $\mathrm{M}' = {}^t\mathrm{PMP}$ et $\mathrm{M} = \mathrm{PM}'\,{}^t\mathrm{P}$.

Définition 3.3.10 (Orientation d'une base)
Soient $(\mathrm{E}, \langle\ |\ \rangle)$ un espace euclidien et $\mathscr{B}$, $\mathscr{B}'$ deux bases de E. On dit qu'elles *ont la même orientation* si, et seulement si, $\det_{\mathscr{B}}\mathscr{B}' = 1$.

Dans la mesure où $\det_{\mathscr{B}}\mathscr{B}' = (\det_{\mathscr{B}'}\mathscr{B})^{-1}$ (voir cours sur les déterminants), on voit que $\mathscr{B}$ et $\mathscr{B}'$ ont la même orientation si, et seulement si, $\mathscr{B}'$ et $\mathscr{B}$ ont la même orientation.

Et si l'on a trois bases orthonormées $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$, telles que $\mathscr{B}$ et $\mathscr{B}'$ ont la même orientation, et $\mathscr{B}'$ et $\mathscr{B}''$ aussi, alors (cours sur les déterminants) :

$$\det_{\mathscr{B}}\mathscr{B}'' = (\det_{\mathscr{B}}\mathscr{B}')(\det_{\mathscr{B}'}\mathscr{B}'') = 1$$

On a donc montré que

Proposition 3.3.11
« Avoir la même orientation » est une relation d'équivalence sur l'ensemble des bases orthonormées d'un espace euclidien E.

Définition 3.3.12 (Orientation de l'espace)
Soit $(\mathrm{E}, \langle\ |\ \rangle)$ un espace euclidien. On dit qu'*on a orienté* E si on a choisi une classe d'équivalence C pour la relation « avoir la même orientation ». Dans ce cas, les bases de la classe C sont dites

directes et celles qui ne sont pas dans C sont dites *indirectes*.

Exemple 3.3.13
Dans l'espace euclidien $\mathbb{R}^3$, on choisit l'orientation donnée par la base canonique (e_1, e_2, e_3). Considérons les vecteurs

$$v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix} \qquad v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\0\\1\end{bmatrix}$$

On a déjà $v_1 \perp v_2$ et ces deux vecteurs sont orthogonaux à e_2. Donc on peut former les deux bases orthonormées (v_1, v_2, e_2) et $(v_1, v_2, -e_2)$. Il est certain qu'une des deux est directe, puisque

$$\det(v_1, v_2, -e_2) = -\det(v_1, v_2, e_2)$$

On calcule $$\det(v_1, v_2, e_2) = \begin{vmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\end{vmatrix} = \frac{1}{2}\begin{vmatrix}1 & -1 & 0\\ 0 & 0 & 1\\ 1 & 1 & 0\end{vmatrix} = -1$$

Ainsi, pour notre choix d'orientation de l'espace, la base orthonormée (v_1, v_2, e_2) est indirecte, tandis que $(v_1, v_2, -e_2)$ est directe. Observons que d'après la propriété d'antisymétrie du déterminant, (v_1, e_2, v_2) est directe, par exemple.

3.3.3 Symétries orthogonales et réflexions

On a déjà vu dans le cours sur les espaces vectoriels que si $E = F \oplus G$ est une décomposition de E en deux sous-espaces supplémentaires, on peut définir la projection p_F sur F parallèlement à G, et la symétrie s_F par rapport à F, parallèlement à G.

Dans le cas où E est euclidien et F est un sous-espace vectoriel, on a une décomposition privilégiée $E = F \oplus F^\circ$, et donc une manière privilégiée de projeter sur F, ou de symétriser par rapport à F. On a d'ailleurs vu (**théorème 2.10**) que la projection orthogonale sur F est précisément la projection sur F parallèlement à F°.

On suppose dans ce paragraphe que $(E, \langle\,|\,\rangle)$ est euclidien de dimension $n \geqslant 2$.

Définition 3.3.14 (Symétrie orthogonale)
Soit F un sous-espace de E. On appelle *symétrie orthogonale* par rapport à F, notée s_F, la symétrie par rapport à F, parallèlement à F°.

Autrement dit, la symétrie orthogonale s_F, par rapport à F, est la symétrie associée à la somme directe orthogonale $E = F \oplus F^\circ$. Rappelons ce que cela signifie : si $x \in E$, il se décompose suivant cette somme directe

$$x = \underbrace{p_F(x)}_{\in F} + \underbrace{p_{F^\circ}(x)}_{\in F^\circ}$$

Par définition, $s_F(x)$ garde la composante de x suivant F, et change la composante suivant F° en son opposé :

$$s_F(x) = p_F(x) - p_{F^\circ}(x) = 2p_F(x) - \mathrm{id}(x)$$

Proposition 3.3.15
Toute symétrie orthogonale est un automorphisme orthogonal de E*. Réciproquement, toute symétrie de* E*, qui est un automorphisme orthogonal de* E*, est une symétrie orthogonale.*

Preuve : Soit F un sous-espace de E et s_F la symétrie orthogonale par rapport à F. Si $x \in E$, on a

$$x = p_F(x) + p_{F^\circ}(x) \qquad s_F(x) = p_F(x) - p_{F^\circ}(x) \qquad \text{avec} \qquad p_F(x) \perp p_{F^\circ}$$

donc
$$\|s_F(x)\|^2 = \|p_F(x)\|^2 + \|p_{F^\circ}(x)\|^2 = \|x\|^2$$

d'après le théorème de Pythagore : s_F préserve la norme. Comme elle est aussi linéaire, c'est un automorphisme orthogonal de E.

Réciproquement, soit s une symétrie de E, qui soit aussi un automorphisme orthogonal de E. Alors il existe des sous-espaces supplémentaires F et G, tels que s soit la symétrie par rapport à F, parallèlement à G. Montrons que F et G sont orthogonaux : soient $x \in F$ et $y \in G$. Par définition de s,

$$s(x) = x \qquad s(y) = -y$$

donc
$$x + y = s(x) - s(y) = s(x - y)$$

Mais s est un automorphisme orthogonal donc préserve la norme et

$$\|x + y\| = \|s(x - y)\| = \|x - y\|$$

Ainsi,
$$4\langle x \mid y\rangle = \|x + y\|^2 - \|x - y\|^2 = 0$$

ce qui prouve bien que $x \perp y$. Par suite, s est une symétrie orthogonale. □

Définition 3.3.16 (Réflexion)
Soit $f \in \mathscr{L}(E)$. On dit que f est une *réflexion* si, et seulement si, f est la symétrie orthogonale par rapport à un hyperplan.

Proposition 3.3.17
Toute réflexion de E *est un automorphisme orthogonal de* E*, de déterminant* -1.

Preuve : Soit r une réflexion de E. Par définition, r est la symétrie orthogonale par rapport à un hyperplan H de E. D'après la **proposition 3.15**, c'est un automorphisme orthogonal de E.

Notons $n \geqslant 2$ la dimension de E. On sait que r est la symétrie orthogonale par rapport à un hyperplan H. Donc H est de dimension $n - 1$ et H° est une droite, de dimension 1. On prend une base $(e_1, \ldots, e_{n-1})$ de H et une base (e_n) de H°. Alors $\mathscr{B} = (e_1, \ldots, e_{n-1}, e_n)$ est une base de E. De plus, par définition de r comme symétrie par rapport à H, parallèlement à H°, on sait que

$$\forall x \in H \qquad r(x) = x \qquad \forall x \in H^\circ \qquad r(x) = -x$$

Par suite,
$$\operatorname{Mat}_{\mathscr{B}} r = \begin{bmatrix} I_{n-1} & 0 \\ 0 & -1 \end{bmatrix}$$

et le déterminant de r vaut bien -1. □

Théorème 3.3.18 (Décomposition en produit de réflexions)
Soit $(E, \langle\ |\ \rangle)$ *un espace euclidien de dimension* $n \geqslant 2$*. Tout automorphisme orthogonal de* E *peut être décomposé en produit d'au plus* n *réflexions.*

Preuve : On démontre ceci par récurrence sur la dimension de E. Le cas des espaces de dimension 2 sera traité dans le prochain paragraphe et on l'admet pour l'instant.

Soit $n \geqslant 2$ un entier. On suppose que tout automorphisme orthogonal d'un espace euclidien de dimension n peut être décomposé en produit d'au plus n réflexions. Soient E un espace euclidien de dimension $n+1$ et $f \in \mathscr{O}(\mathrm{E})$. On se donne une base orthonormée $\mathscr{B} = (e_1, \dots, e_n, e_{n+1})$ de E et on distingue deux cas :

- **Si $f(e_{n+1}) = e_{n+1}$** : Comme f est orthogonal et que $e_1, \dots, e_n$ sont orthogonaux à e_{n+1}, il vient que $f(e_1), \dots, f(e_{n+1})$ sont aussi orthogonaux à e_{n+1}. Donc

$$f\big(\mathrm{Vect}\,(e_1, \dots, e_n)\big) \subset \{e_{n+1}\}^\circ = \mathrm{Vect}\,(e_1, \dots, e_n)$$

Ainsi, f induit un endomorphisme de $\mathrm{F} = \mathrm{Vect}\,(e_1, \dots, e_n)$; plus précisément, si l'on définit

$$\forall x \in \mathrm{F} \qquad \tilde{f}(x) = f(x)$$

alors $\tilde{f}$ est un endomorphisme de F. En outre, $\tilde{f} \in \mathscr{O}(\mathrm{F})$ puisque

$$\forall x, y \in \mathrm{F} \qquad \langle \tilde{f}(x) \mid \tilde{f}(y) \rangle = \langle f(x) \mid f(y) \rangle = \langle x \mid y \rangle$$

D'après l'hypothèse de récurrence, il existe des réflexions $\tilde{r}_1, \dots, \tilde{r}_k$ de F, avec $k \leqslant n$, telles que $\tilde{f} = \tilde{r}_1 \cdots \tilde{r}_k$.

Pour chaque $j \in [\![1 ; k]\!]$, on note r_j l'endomorphisme de E qui vaut $\tilde{r}_j$ sur F et qui envoie e_{n+1} sur e_{n+1}. Alors r_j est une symétrie puisque $r_j^2 = \mathrm{id}$ sur F et sur $\mathrm{Vect}\, e_{n+1}$. De plus, si $x \in \mathrm{E}$, il se décompose de manière unique en $x = x_\mathrm{F} + y$ avec $y = \langle x \mid e_n \rangle\, e_n$ orthogonal à F ; donc

$$r_j(x) = r_j(x_\mathrm{F}) + r_j(y) = \underbrace{\tilde{r}_j(x_\mathrm{F})}_{\in \mathrm{F}} + y$$

et $$\|r_j(x)\|^2 = \|\tilde{r}_j(x_\mathrm{F})\|^2 + \|y\|^2 = \|x_\mathrm{F}\|^2 + \|y\|^2 = \|x\|^2$$

ce qui démontre que $r_j \in \mathscr{O}(\mathrm{E})$. D'après la proposition **3.15**, r_j est une symétrie orthogonale. Enfin, il est facile de voir que r_j est une réflexion puisque

$$\mathrm{Ker}\,(\tilde{r}_j - \mathrm{id}) \subset \mathrm{Ker}\,(r_j - \mathrm{id}) \qquad \text{et} \qquad e_{n+1} \in \mathrm{Ker}\,(r_j - \mathrm{id})$$

donc $$\mathrm{Ker}\,(\tilde{r}_j - \mathrm{id}) \oplus \mathrm{Vect}\,(e_{n+1}) \subset \mathrm{Ker}\,(r_j - \mathrm{id})$$

Mais $$\dim \mathrm{Ker}\,(\tilde{r}_j - \mathrm{id}) = n - 1 \qquad \text{donc} \qquad \dim \mathrm{Ker}\,(r_j - \mathrm{id}) \geqslant n$$

et cette dimension n'est pas $n+1$, puisque r_j est une réflexion sur F, donc a au moins un vecteur qui n'est pas invariant. Ce qui démontre bien que r_j est une réflexion.

Mais on a $f = r_1 \cdots r_k$ puisque cette relation est vérifiée sur F et sur $\mathrm{Vect}\,(e_{n+1})$. Donc f peut être décomposé en produit d'au plus n réflexions.

- **Si $f(e_{n+1}) \neq e_{n+1}$** : Dans ce cas, on remarque que

$$e_{n+1} = \frac{1}{2}(e_{n+1} - f(e_{n+1})) + \frac{1}{2}(e_{n+1} + f(e_{n+1}))$$

et $$\langle e_{n+1} - f(e_{n+1}) \mid e_{n+1} + f(e_{n+1}) \rangle = \|e_{n+1}\|^2 - \|f(e_{n+1})\|^2 = 0$$

parce que $f \in \mathscr{O}(\mathrm{E})$ et préserve la norme. Si on note g la réflexion autour de l'hyperplan $(e_{n+1} - f(e_{n+1}))^\circ$, on a par définition

$$g(e_{n+1} - f(e_{n+1})) = -(e_{n+1} - f(e_{n+1})) = f(e_{n+1}) - e_{n+1}$$

et $\qquad g(e_{n+1} + f(e_{n+1})) = e_{n+1} + f(e_{n+1}) \qquad$ car $\qquad (e_{n+1} + f(e_{n+1})) \perp (e_{n+1} - f(e_{n+1}))$

Ainsi,
$$g(e_{n+1}) = f(e_{n+1})$$

et
$$gf(e_{n+1}) = g^2(e_{n+1}) = e_{n+1}$$

car g est une symétrie. Donc $gf \in \mathscr{O}(\mathrm{E})$ et il fixe e_{n+1}. D'après le premier cas étudié, gf se décompose en produit d'au plus n réflexions. Mais $f = g(gf)$ se décompose alors en produit d'au plus $n+1$ réflexions. Ce qui achève la récurrence. ☐

La preuve de ce théorème explique comment faire, en pratique, pour décomposer un automorphisme orthogonal f en produit de réflexions. On commence par se donner une base orthonormée $(e_1, \dots, e_n)$ de l'espace. On note g_n la réflexion autour de l'hyperplan orthogonal à $e_n - f(e_n)$. Alors $g_n f$ est un automorphisme orthogonal qui fixe e_n et induit un automorphisme orthogonal de $\mathrm{Vect}(e_1, \dots, e_{n-1})$. On recommence alors la même procédure : on note g_{n-1} la réflexion autour de l'hyperplan $e_{n-1} - g_n f(e_{n-1})$. Alors $g_{n-1} g_n f$ est un automorphisme orthogonal qui fixe e_{n-1} et e_n. Et ainsi de suite.

Exemple 3.3.19
On considère la matrice

$$\mathrm{M} = \frac{1}{3}\begin{bmatrix} -2 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & -2 \end{bmatrix} \in \mathrm{O}_3(\mathbb{R})$$

M représente un automorphisme orthogonal de $\mathbb{R}^3$ dans la base canonique ; on identifie l'automorphisme et sa matrice. Posons

$$u_3 = e_3 - \mathrm{M}e_3 = \frac{1}{3}\begin{bmatrix} -1 \\ -2 \\ 5 \end{bmatrix}$$

La projection orthogonale P_3 sur $\mathrm{Vect}\, u_3$ est donnée par

$$\forall x \in \mathbb{R}^3 \qquad \mathrm{P}_3(x) = \left\langle x \,\middle|\, \frac{u_3}{\|u_3\|} \right\rangle \frac{u_3}{\|u_3\|} = \frac{\langle x \mid u_3\rangle}{\|u_3\|^2}\, u_3$$

et la réflexion autour de $(u_3)^\circ$ est donnée par

$$\forall x \in \mathbb{R}^3 \qquad \mathrm{S}_3 = \mathrm{I}_3 - 2\mathrm{P}_3$$

Après calculs, $\qquad \mathrm{P}_3 = \dfrac{1}{30}\begin{bmatrix} 1 & 2 & -5 \\ 2 & 4 & -10 \\ -5 & -10 & 25 \end{bmatrix} \qquad \mathrm{S}_3 = \mathrm{I}_3 - 2\mathrm{P}_3 = \dfrac{1}{15}\begin{bmatrix} 14 & -2 & 5 \\ -2 & 11 & 10 \\ 5 & 10 & -10 \end{bmatrix}$

Également,
$$\mathrm{S}_3\mathrm{M} = \frac{1}{5}\begin{bmatrix} -4 & 4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

On voit que, comme prévu, $\mathrm{S}_3\mathrm{M}$ induit un automorphisme orthogonal de $\mathrm{Vect}(e_1, e_2)$. On recommence, en travaillant cette fois avec e_2. Posons

$$u_2 = e_2 - \mathrm{S}_3\mathrm{M}e_2 = \frac{1}{5}\begin{bmatrix}-4\\2\\0\end{bmatrix}$$

La projection orthogonale P_2 sur $\mathrm{Vect}\, u_2$ et la réflexion S_2 autour de $(u_2)^\perp$ sont données par

$$\forall x \in \mathbb{R}^3 \qquad \mathrm{P}_2(x) = \frac{\langle x \mid u_2\rangle}{\| u_2\|^2}\, u_2 \qquad \mathrm{S}_2 = \mathrm{I}_3 - 2\mathrm{P}_2$$

Tous calculs faits, $\qquad \mathrm{P}_2 = \dfrac{1}{5}\begin{bmatrix}4 & -2 & 0\\ -2 & 1 & 0\\ 0 & 0 & 0\end{bmatrix} \qquad \mathrm{S}_2 = \mathrm{I}_3 - 2\mathrm{P}_2 = \dfrac{1}{5}\begin{bmatrix}-3 & 4 & 0\\ 4 & 3 & 0\\ 0 & 0 & 5\end{bmatrix}$

On remarque que $\mathrm{S}_2 = \mathrm{S}_3\mathrm{M}$, donc $\mathrm{S}_3\mathrm{M}$ était déjà une réflexion. Du coup, une décomposition de M en produit de réflexions est

$$\mathrm{M} = \mathrm{S}_3\mathrm{S}_2 = \frac{1}{15}\begin{bmatrix}14 & -2 & 5\\ -2 & 11 & 10\\ 5 & 10 & -10\end{bmatrix} \times \frac{1}{5}\begin{bmatrix}-3 & 4 & 0\\ 4 & 3 & 0\\ 0 & 0 & 5\end{bmatrix}$$

Cette procédure nous a même permis d'identifier géométriquement ces transformations : appliquer M, c'est la même chose que

- faire une réflexion autour du plan $(u_2)^\circ$, d'équation $-2x + y = 0$;
- suivie d'une réflexion autour du plan $(u_3)^\circ$, d'équation $x + 2y - 5z = 0$.

3.4 Automorphismes orthogonaux en dimension 2

L'étude du groupe orthogonal en dimension 2 est très simple. Étant donné un réel θ, on note

$$\mathrm{R}(\theta) = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix} \qquad \mathrm{S}(\theta) = \begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}$$

On peut remarquer que ces matrices sont orthogonales, de déterminants respectifs 1 et −1. Par définition, $\mathrm{R}(\theta)$ est une rotation de $\mathbb{R}^2$. Un simple dessin montre qu'il s'agit effectivement de la rotation d'angle θ.

Identifions la nature géométrique de $\mathrm{S}(\theta)$. À tout hasard, on se demande s'il ne s'agit pas, peut-être, d'une réflexion. Auquel cas, il suffit de trouver $\mathrm{Ker}\,(\mathrm{S}(\theta) - \mathrm{I}_2)$ et $\mathrm{Ker}\,(\mathrm{S}(\theta) + \mathrm{I}_2)$. Un calcul facile prouve que

$$\mathrm{Ker}\,(\mathrm{S}(\theta) - \mathrm{I}_2) = \mathrm{Vect}\left(\begin{bmatrix}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}\end{bmatrix}\right) \qquad \mathrm{Ker}\,(\mathrm{S}(\theta) + \mathrm{I}_2) = \mathrm{Vect}\left(\begin{bmatrix}-\sin\frac{\theta}{2}\\ \cos\frac{\theta}{2}\end{bmatrix}\right)$$

Géométriquement, $\mathrm{S}(\theta)$ est donc la symétrie orthogonale autour de la droite qui fait un angle $\frac{\theta}{2}$ avec l'axe des abscisses.

On remarque aussi, par le calcul, que conformément à l'intuition,

$$\forall \theta, \varphi \in \mathbb{R} \qquad \mathrm{R}(\theta)\,\mathrm{R}(\varphi) = \mathrm{R}(\theta + \varphi)$$

$$\forall \theta \in \mathbb{R} \qquad \mathrm{R}(\theta)^{-1} = {}^t\mathrm{R}(\theta) = \mathrm{R}(-\theta)$$

et

$$\forall \theta \in \mathbb{R} \qquad \mathrm{R}(\theta) = \mathrm{I}_2 \iff \theta \in 2\pi\mathbb{Z}$$

De manière plus abstraite, mais aussi plus jolie, on vient de montrer que $\mathrm{R} : \theta \longmapsto \mathrm{R}(\theta)$ est un morphisme de groupes de $(\mathbb{R}, +)$ dans $(\mathrm{SO}_2(\mathbb{R}), \cdot)$, de noyau $2\pi\mathbb{Z}$.

Montrons qu'il est surjectif. Donnons-nous

$$\mathrm{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SO}_2(\mathbb{R})$$

Ceci signifie que les colonnes de M forment une base orthonormée de $\mathbb{R}^2$ pour le produit scalaire canonique, et que $\det \mathrm{M} = 1$. Ce qui fournit les relations :

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \\ ad - bc = 1 \end{cases}$$

À l'aide des deux premières relations et de l'étude des fonctions sinus et cosinus, il existe des réels θ et φ dans $[0\,;2\pi[$, tels que

$$a = \cos\theta \qquad c = \sin\theta \qquad b = \sin\varphi \qquad d = -\cos\varphi$$

Ensuite, $$1 = ad - bc = \cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta - \varphi)$$

donc θ et φ sont égaux modulo 2π. Mais comme ils sont dans $[0\,;2\pi[$, ils sont égaux. Ainsi, $\mathrm{M} = \mathrm{R}(\theta)$ et on a montré que

Théorème 3.4.1 (Paramétrisation de $\mathrm{SO}_2(\mathbb{R})$**)**
L'application R *est un morphisme surjectif de* $(\mathbb{R}, +)$ *sur* $\mathrm{SO}_2(\mathbb{R})$, *de noyau* $2\pi\mathbb{Z}$. *En particulier,* $\mathrm{SO}_2(\mathbb{R})$ *est commutatif et*

$$\mathrm{SO}_2(\mathbb{R}) = \{\mathrm{R}(\theta) \mid \theta \in \mathbb{R}\}$$

Le même raisonnement montre que tout automorphisme orthogonal de déterminant -1 est une réflexion de la forme $\mathrm{S}(\theta)$. Comme toute matrice orthogonale est de déterminant 1 ou -1,

$$\mathrm{O}_2(\mathbb{R}) = \underbrace{\{\mathrm{R}(\theta) \mid \theta \in \mathbb{R}\}}_{=\mathrm{SO}_2(\mathbb{R})} \cup \{\mathrm{S}(\theta) \mid \theta \in \mathbb{R}\}$$

Intéressons-nous maintenant à la structure multiplicative de ces réflexions. $\mathrm{S}(\theta)$. Qu'obtient-on si on les compose ? Le calcul montre que

$$\forall \theta, \varphi \in \mathbb{R} \qquad \mathrm{S}(\theta)\,\mathrm{S}(\varphi) = \mathrm{R}(\theta - \varphi)$$

Par conséquent, toute rotation peut s'écrire comme produit de deux réflexions : en effet, si $\theta \in \mathbb{R}$, on a $\mathrm{R}(\theta) = \mathrm{S}(\theta)\mathrm{S}(0)$. Mais cette décomposition n'est pas unique, évidemment : il suffit de prendre deux réels quelconques φ et ψ, tels que $\varphi - \psi = \theta\ [2\pi]$, pour avoir $\mathrm{R}(\theta) = \mathrm{S}(\varphi)\mathrm{S}(\psi)$.

Enfin, que donne le produit d'une rotation et d'une réflexion du type $\mathrm{S}(\theta)$? Il suffit encore de faire le calcul

$$\forall \theta, \varphi \in \mathbb{R} \qquad \mathrm{S}(\theta)\mathrm{R}(\varphi) = \mathrm{S}(\theta)\mathrm{S}(\theta)\mathrm{S}(\theta - \varphi) = \mathrm{S}(\theta - \varphi)$$

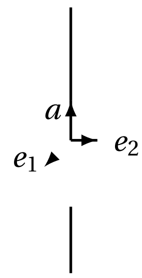
et $$\forall \theta, \varphi \in \mathbb{R} \qquad \mathrm{R}(\varphi)\mathrm{S}(\theta) = \mathrm{S}(\varphi + \theta)\mathrm{S}(\theta)\mathrm{S}(\theta) = \mathrm{S}(\varphi + \theta)$$

Il est intéressant de remarquer comme chacun de ces calculs est simple, et fournit immédiatement un résultat géométrique sans le moindre effort. Par exemple : si l'on fait une rotation d'angle φ, suivie d'une réflexion autour de la droite qui fait un angle $\theta/2$ avec l'horizontale, on obtient le même résultat que si l'on fait une réflexion autour de la droite qui fait un angle $\frac{\theta-\varphi}{2}$ avec l'horizontale. Ce n'est pas évident à voir géométriquement ; mais c'est trivial par le calcul matriciel.

3.5 Automorphismes orthogonaux en dimension 3

Commençons par expliquer comment il est possible d'orienter un plan, connaissant un vecteur normal à celui-ci.

Proposition 3.5.1 (Orientation d'un plan par un vecteur normal)
Soit E *un espace euclidien orienté de dimension 3. Soit* $a \in \mathrm{E}$ *de norme 1. Il existe une unique orientation de* $\mathrm{P} = (a)^{\circ}$ *telle que, pour toute base orthonormée directe* (e_1, e_2) *de* P*, la famille* (e_1, e_2, a) *est une base orthonormée directe de* E.

> Cette proposition peut faire peur, mais ce qu'elle dit est intuitivement très simple : l'orientation d'une droite de E détermine de manière « naturelle » une orientation de son plan orthogonal.
>
> 
>
>
> Supposons avoir choisi l'orientation habituelle de $\mathbb{R}^3$, dans laquelle la base caonique est directe. On prend une base orthonormée (e_1, e_2) de $(a)^{\circ}$ comme sur le dessin, de manière à avoir (e_1, e_2, a) orthonormée directe. On souhaite choisir une orientation de $(a)^{\circ}$: naturellement, on décide de dire que (e_1, e_2) est directe.

Preuve : On sait que P a des bases orthonormées et on en prend une $(\varepsilon_1, \varepsilon_2)$. Alors $(\varepsilon_1, \varepsilon_2, a)$ est une base orthonormée de E. Si elle est directe, on pose

$$u_1 = \varepsilon_1 \qquad u_2 = \varepsilon_2$$

et si elle est indirecte, on pose

$$u_1 = \varepsilon_2 \qquad u_2 = \varepsilon_1$$

On a ainsi une base orthonormée directe (u_1, u_2, a) de E, telle que (u_1, u_2) est une base orthonormée de P.

On choisit alors comme orientation pour P celle donnée par (u_1, u_2). Par définition, ceci veut dire que, si (e_1, e_2) est une base orthonormée de P, elle est directe si, et seulement si,

$$\det_{(u_1,u_2)}(e_1, e_2) = 1$$

Maintenant, si (e_1, e_2) est une base orthonormée directe de P, la matrice de la famille (e_1, e_2) dans (u_1, u_2) est orthogonale, de déterminant 1. Donc il existe $\theta \in \mathbb{R}$ tel que la matrice de (e_1, e_2) dans la base (u_1, u_2) soit

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Par suite, la matrice de (e_1, e_2, a) dans la base directe (u_1, u_2, a) est

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Son déterminant vaut 1, donc (u_1, u_2, a) est directe. Ceci démontre bien l'existence d'une orientation de P qui a la propriété demandée.

Pour l'unicité, c'est simple : supposons avoir choisi l'autre orientation de P. Ceci signifie que (u_1, u_2) est indirecte dans P, donc (le déterminant est alterné) que (u_2, u_1) est directe dans P. Alors la base (u_2, u_1, a) est indirecte dans E, toujours d'après les propriétés du déterminant. Et cette orientation ne satisfait pas la propriété voulue. □

Identifions maintenant les rotations d'un espace euclidien de dimension 3.

Proposition 3.5.2
Soient E *un espace euclidien de dimension 3 et* $f \in \mathscr{SO}(\mathrm{E})$. *Il existe une base orthonormée directe* $\mathscr{B} = (u, v, w)$ *de* E, *et* $\theta \in \mathbb{R}$, *tels que*

$$Mat_{\mathscr{B}}(f) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

On dit que f est la rotation d'angle θ autour du vecteur w.

Preuve : Si f est l'identité, la proposition est évidente puisque sa matrice est I_3 dans n'importe quelle base orthonormée ; il suffit alors de prendre $\theta = 0$.

Supposons alors que f n'est pas l'identité. Comme E est de dimension 3, f peut se décomposer comme produit d'une, deux ou trois réflexions. Mais f est dans $\mathscr{SO}(\mathrm{E})$ donc son déterminant vaut 1 ; par suite, f est le produit de deux réflexions distinctes r_1 et r_2. Par définition, $\mathrm{Ker}(r_1 - \mathrm{id})$ et $\mathrm{Ker}(r_2 - \mathrm{id})$ sont de dimension 2. D'après la relation de Grassmann, leur intersection est de dimension 1 : il existe w, de norme 1, tel que $r_1(w) = r_2(w) = w$. Par suite,

$$f(w) = r_1 r_2(w) = w$$

On note $\mathrm{P} = (w)^\circ$, qui est un plan. On l'oriente par son vecteur normal w, à l'aide de la **proposition 5.1**. Comme f préserve l'orthogonalité et que $f(w) = w$, P est stable par f : en effet,

$$\forall x \in \mathrm{P} \qquad \langle f(x) \mid w\rangle = \langle f(x) \mid f(w)\rangle = \langle x \mid w\rangle = 0$$

Donc f induit un automorphisme orthogonal de P. Et f ne fixe aucun vecteur de P, puisque $\mathrm{Ker}(f - \mathrm{id}) = \mathrm{Vect}\, w$. D'après l'étude des automorphismes orthogonaux en dimension 2, f est une rotation sur P. Alors si (u, v) est une base orthonormée directe de P, on trouve $\theta \in \mathbb{R}$ tels que la matrice de f dans la base (u, v) soit

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Par suite, (u, v, w) est une base orthonormée directe de E, dans laquelle la matrice de f a la forme voulue. □

La preuve de ce théorème nous dit, à nouveau, comment faire pour identifier une rotation f. On commence par chercher $\mathrm{Ker}(f - \mathrm{id})$, qui est nécessairement de dimension 1 et on prend un vecteur w dedans, de norme 1.

Ensuite, on détermine le plan $\mathrm{P} = (w)^\circ$ et on en choisit une base orthonormée (u, v) telle que (u, v, w) soit orthonormée directe. La matrice de f dans cette base aura alors la bonne forme.

Exemple 3.5.3
On reprend la matrice de l'**exemple 3.19** :

$$\mathrm{M} = \frac{1}{3}\begin{bmatrix} -2 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & -2 \end{bmatrix}$$

Le calcul de son déterminant, ou bien sa décomposition comme produit de deux réflexions, prouvent qu'il s'agit d'une rotation. On recherche $\operatorname{Ker}(\mathrm{M}-\mathrm{I}_3)$ et le calcul donne

$$\operatorname{Ker}(\mathrm{M}-\mathrm{I}_3) = \operatorname{Vect}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

et on pose
$$w = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Pour trouver une base orthonormée de $(w)^\circ$, il y a plusieurs méthodes. On peut en chercher une base, et utiliser Schmidt pour l'orthonormer. Mais on peut aller plus vite, puisqu'on est dans $\mathbb{R}^3$: il est clair que

$$u = \frac{1}{\sqrt{5}}\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$$

est orthogonal à w et de norme 1. On pose alors

$$v = w \wedge u = \frac{1}{\sqrt{30}}\begin{bmatrix} 5 \\ -2 \\ -1 \end{bmatrix}$$

D'après les propriétés du produit vectoriel, (u, v, w) est une base orthonormée directe de $\mathbb{R}^3$. On peut alors trouver l'angle de la rotation M en décomposant $\mathrm{M}u$ dans la base orthonormée (u, v) de $(w)^\circ$:

$$\mathrm{M}u = \frac{1}{\sqrt{5}}\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} = -u$$

> Sur cet exemple, la décomposition de $\mathrm{M}u$ s'obtient immédiatement. Mais en général, c'est à peine plus difficile puisque $\mathrm{M}u = \langle \mathrm{M}u \mid u\rangle\, u + \langle \mathrm{M}u \mid v\rangle\, v$.

Par suite, dans la base orthonormée directe (u, v, w), la matrice de M est

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ou encore
$$\mathrm{M} = \frac{1}{30}\begin{bmatrix} 0 & 5 & \sqrt{5} \\ \sqrt{6} & -2 & 2\sqrt{5} \\ -2\sqrt{6} & -1 & \sqrt{5} \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & \sqrt{6} & -2\sqrt{6} \\ 5 & -2 & -1 \\ \sqrt{5} & 2\sqrt{5} & \sqrt{5} \end{bmatrix}$$

M est la rotation autour de w, d'angle π.