ESPACES PREHILBERTIENS

PLAN

I : Rappels de 1ère année

Les définitions et propriétés vues dans le chapitre L1/ESPEUCL.PDF sont rappelées sans démonstration. Quelques compléments sont effectués par moment.

1- Produit scalaire

Soit E un espace vectoriel sur $\mathbb{R}$. Un **produit scalaire** est une application de $E \times E$ dans $\mathbb{R}$, qui à un couple (x, y) de vecteurs de E associe un réel noté $<x, y>$, et telle que :

$\forall x \in E$, l'application $y \in E \to <x, y>$ est linéaire
$\forall y$, l'application $x \in E \to <x, y>$ est également linéaire.
$\forall x \in E, \forall y \in E, <x, y> = <y, x>$
$\forall x \in E, <x, x> \geq 0$
$\forall x, <x, x> = 0 \Rightarrow x = 0$

On pose $\| x \| = \sqrt{<x, x>}$, **norme euclidienne** de x associée au produit scalaire.

E, espace vectoriel muni d'un produit scalaire, est appelé **espace préhilbertien**. Si, de plus, il est de dimension finie, il est dit **euclidien**.

EXEMPLES :

❑ Si E = $\mathbb{R}^n$, on pose <X,Y> = $\sum_{i=1}^{n} x_i y_i$ (produit scalaire canonique sur $\mathbb{R}^n$)

❑ Si E = $C^0([a,b])$, on pose $<f,g> = \int_a^b f(t)g(t)\,dt$. On peut prendre la même définition pour les fonctions continues par morceaux, mais dans ce dernier cas, on peut avoir $<f, f> = 0$ sans que f ne soit nul : prendre par exemple $f = 0$ sur $[a, b[$ et $f(b) \neq 0$. Par contre si f est continue, on a bien $<f, f> = 0 \Rightarrow f = 0$ car $\int_a^b f^2(t)\,dt = 0$ est l'intégrale nulle d'une fonction continue positive ou nulle, donc cette fonction est nulle. Ainsi, on dispose d'un produit scalaire sur $C^0([a, b])$, mais pas sur l'espace des fonctions continues par morceaux.

La norme euclidienne vérifie les propriétés suivantes :

$\forall x \in E, \| x \| \geq 0$
$\forall x \in E, \| x \| = 0 \Rightarrow x = 0$
$\forall x \in E, \forall \lambda \in \mathbb{R}, \| \lambda x \| = |\lambda| \, \| x \|$
$\forall (x, y) \in E^2, \| x + y \|^2 = \| x \|^2 + 2<x, y> + \| y \|^2$
$\forall (x, y) \in E^2, |<x,y>| \leq \| x \| \| y \|$ (**Inégalité de Cauchy-Schwarz**)
$\forall (x, y) \in E^2, \| x + y \| \leq \| x \| + \| y \|$ (**Inégalité triangulaire ou de Minkowski**)

On retrouve le produit scalaire à partir de la norme euclidienne au moyen des **formules de polarisation** :

$$<x, y> = \frac{\| x + y \|^2 - \| x \|^2 - \| y \|^2}{2} = \frac{\| x + y \|^2 - \| x - y \|^2}{4}$$

On dispose :

❑ du **théorème de Pythagore** :

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2 \Leftrightarrow <x,y> = 0$$

En effet, $\| x + y \|^2 = \| x \|^2 + \| y \|^2 + 2<x, y>$
Plus généralement, si les x_i sont deux à deux orthogonaux, on a :

$$\| \sum_{i=1}^{n} x_i \|^2 = \sum_{i=1}^{n} \| x_i \|^2$$

car $\| \sum_{i=1}^{n} x_i \|^2 = \sum_{i=1}^{n} \| x_i \|^2 + 2 \sum_{i<j} <x_i, x_j> = \sum_{i=1}^{n} \| x_i \|^2$ si tous les $<x_i, x_j>$ sont nuls pour $i \neq j$.

❑ de **l'identité du parallélogramme** :

$$\| x + y \|^2 + \| x - y \|^2 = 2 \| x \|^2 + 2 \| y \|^2$$

2- Orthogonalité

Deux vecteurs x et y sont dits **orthogonaux** si $<x, y> = 0$. x est dit **unitaire** si $\| x \| = 1$. Une famille $(x_i)_{i\in I}$ (I fini ou non), est dite **orthogonale** si pour tout i différent de j, $<x_i, x_j> = 0$. Elle est dite **orthonormale** ou **orthonormée** si, de plus, $\| x_i \| = 1$ pour tout i.

Deux sous-espaces vectoriels F et G d'un espace vectoriel euclidien sont dits **orthogonaux** si :

$$\forall x \in F, \forall y \in G, <x,y> = 0$$

Si on est en dimension finie, pour vérifier que deux sous-espaces F et G sont orthogonaux, il suffit de vérifier que les vecteurs d'une base de F sont orthogonaux aux vecteurs d'une base de G. Par ailleurs, deux sous-espaces vectoriels orthogonaux sont en somme directe.

Plus généralement, on a :

PROPOSITION

Soient $F_1, F_2, \dots, F_p$ *une famille de sous-espaces vectoriels deux à deux orthogonaux. Alors ils sont en somme directe.*

<u>Démonstration</u> :

❑ Si $x_1 + \dots + x_p = 0$ avec $x_i \in F_i$, alors, en faisant le produit scalaire par x_i, on obtient $<x_i, x_i> = 0$ puisque tous les autres produits scalaires sont nuls. D'où $x_i = 0$, et ceci, quel que soit x_i.

Soit F un sous-espace vectoriel d'un espace vectoriel E préhilbertion. On appelle **orthogonal** de F l'ensemble des vecteurs y de E tels que :

$$\forall x \in F, <x, y> = 0$$

On note cet ensemble F^{o} ou $F^{\perp}$.

$F^{\perp}$ est un sous-espace vectoriel de E, orthogonal à F. Si F est de dimension finie, pour appartenir à $F^{\perp}$, il suffit de vérifier l'orthogonalité avec une base de F.

Si $F \subset G$, alors $G^{\perp} \subset F^{\perp}$.

3- Bases orthonormales

Une base $(e_1, \dots, e_n)$ d'un espace vectoriel de dimension finie est dite **orthogonale** si les vecteurs de base sont deux à deux orthogonaux.

Une base $(e_1, \dots, e_n)$ est dite **orthonormale** ou **orthonormée** si elle est orthogonale et si tous ses vecteurs sont unitaires.

Un système de vecteurs non nuls constitué de vecteurs deux à deux orthogonaux forme un système est libre. Pour que ce soit une base, il suffit que le nombre de ces vecteurs soit égal à la dimension de l'espace.

Une base orthonormée vérifie :

$$<e_i, e_j> = \delta_{ij} = \left[\begin{array}{l} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{array} \right.$$

δ_{ij} s'appelle **symbole de Kronecker**.

L'intérêt d'une base orthonormée est que le produit scalaire s'y exprime très simplement. Si $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{i=1}^{n} y_i\, e_i$, alors :

$$<x,y> = \sum_{i=1}^{n} x_i\, y_i = X^{\mathrm{T}}Y \qquad \text{où } X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix},\ X^{\mathrm{T}} = (x_1 \dots x_n) \text{ et } Y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$$

et $$\| x \| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{X^{\mathrm{T}}X}$$

L'implication suivante est d'usage fréquent. Si u et v sont deux endomorphismes de E préhilbertien, alors :

$$\forall x \in E,\ \forall y \in E,\ <x,u(y)> = <x,v(y)> \Rightarrow u = v$$

En effet :

$\forall x \in E,\ \forall y \in E,\ <x,u(y)> = <x,v(y)>$

$\Rightarrow \quad \forall x \in E,\ \forall y \in E,\ <x,u(y) - v(y)> = 0$

$\Rightarrow \quad \forall y \in E,\ \| u(y) - v(y) \| = 0$ en prenant $x = u(y) - v(y)$

$\Rightarrow \quad \forall y \in E,\ u(y) = v(y)$

$\Rightarrow \quad u = v$

En dimension finie, la traduction matricielle de l'implication donne, pour deux matrices A et B éléments de $\mathcal{M}_n(\mathbb{R})$:

$$\forall X \in \mathbb{R}^n,\ \forall Y \in \mathbb{R}^n,\ X^{\mathrm{T}}AY = X^{\mathrm{T}}BY \Rightarrow A = B$$

Il suffit de considérer A et B comme les matrices de deux endomorphismes u et v dans une base orthonormée pour retrouver l'implication précédente.

Dans une base quelconque, on a :

$$<x,y> = \sum_{i,j} x_i\, y_j <e_i,e_j>$$

ou encore, si l'on note M la matrice de terme général $<e_i,e_j>$:

$$<x,y> = X^{\mathrm{T}}MY$$

Puisque $<e_i, e_j> = <e_j, e_i>$, M est symétrique.

Tout espace euclidien possède une base orthonormée. On en obtient une à partir d'une base quelconque en appliquant le procédé d'**orthogonalisation de Gram-Schmidt**.

Toute forme linéaire f dans un espace euclidien peut s'exprimer au moyen du produit scalaire. Soit en effet $(e_1, \dots, e_n)$ est une base orthonormale de E. Si $x = \sum_{i=1}^{n} x_i\, e_i$, avec x_i élément de $\mathbb{R}$, alors :

$$f(x) = f(\sum_{i=1}^{n} x_i\, e_i) = \sum_{i=1}^{n} x_i\, f(e_i) = <a,x>$$

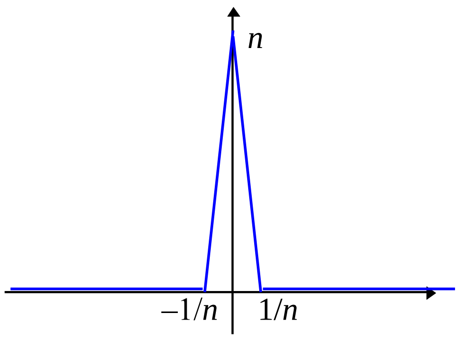
où $a = \sum_{i=1}^{n} f(e_i)\, e_i$. Si f est non nulle, a est un vecteur orthogonal à l'hyperplan Ker(f).

La propriété précédente n'existe pas en dimension infinie. Un contre-exemple est donné par l'impulsion de Dirac. Soit E l'espace $C^0([-1,1])$, muni du produit scalaire :

$$<f,g> = \int_{-1}^{1} f(t)g(t)\, dt$$

L'application $\delta : E \to \mathbb{R}$, qui à f associe $f(0)$ ne peut être définie à partir d'un produit scalaire avec une certaine fonction g. En effet, si on prend f_n paire définie par $f_n(x) = 1 - nx$ sur $[0, \frac{1}{n}]$ et 0 sur $[\frac{1}{n}, 1]$, alors il faudrait que g soit telle que, pour tout n, $\delta(f_n) = 1 = \int_{-1/n}^{1/n} f_n(t)\, g(t)\, dt$. Or l'intégrale est majorée par $\frac{2}{n}$ Max $\{|g(t)|, t \in [-1,1]\}$, quantité qui tend vers 0 quand n tend vers ∞ et ne peut donc rester égale à 1.

δ est une forme linéaire sur l'espace des fonctions continues, appelée **distribution de Dirac** (le mot distribution est utilisé de préférence à fonction car δ n'est pas une fonction de $\mathbb{R}$ dans $\mathbb{R}$). Cette forme linéaire ne peut s'exprimer comme produit scalaire. Cependant, on n'en est pas loin, car elle peut s'exprimer comme limite de produits scalaires. Considérons en effet les fonctions g_n suivantes :



Alors $\delta(f) = f(0) = \lim_{n\to+\infty} \int_{-1}^{1} f(t)\, g_n(t)\, dt$ pour toute fonction f continue. En fait, l'intégrale vaut $\int_{-1/n}^{1/n} f(t)\, g_n(t)\, dt$. Intuitivement, on comprend pourquoi cette intégrale tend vers $f(0)$. Quand n tend vers l'infini, t étant compris entre $-\frac{1}{n}$ et $\frac{1}{n}$, $f(t)$ vaut à peu près $f(0)$. L'intégrale vaut environ :

$$\int_{-1/n}^{1/n} f(0)\, g_n(t)\, dt = f(0) \int_{-1/n}^{1/n} g_n(t)\, dt = f(0).$$

D'une manière plus rigoureuse, on a :

$\forall \varepsilon > 0, \exists \alpha > 0, \forall t \in]-\alpha, \alpha[, |f(t) - f(0)| < \varepsilon$ (continuité de f en 0)

Mais par ailleurs, ε étant choisi :

$\exists N, \forall n \geq N, [-\frac{1}{n}, \frac{1}{n}] \subset]-\alpha,\alpha[$

Pour de tels n, on a, compte tenu du fait que $\int_{-1/n}^{1/n} f(0)\, g_n(t)\, \mathrm{d}t = f(0)$:

$$\left| \int_{-1}^{1} f(t)\, g_n(t)\, \mathrm{d}t - f(0) \right| = \left| \int_{-1/n}^{1/n} (f(t) - f(0))\, g_n(t)\, \mathrm{d}t \right|$$

$$\leq \int_{-1/n}^{1/n} |f(t) - f(0)|\, g_n(t)\, \mathrm{d}t$$

$$\leq \int_{-1/n}^{1/n} \varepsilon\ g_n(t)\, \mathrm{d}t = \varepsilon \int_{-1/n}^{1/n} g_n(t)\, \mathrm{d}t = \varepsilon$$

Ainsi :

$$\forall\, \varepsilon > 0, \exists\, \mathrm{N}, \forall\, n \geq \mathrm{N}, \left| \int_{-1}^{1} f(t)\, g_n(t)\, \mathrm{d}t - f(0) \right| \leq \varepsilon$$

ce qu'on voulait montrer.

L'interprétation précédente fait que δ est souvent qualifiée d'impulsion, modélisée par une fonction de support infiniment bref et ayant une valeur infiniment grande.

4- Projecteurs orthogonaux et symétries orthogonales

PROPOSITION

Si F *est un sous-espace vectoriel de* **dimension finie** *d'un espace préhilbertien* E, *alors* F *et* $\mathrm{F}^{\perp}$ *sont supplémentaires*.

Démonstration :

❑ E, lui, peut être de dimension infinie. Si $(e_1, .., e_m)$ est une base orthonormée de F, la décomposition d'un élément quelconque x de E en un élément de F et un élément de $\mathrm{F}^{\perp}$ s'obtient en écrivant :

$$x = \underbrace{\sum_{i=1}^{m} \langle e_i, x\rangle\, e_i}_{\in \mathrm{F}} + \underbrace{\left(x - \sum_{i=1}^{m} \langle e_i, x\rangle\, e_i\right)}_{\in \mathrm{F}^{\perp}}$$

L'application p_{F} définie par $p_{\mathrm{F}}(x) = \sum_{i=1}^{m} \langle e_i,x\rangle\, e_i$ est le projecteur orthogonal sur F. La quantité $\| x - p_{\mathrm{F}}(x) \|$ s'appelle **distance** de x à F. Si on considère la quantité $\| x - z \|$ lorsque z décrit F, celle-ci atteint son minimum pour $z = p_{\mathrm{F}}(x)$. On a également :

$$\| p_{\mathrm{F}}(x) \| \leq \| x \|$$

ce qui s'écrit aussi :

$$\sum_{i=1}^{n} |\langle e_i,x\rangle|^2 \leq \| x \|^2 \qquad \textbf{(inégalité de Bessel)}$$

Si F est une droite D de vecteur directeur unitaire u, $\mathrm{F}^{\perp}$ est alors un hyperplan H de vecteur normal u. La décomposition de x se réduit à :

$$x = \underbrace{\langle u, x\rangle\, u}_{} + \underbrace{(x - \langle u, x\rangle\, u)}_{}$$

$\in$ D $\qquad\qquad \in$ H

Le projecteur orthogonal sur D est l'application $p_D : x \in E \to <u, x> u$.
Le projecteur orthogonal sur H est l'application $p_H : x \in E \to x - <u, x> u$.
La symétrie orthogonale par rapport à D est l'application $s_D : x \to 2<u,x> u - x = 2p_D(x) - x$
La symétrie orthogonale par rapport à H est l'application $s_H : x \to x - 2<u,x> u = x - 2p_D(x) = -s_D(x)$
Dans le cas de la dimension 2 ou 3, cela permet de traiter tout projecteur orthogonal ou toute symétrie orthogonale.

EXEMPLE :

❑ Voici un exemple de projection dans un contexte non géométrique. Considérons la fonction sinus sur $[0, \pi]$. On cherche à approximer cette fonction par un polynôme de degré 2 sur $[0, \pi]$. Il existe plusieurs sens à donner à cette approximation, mais une des façons possibles de procéder est la suivante. On prend $E = C^0([0, \pi])$ muni du produit scalaire $<f, g> = \int_0^\pi f(t)g(t)\, dt$. Soit F le sous-espace vectoriel des polynômes de degré inférieur ou égal à 2. On prend comme approximation de $f = \sin$ le projeté orthogonal $p_F(f)$ de f sur F. Concrètement, on cherche a, b, c tels que $p_F(f)(x) = a + bx + cx^2$. Pour cela, on écrit que la fonction $f - p_F(f)$ est orthogonale à F, i.e. :

$$\begin{cases} <f - p_F(f), 1> = 0 \\ <f - p_F(f), x> = 0 \\ <f - p_F(f), x^2> = 0 \end{cases}$$

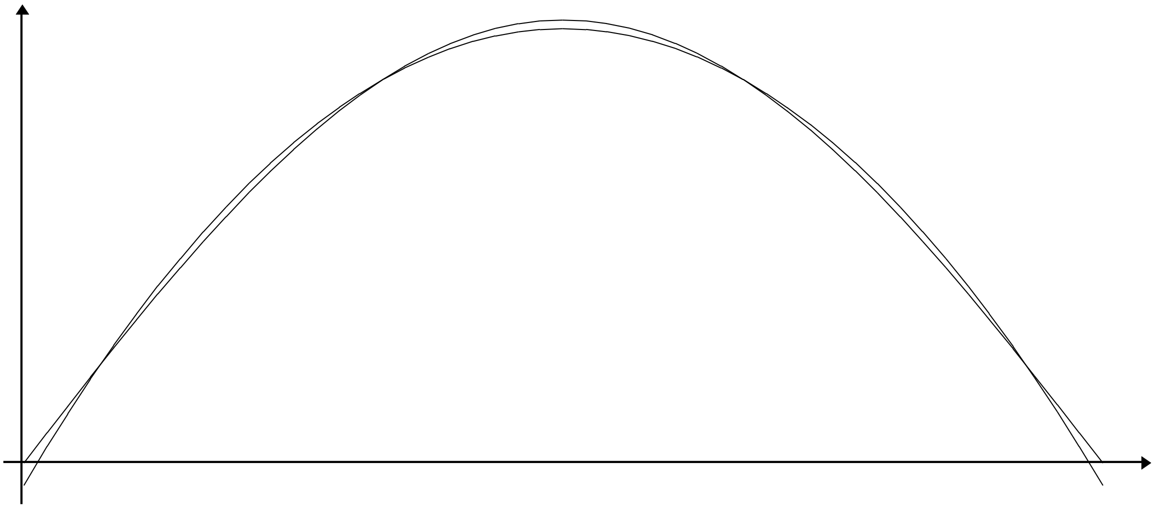
ce qui donne les équations :

$$\begin{cases} \displaystyle\int_0^\pi \sin(x) - a - bx - cx^2\, dx = 0 \\ \displaystyle\int_0^\pi (\sin(x) - a - bx - cx^2)x\, dx = 0 \\ \displaystyle\int_0^\pi (\sin(x) - a - bx - cx^2)x^2\, dx = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} 2 - a\pi - \dfrac{b\pi^2}{2} - \dfrac{c\pi^3}{3} = 0 \\ \pi - \dfrac{a\pi^2}{2} - \dfrac{b\pi^3}{3} - \dfrac{c\pi^4}{4} = 0 \\ \pi^2 - 4 - \dfrac{a\pi^3}{3} - \dfrac{b\pi^4}{4} - \dfrac{c\pi^5}{5} = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} a = \dfrac{12(\pi^2 - 10)}{\pi^3} \\ b = -\dfrac{60(\pi^2 - 12)}{\pi^4} \\ c = \dfrac{60(\pi^2 - 12)}{\pi^5} \end{cases}$$

Dans le graphe ci-dessous, le lecteur devrait pouvoir discerner quel est le graphe de sinus et celui du polynôme de degré 2 :



PROPOSITION
Soit E *un espace préhilertien de dimension quelconque et* F *un sous-espace vectoriel. On a* $F \subset (F^{\perp})^{\perp}$. *Si* E *est de dimension finie,* $F = (F^{\perp})^{\perp}$.

Démonstration :
❑ Si x appartient à F, alors :
$$\forall\, y \in F^{\perp},\ \langle x, y\rangle = 0$$
donc $x \in (F^{\perp})^{\perp}$
Mais si E est de dimension finie, on a de plus :
$\dim(E) = \dim(F) + \dim(F^{\perp})$ car F et $F^{\perp}$ sont supplémentaires
$\dim(E) = \dim(F^{\perp}) + \dim((F^{\perp})^{\perp})$ car $F^{\perp}$ et $(F^{\perp})^{\perp}$ sont supplémentaires
donc $\dim((F^{\perp})^{\perp}) = \dim(F)$. Comme les deux sous-espaces vectoriels ont même dimension et que l'un est inclus dans l'autre, ils sont égaux.

En dimension infinie, on peut n'avoir qu'une inclusion stricte.

EXEMPLES :

❑ Soit E l'espace $l^2(\mathbb{R})$ des suites réelles $u = (u_n)_{n\in\mathbb{N}}$ telles que $\sum_{n=0}^{\infty} u_n^{\ 2}$ converge, muni du produit scalaire $\langle u,v\rangle = \sum_{n=0}^{\infty} u_n v_n$. Considérons F le sous-espace vectoriel des suites nulles à partir d'un certain rang. Cet espace est engendré par la base $(e_n)_{n\in\mathbb{N}}$ où $e_n = (0, 0, ..., 0, 1, 0, ...)$, le 1 se trouvant au $n^{\text{ème}}$ rang. Soit u élément de $F^{\perp}$. Alors, pour tout n, $\langle e_n, u\rangle = 0$ donc $u_n = 0$. Donc $u = 0$. Donc $F^{\perp} = \{0\}$, et $F \oplus F^{\perp} = F \neq E$. $(F^{\perp})^{\perp} = E$ et $F \neq (F^{\perp})^{\perp}$.

❑ Soit $E = C^0([0, 1])$ muni du produit scalaire $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Considérons le sous-espace vectoriel F des fonctions f telles que $\int_0^{1/2} f(t)\,dt = 0$. F est le noyau de la forme linéaire φ :

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

Il s'agit donc d'un hyperplan. Montrons que $F^\perp = \{0\}$. Par l'absurde, supposons que $F^\perp \neq \{0\}$ et soit g un élément unitaire de $F^\perp$. g n'étant pas élément de F et F étant un hyperplan, la droite engendrée par g est supplémentaire de F (revoir au besoin le chapitre L1/LINEF.PDF), donc $E = F \oplus F^\perp$. Toute fonction f de E peut se décomposer sous la forme $f = h + \lambda g$, avec h élément de F et λ réel. On a donc $\langle f, g\rangle = \langle h, g\rangle + \lambda \langle g, g\rangle = \lambda$, donc $h = f - \langle f, g\rangle\, g$. Comme h appartient à F, on a $\varphi(h) = 0$, donc pour tout f de E, $\varphi(f) = \langle f, g\rangle\, \varphi(g)$. $\varphi(g) \neq 0$ car $g \notin F$. Prenons en particulier la fonction f définie par :

$$f(x) = \begin{cases} 0 \text{ sur } [0, \frac{1}{2}] \\ x - \frac{1}{2} \text{ sur } [\frac{1}{2}, 1] \end{cases}$$

Comme $fg = 0$ sur $[0, \frac{1}{2}]$, $\varphi(fg) = \int_0^{1/2} f(t)g(t)\,dt = 0$, donc $\langle fg, g\rangle\, \varphi(g) = 0$ donc $\langle fg, g\rangle = 0$, donc :

$$\int_0^1 f(x)g(x)^2\,dx = 0$$

$$\Rightarrow \quad \int_{1/2}^1 f(x)g(x)^2\,dx = 0 \qquad \text{puisque } f = 0 \text{ sur } [0, \tfrac{1}{2}]$$

La fonction fg^2 étant continue sur $[\frac{1}{2}, 1]$, positive ou nulle, et d'intégrale nulle, elle est identiquement nulle (revoir au besoin les propriétés de l'intégrale dans le chapitre L1/INTEGRAL.PDF). Comme $f > 0$ sur $]\frac{1}{2}, 1]$, $g = 0$ sur le même intervalle, et par continuité, $g(\frac{1}{2}) = 0$.

Reprenons alors la relation $\varphi(f) = \langle f, g\rangle\, \varphi(g)$. Pour tout f de E, elle s'écrit :

$$\int_0^{1/2} f(x)\,dx = \int_0^1 f(x)g(x)\,dx \times \varphi(g)$$

$$= \int_0^{1/2} f(x)g(x)\varphi(g)\,dx \qquad \text{puisque } g = 0 \text{ sur } [\tfrac{1}{2}, 1]$$

donc $\displaystyle\int_0^{1/2} f(x)\,(1 - g(x)\varphi(g))\,dx = 0$

En prenant $f(x) = 1 - g(x)\varphi(g)$, on obtient $\displaystyle\int_0^{1/2} (1 - g(x)\varphi(g))^2\,dx = 0$ donc $1 - g(x)\varphi(g) = 0$ pour tout x de $[0, \frac{1}{2}]$ et en particulier, $g(\frac{1}{2})\varphi(g) = 1$, ce qui est contradictoire avec $g(\frac{1}{2}) = 0$.

❑ Dans une annexe du chapitre L3/METRIQUE.PDF, on montre que, si φ est continue sur [0, 1] et si, pour tout n, $\displaystyle\int_0^1 \varphi(t)\, t^n\,dt = 0$, alors $\varphi = 0$. Soit $E = C^0([0, 1])$ le même espace préhilbertien que dans l'exemple précédent. La propriété énoncée exprime que, dans E, le sous-espace vectoriel F des fonctions polynomiales admet pour orthogonal $F^\perp = \{0\}$. On a donc $(F^\perp)^\perp = E \neq F$.

❑ Deux autres exemples sont donnés dans les exercices du chapitre *Polynômes orthogonaux* L2/POLORTHO.PDF.

II : Endomorphismes dans les espaces euclidiens

On se place ici dans des espaces de dimension finie sur $\mathbb{R}$.

1- Adjoint d'un endomorphisme

Considérons un espace euclidien E muni d'une base **orthonormée** $(e_1, ..., e_n)$ et u un endomorphisme de E de matrice M dans la base $(e_1, ..., e_n)$. Si x est un vecteur de E, on note X la colonne de ses composantes dans la même base. $u(x)$ a alors pour composantes MX. Si y est un autre vecteur de E, de composantes Y, on rappelle que $<x, y> = X^TY$, où X^T est la ligne transposée de la colonne X. Considérons maintenant la quantité :

$$\begin{aligned} <u(x),y> &= (MX)^TY \\ &= X^TM^TY \end{aligned}$$

Il y a deux façons de lire cette expression. Ou bien on applique M sur X puis on fait le produit scalaire par Y, ce qui donne $(X^TM^T)Y$, ou bien on applique M^T sur Y puis on fait le produit scalaire par X ce qui donne $X^T(M^TY)$. Introduisons donc l'application linéaire u^* de matrice M^T dans la base $(e_1, ..., e_n)$. Comme $(X^TM^T)Y = X^T(M^TY)$, on a :

$$\boxed{\forall x, \forall y, <u(x), y> = <x, u^*(y)>}$$

u^* s'appelle endomorphisme **adjoint** de u. La relation précédente définit u^* sans ambiguïté indépendamment de toute base. En effet, l'application $x \to <u(x), y>$ étant une forme linéaire, en vertu de l'isomorphisme entre E et son dual E*, il existe un vecteur a tel que, pour tout x :

$$<u(x), y> = <x, a>$$

a dépend évidemment de y n'est autre que le $u^*(y)$ vu ci-dessus. Il est alors facile de vérifier que u^* est une fonction linéaire de y. En outre l'unicité de a entraîne celle de $u^*(y)$ et donc celle de u^*.

Les propriétés de $u \in L(E) \to u^* \in L(E)$ sont identiques à celles de $M \in \mathcal{M}_n(\mathbb{R}) \to M^T$. En particulier :

- La transposition est linéaire et involutive. Il en est de même de *. Autrement dit, $(u + v)^* = u^* + v^*$ et $(\lambda u)^* = \lambda u^*$. Enfin $(u^*)^* = u$. Ces propriétés peuvent se montrer directement sans recourir à une base pour se ramener à la transposition. Considérons par exemple la dernière :

$\qquad \forall x, \forall y, <u(x),y> = <x,u^*(y)>$
$\Rightarrow \quad \forall x, \forall y, <u(x),y> = <u^*(y),x> \qquad$ par symétrie du produit scalaire
$\Rightarrow \quad \forall x, \forall y, <u(x),y> = <y,(u^*)^*(x)> \qquad$ par définition de l'adjoint de u^*
$\Rightarrow \quad \forall x, \forall y, <u(x),y> = <(u^*)^*(x),y> \qquad$ par symétrie du produit scalaire
$\Rightarrow \quad u = (u^*)^*$

Ou encore la linéarité :

$\qquad \forall x, \forall y, <u(x),y> = <x,u^*(y)>$
$\Rightarrow \quad \forall x, \forall y, \forall \lambda, \lambda<u(x),y> = \lambda<x,u^*(y)>$
$\Rightarrow \quad \forall x, \forall y, \forall \lambda, <\lambda u(x),y> = <x,\lambda u^*(y)>$
$\Rightarrow \quad \forall x, \forall y, \forall \lambda, <x,(\lambda u)^*(y)> = <x,\lambda u^*(y)>$

$\Rightarrow \quad \forall \lambda, (\lambda u)^* = \lambda u^*$
etc...

- On a aussi $(AB)^T = B^T A^T$, d'où nécessairement $(u \circ v)^* = v^* \circ u^*$, ce qu'on peut également vérifier directement :
  $\forall x, \forall y, <u(x),y> = <x,u^*(y)>$

$\Rightarrow \quad \forall x, \forall y, <u(v(x)),y> = <v(x),u^*(y)>$
$\Rightarrow \quad \forall x, \forall y, <(u \circ v)(x)),y> = <x,v^*(u^*(y))>$
$\Rightarrow \quad \forall x, \forall y, <x,(u \circ v)^*(y)> = <x,(v^* \circ u^*)(y)>$
$\Rightarrow \quad (u \circ v)^* = v^* \circ u^*$

- Dans le cas d'un endomorphisme bijectif (automorphisme) u, en prenant $v = u^{-1}$, on a :
  $(u^{-1})^* \circ u^* = (u \circ u^{-1})^* = \text{Id}^* = \text{Id}$

donc $\quad (u^{-1})^* = (u^*)^{-1}$

- $\text{rg}(u^*) = \text{rg}(u)$ et $\text{Tr}(u^*) = \text{Tr}(u)$, résultant des propriétés analogues sur les matrices : $\text{rg}(M^T) = \text{rg}(M)$ et $\text{Tr}(M^T) = \text{Tr}(M)$.

Une dernière propriété enfin :
PROPOSITION :
Soit F *un sous-espace vectoriel de* E, *stable par un endomorphisme u. Alors* $F^\perp$ *est stable par u*.*

<u>Démonstration</u> :
❑ Soit x un élement de $F^\perp$. Il s'agit de montrer que $u^*(x)$ appartient à $F^\perp$, autrement dit, que, pour tout y de F, on a $<y, u^*(x)> = 0$. Or :

$<y, u^*(x)> = <u(y), x>$ avec $u(y)$ dans F, puisque F est stable par u
$= 0$ puisque $u(y) \in F$ et $x \in F^\perp$

On peut de même définir l'**adjoint d'une application linéaire** u d'un espace euclidien (E, < , >) dans un espace euclidien (F, (,)). Il s'agit de l'application linéaire u^* de F dans E vérifiant :

$$\forall x \in E, \forall y \in F, (u(x), y) = <x, u^*(y)>$$

Si on dispose d'une base orthonormée $(e_1, ..., e_p)$ de E et d'une base orthonormée $(\varepsilon_1, ..., \varepsilon_n)$ de F, notons M la matrice de u dans ces deux bases, élément de $\mathcal{M}_{np}(\mathbf{R})$. Notons $X \in \mathbf{R}^p$ les composantes de x et $Y \in \mathbf{R}^n$ les composantes de y. Le produit scalaire $(u(x), y)$ s'écrit :

$$(u(x), y) = (MX)^T Y = (X^T M^T) Y = X^T (M^T Y) = <x, u^*(y)>$$

avec donc ici aussi M^T matrice de u^*, mais élément de $\mathcal{M}_{pn}(\mathbf{R})$.

2- Isométries vectorielles

On s'intéresse aux endomorphismes u qui préservent la norme, c'est–à–dire tels que :

$$\forall x \in E, \ \| u(x) \| = \| x \|$$

De tels endomorphismes sont qualifiés d'**isométries vectorielles** ou d'**endomorphismes orthogonaux**. Ils disposent des propriétés suivantes :

PROPOSITION :
Il y a équivalence entre :
(i) $\forall x \in E, \ \| u(x) \| = \| x \|$. *On dit que u conserve la norme.*

(ii) $\forall x \in E, \forall y \in E$, $<u(x), u(y)> = <x, y>$. *On dit que u conserve le produit scalaire.*
(iii) *u transforme une base orthonormée en une base orthonormée.*
(iv) $u^*u = uu^* = \mathrm{Id}$

Démonstration :
❑ (ii) $\Rightarrow$ (i) est trivial en prenant $x = y$

❑ (i) $\Rightarrow$ (ii) en utilisant une formule de polarisation, par exemple $<x, y> = \dfrac{\| x + y \|^2 - \| x - y \|^2}{4}$. Il en résulte que, si u conserve la norme, alors u conserve le produit scalaire :

$$<u(x), u(y)> = \frac{\| u(x) + u(y) \|^2 - \| u(x) - u(y) \|^2}{4}$$
$$= \frac{\| u(x + y) \|^2 - \| u(x - y) \|^2}{4} = \frac{\| x + y \|^2 - \| x - y \|^2}{4} = <x, y>$$

❑ (ii) $\Rightarrow$ (iii), car si (e_i) est une base orthonormée quelconque, alors, pour tout i, j :
$<e_i, e_j> = \delta_{ij}$ (symbole de Kronecker)
donc $<u(e_i), u(e_j)> = <e_i, e_j> = \delta_{ij}$ ce qui signifie que $(u(e_i))$ est orthonormée.

❑ (iii) $\Rightarrow$ (i) : supposons qu'un opérateur u transforme une base orthonormée (e_i) en une base orthonormée (ε_i). Si $x = \sum_{i=1}^n x_i e_i$, on a :

$$\| u(x) \| = \| \sum_{i=1}^n x_i\, u(e_i) \| = \| \sum_{i=1}^n x_i \varepsilon_i \| = \sqrt{\sum_{i=1}^n x_i^2} = \| x \|$$

❑ (ii) $\Leftrightarrow$ (iv) :
$\forall x \in E, \forall y \in E$, $<u(x), u(y)> = <x, y>$
$\Leftrightarrow$ $\forall x \in E, \forall y \in E$, $<x, u^*u(y)> = <x, y>$
$\Leftrightarrow$ $u^*u = \mathrm{Id}$
$\Leftrightarrow$ $u^* = u^{-1}$ l'implication $\Rightarrow$ étant valide pour des endomorphismes en dimension finie
$\Leftrightarrow$ $uu^* = \mathrm{Id}$

La matrice de u dans une base orthonormée vérifie donc les propriétés suivantes : les vecteurs colonnes sont de norme 1, les colonnes sont deux à deux orthogonales. On notera que le sens "*une*" utilisée dans la propriété (iii) peut se prendre aussi bien dans le sens "*une quelconque*" que dans le sens "*une particulière*". Le sens (ii) $\Rightarrow$ (iii) s'applique à une base quelconque orthonormée, mais il suffit que la propriété (iii) soit vraie dans une base orthonormée particulière pour prouver i).

PROPRIETES :
(i) *Une isométrie est bijective.*
(ii) *L'ensemble des isométries muni de la composition des endomorphismes est un sous-groupe de* GL(E), *appelé* **groupe orthogonal** *et noté* O(E). *On rappelle qu'une isométrie s'appelle également endomorphisme orthogonal.*
(iii) *Une isométrie admet pour déterminant* 1 *ou* –1.
(iv) *L'ensemble des isométries de déterminant* 1 *muni de la composition des endomorphismes est un sous-groupe de* O(E), *appelé* **groupe spécial orthogonal** *et noté* SO(E).

(v) *Si* F *est un sous-espace vectoriel stable par une isométrie u, alors* $F^{\perp}$ *est aussi stable par u.*

On rappelle que (GL(E), $\circ$) est le groupe des endomorphismes bijectifs de E, ou automorphismes.

<u>Démonstration</u> :

❑ (i) E étant de dimension finie, il suffit de montrer qu'un tel opérateur est injectif. Or :

$$u(x) = 0 \Rightarrow \| u(x) \| = 0 \Rightarrow \| x \| = 0 \Rightarrow x = 0$$

La bijectivité résulte aussi de la propriété $u^*u = \mathrm{Id}$ en dimension finie.

❑ (ii) O(E) est non vide car il contient Id. Le i) montre en outre que O(E) est inclus dans GL(E). Montrons la stabilité pour la composition et l'inverse.
Si u et v sont éléments de O(E), alors, pour tout x :

$$\begin{aligned} \| u \circ v(x) \| = \| u(v(x)) \| &= \| v(x) \| &&\text{car } u \text{ conserve la norme} \\ &= \| x \| &&\text{car } v \text{ conserve la norme} \end{aligned}$$

donc $u \circ v$ conserve la norme.
Par ailleurs, $\| u^{-1}(x) \| = \| u(u^{-1}(x)) \|$ car u conserve la norme
$= \| x \|$
donc u^{-1} conserve la norme.

❑ (iii) Si (e_i) est une base orthonormée dans laquelle u possède une matrice M, alors u^* a pour matrice M^T, et u^*u a pour matrice M^TM. Comme $u^*u = \mathrm{Id}$, on a $M^TM = \mathrm{Id}$. Donc :

$$\det(M^TM) = 1 = \det(M^T)\det(M) = \det(M)^2 = \det(u)^2$$

❑ iv) SO(E) est non vide car il contient Id. Le fait que ce soit un sous-groupe de O(E) résulte du fait que $\det(u \circ v) = \det(u)\det(v)$ et que $\det(u^{-1}) = \dfrac{1}{\det(u)}$. Si u et v sont des éléments de O(E) de déterminant 1, il en est alors de même de $u \circ v$ et de u^{-1}. Les éléments de SO(E) sont appelées **rotations** ou **isométries directes**. Les autres isométries sont qualifiées d'**indirectes**.

❑ v) Soit F stable par u. On a $u(F) \subset F$ et même $u(F) = F$ car u est un automorphisme donc conserve la dimension de F. Donc $F = u^{-1}(u(F)) = u^{-1}(F)$. Ainsi F est aussi stable par u^{-1}. Montrons que $F^{\perp}$ est stable par u. Pour tout x de $F^{\perp}$, on a :

$$\begin{aligned} \forall y \in F, \langle u(x), y\rangle = \langle u(x), u(u^{-1}(y)) \rangle & \\ = \langle x, u^{-1}(y) \rangle & \qquad \text{car } u \text{ est une isométrie} \\ = 0 & \qquad \text{car } x \in F^{\perp} \text{ et } u^{-1}(y) \in F \end{aligned}$$

donc $u(x) \in F^{\perp}$.
On peut aussi écrire que :

F stable par u
$\Rightarrow$ $F^{\perp}$ stable par u^*
$\Rightarrow$ $F^{\perp}$ stable par u^{-1} car $u^* = u^{-1}$ puisque $u^*u = \mathrm{Id}$
$\Rightarrow$ $u^{-1}(F^{\perp}) \subset F^{\perp}$
$\Rightarrow$ $u^{-1}(F^{\perp}) = F^{\perp}$ car u^{-1} est un automorphisme donc conserve la dimension
$\Rightarrow$ $F^{\perp} = u(F^{\perp})$ en composant par u
donc $F^{\perp}$ est stable par u.

Passons aux matrices.

PROPOSITION :

Il y a équivalence entre :

(i) M *est la matrice d'une isométrie u dans une base orthonormée.*

(ii) $M^T M = I_n$

(iii) *Les colonnes de* M *forment une base orthonormée de* $\mathbb{R}^n$.

Une telle matrice s'appelle **matrice orthogonale**.

Démonstration :

❑ L'équivalence (i) ⇔ (ii) a déja été rencontrée dans le (iii) de la proposition précédente. L'équivalence avec le (iii) actuel est la traduction matricielle du fait que *u* transforme une base orthonormée en une base orthonormée.

Les propriétés de telles matrices sont analogues à celles des applications correspondantes, en particulier :

PROPRIETES :

(i) *Une matrice orthogonale est inversible.*

(ii) *L'ensemble des matrices orthogonales* $n \times n$ *muni du produit des matrices est un sous-groupe de* $GL_n(\mathbb{R})$, **groupe orthogonal** *et noté* $O_n(\mathbb{R})$.

(iii) *Une matrice orthogonale admet pour déterminant* 1 *ou* –1.

(iv) *L'ensemble des matrices orthogonales de déterminant* 1 *muni du produit des matrices est un sous-groupe de* $O_n(\mathbb{R})$ *appelé* **groupe spécial orthogonal** *et noté* $SO_n(\mathbb{R})$.

EXEMPLES :

❑ Les symétries orthogonales sont des isométries. Cela est évident en utilisant la définition : Soit *s* symétrie orthogonale par rapport à F. On a :

$$E = F \oplus F^\perp \to E$$

$$x = y + z \to s(x) = y - z$$

et $\| x \|^2 = \| y \|^2 + \| z \|^2 = \| s(x) \|$ d'après le théorème de Pythagore.

On peut également remarquer que, si (e_i) est une base orthonormée de E dont les p premiers vecteurs forment une base de F (et donc les $n - p$ restants une base de $F^\perp$), alors la matrice de la symétrie s est $\begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}$ qui est une matrice orthogonale puisque ses colonnes forment une base orthonormée. Son déterminant vaut $(-1)^{n-p}$ et donc s appartient à SO(E) si et seulement si $n - p$ est pair. Ainsi, certaines symétries peuvent également être classée parmi les rotations. C'est le cas en dimension 3 pour les symétries par rapport à une droite, qui peuvent être également considérées comme des demi-tours autour de cette droite.

❑ Soit H un hyperplan et s_H la symétrie orthogonale par rapport à H. Alors s_H s'appelle une **réflexion** et son déterminant vaut –1.

❑ Une fonction f de E euclidien dans E vérifiant :

$$f(0) = 0 \qquad \text{et} \qquad \forall x, \forall y, \| f(x) - f(y) \| = \| x - y \|$$

est nécessairement une isométrie linéaire.

En effet, en posant $y = 0$, on a $\| f(x) \| = \| x \|$. En développant les carrés $\| f(x) - f(y) \|^2 = \| x - y \|^2$ et en simplifiant par les termes $\| f(x) \|^2 = \| x \|^2$ et $\| f(y) \|^2 = \| y \|^2$, on obtient l'égalité $<f(x), f(y)> = <x, y>$. Donc, pour tout y :

$$\begin{aligned} <f(\lambda x) - \lambda f(x), f(y)> &= <f(\lambda x), f(y)> - <\lambda f(x), f(y)> \\ &= <f(\lambda x), f(y)> - \lambda <f(x), f(y)> \\ &= <\lambda x, y> - \lambda <x, y> \\ &= 0 \end{aligned}$$

Donc $f(\lambda x) - \lambda f(x)$ est orthogonal à $f(\mathrm{E})$ et donc à $\mathrm{Vect}(f(\mathrm{E}))$. Comme $f(\lambda x) - \lambda f(x) \in \mathrm{Vect}(f(\mathrm{E}))$, $f(\lambda x) - \lambda f(x)$ est orthogonal à lui-même, donc est nul. Donc $f(\lambda x) = \lambda f(x)$.
On procède de même pour $f(x + y)$.

❑ Le procédé d'orthogonalisation de Schmidt permet de montrer la propriété suivante :
Soit $\mathrm{A} \in \mathrm{GL}_n(\mathbf{R})$. *Alors il existe* $\mathrm{P} \in \mathrm{O}_n(\mathbf{R})$ *et* T *une matrice triangulaire supérieure inversible telle que* $\mathrm{A} = \mathrm{PT}$.
En effet, si on considère A comme la matrice de passage P_{ce} de la base canonique (c) de $\mathbf{R}^n$ à une base (e) quelconque, le procédé d'orthogonalisation de Schmidt fournit une base (ε) orthonormée telle que la matrice de passage $\mathrm{P}_{\varepsilon e}$ de (ε) à (e) soit triangulaire supérieure. On a alors :

$$\mathrm{A} = \mathrm{P}_{ce} = \mathrm{P}_{c\varepsilon}\mathrm{P}_{\varepsilon e}$$

où $\mathrm{P}_{c\varepsilon}$ est la matrice de passage de la base (c) à la base (ε). Ces deux bases étant orthonormée, $\mathrm{P}_{c\varepsilon}$ est une matrice P orthogonale. La matrice triangulaire T demandée n'est autre que $\mathrm{P}_{\varepsilon e}$.

Soit par exemple $\mathrm{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 1 & 3 \end{pmatrix}$, de sorte que $e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $e_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $e_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Le procédé d'orthogonalisation de Schmidt conduit à :

$$\varepsilon_1 = \frac{1}{\sqrt{3}} e_1$$

$$\varepsilon_2 = \frac{1}{\sqrt{6}} e_2 \qquad \text{car } e_1 \text{ est orthogonal à } e_2$$

$$\varepsilon_3 \text{ colinéaire à } e_3 - 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ d'où } \varepsilon_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} e_3 - \sqrt{2}\, e_1$$

On a donc :

$$\mathrm{P} = \mathrm{P}_{c\varepsilon} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\mathrm{P}_{e\varepsilon} = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & -\sqrt{2} \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$T = P_{\varepsilon e} = P_{e\varepsilon}^{-1} = \begin{pmatrix} \sqrt{3} & 0 & 2\sqrt{3} \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

Ainsi :

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 2\sqrt{3} \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

PROPOSITION

Les réflexions engendrent le groupe orthogonal O(E).

<u>Démonstration</u> :

Il s'agit de montrer que tout endomorphisme orthogonal est la composée de réflexions. On montrera de plus que le nombre de réflexions à utiliser est inférieur ou égal à dim(E). On raisonne par récurrence sur dim(E). Id est considéré comme un produit vide de réflexions.

❑ Si dim(E) = 1, alors les seuls endomorphismes orthogonaux sont Id et – Id qui est une réflexion.

❑ Si dim(E) = 2, alors, dans une base orthonormée, la matrice d'un endomorphisme orthogonal peut être :

une matrice de réflexion et il n'y a rien à montrer

une matrice $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ de rotation. Cette matrice est le produit de deux matrices de réflexion, par exemple :

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

❑ Supposons la propriété vérifiée pour les espaces vectoriels euclidiens de dimension $n - 1$ et montrons-la pour un espace vectoriel de dimension n. Soit u un endomorphisme orthogonal. Soit x unitaire.

Si $u(x) = x$, soit F l'hyperplan orthogonal à x. Vect(x) étant stable par u et u étant orthogonal, le sous-espace vectoriel F orthogonal à x est également stable par u. La restriction $u_{|F}$ est un endomorphisme orthogonal de F, sous-espace vectoriel de dimension $n - 1$. Donc, d'après l'hypothèse de récurrence, $u_{|F}$ est le produit d'au plus $n - 1$ réflexions s de F. On étend chacune de ces réflexions à E tout entier en posant $s(x) = x$. L'extension ainsi définie est une réflexion de E et u est le produit de ces réflexions.

Si $u(x) \neq x$, soit $y = u(x) - x$ qui est orthogonal à $u(x) + x$. Soit f la réflexion par rapport à l'hyperplan orthogonal à y. On a donc :

$$f(y) = -y$$

donc :

$$f(u(x)) - f(x) = -u(x) + x$$

De plus, comme $u(x) + x$ est dans l'hyperplan invariant par f, on a :

$$f(u(x)) + f(x) = u(x) + x$$

donc :

$$f(u(x)) = x$$

$f \circ u$ est un endomorphisme orthogonal laissant x invariant. On s'est donc ramené au cas précédent. $f \circ u$ est la composée d'au plus $n - 1$ réflexions, et, en composant par f à gauche, u est la composée d'au plus n réflexions.

3- Isométries en dimension 2 ou 3

❑ En dimension 2, dans une base orthonormée, la matrice d'un opérateur orthogonal u est de la forme $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $a^2 + c^2 = b^2 + d^2 = 1$, $ac + bd = 0$. La première colonne vérifiant $a^2 + b^2 = 1$, on peut trouver θ tel que $a = \cos(\theta)$ et $c = \sin(\theta)$. La deuxième colonne est orthogonale à la première, donc colinéaire au vecteur $\begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$. Etant unitaire, il n'y a que deux choix possibles : $\pm \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$.

Les matrices orthogonales sont donc de deux types :

$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ de déterminant 1. C'est la rotation R_θ d'angle θ.

$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$ de déterminant -1. C'est la symétrie orthogonale $S_{\theta/2}$ par rapport à la droite faisant un angle $\theta/2$ avec le premier vecteur de base.

On vérifiera aisément que, pour tout θ et φ, $R_\theta \circ R_\varphi = R_{\theta+\varphi}$.

❑ Soit u un endomorphisme orthogonal en dimension 3. Nous raisonnerons sur la dimension du sous-espace $\mathrm{Inv}(u)$ des vecteurs invariants par u :

- Si $\dim(\mathrm{Inv}(u)) = 3$, alors $u = \mathrm{Id}$.

- Si $\dim(\mathrm{Inv}(u)) = 2$, alors u est la réflexion par rapport au plan $\mathrm{Inv}(u)$.

En effet, la droite orthogonale à $\mathrm{Inv}(u)$ est globalement invariante et son vecteur directeur a, n'appartenant pas à $\mathrm{Inv}(u)$, est nécessairement transformé en un vecteur de même norme que a, élément de la même droite et différent de a. Ce ne peut être $-a$. Ainsi, $u(a) = -a$.

- Si $\dim(\mathrm{Inv}(u)) = 1$, alors le sous-espace orthogonal à $\mathrm{Inv}(u)$ est un plan stable par u.

La restriction de u à ce plan est une isométrie de ce plan, n'admettant aucun vecteur invariant non nul. Cette restriction est donc une rotation du plan, d'angle θ non nul. u s'appelle rotation de l'espace, dont l'axe est la droite $\mathrm{Inv}(u)$. La matrice de u dans une base orthonormée formée d'une base orthonormée du plan $(\mathrm{Inv}(u))^\perp$ et d'un vecteur unitaire de la droite $\mathrm{Inv}(u)$ est :

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Pour $\theta = 0$, on a $u = \mathrm{Id}$. L'identité peut être vue comme une rotation d'angle nul et d'axe quelconque.

Si a est un vecteur directeur unitaire de l'axe et si x est un vecteur non nul du plan orthogonal à $\mathrm{Inv}(u)$, alors $a \wedge x$ est lui aussi dans ce plan, orthogonal à x, de même norme, et la base $(a, x, a \wedge x)$ est directe. L'image par u de x est alors :

$$u(x) = \cos(\theta)x + \sin(\theta)\, a \wedge x$$

Si $\theta = \pi$, on parle de **demi-tour** ou de **retournement**. C'est aussi la symétrie orthogonale s_D par rapport à l'axe D du demi-tour.

Pour un vecteur x quelconque, la composante de x appartenant à l'axe (à savoir $\langle x, a\rangle a$) est invariante par u, alors que la composante de x appartenant au plan (à savoir $x - \langle x, a\rangle a$) est transformé selon la formule précédente. On a donc, dans le cas général :

$$u(x) = \langle x, a\rangle a + \cos(\theta)(x - \langle x, a\rangle a) + \sin(\theta)\, a \wedge (x - \langle x, a\rangle a)$$

donc :

$$\boxed{u(x) = \cos(\theta)x + (1 - \cos(\theta))\langle x, a\rangle a + \sin(\theta)\, a \wedge x}$$

- Si $\dim(\mathrm{Inv}(u)) = 0$, remarquons d'abord que $u + \mathrm{Id}$ possède un noyau non nul.

En effet, le polynôme caractéristique de u $\det(X\mathrm{Id} - u)$ est de degré 3, donc admet au moins une racine réelle λ. Un élément v non nul du sous-espace propre $E_\lambda = \mathrm{Ker}(\lambda\mathrm{Id} - u)$ vérifie $u(v) = \lambda v$. Comme $\| u(x) \| = \| x \|$, on a nécessairement $\lambda = \pm 1$. Mais λ ne peut être égal à 1 puisqu'il n'y a pas de vecteur invariant non nul, donc $\lambda = -1$.

Soit H le plan orthogonal à v. Comme ci-dessus, la restriction de u à H est une isométrie du plan H n'ayant pas de vecteur invariant non nul. Cette restriction est donc une rotation du plan H, d'angle θ non nul.La matrice de u dans une base orthonormée formée d'une base orthonormée de H et d'un vecteur unitaire orthogonal à H (et donc colinéaire à v) est :

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

On remarque que $\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ $\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ce qui donne une décomposition $r_\theta \circ s_H = s_H \circ r_\theta$ de u en une rotation r_θ d'axe v et la réflexion s_H par rapport à H orthogonal à v, cette décomposition étant commutative. Pour $\theta = 0$, on retrouve simplement s_H.

Ce type d'opérateur est parfois appelé **antirotation**. Un cas particulier est $u = -\mathrm{Id}$, obtenu pour $\theta = \pi$. Si $u \neq -\mathrm{Id}$, le sous-espace propre associé à la valeur propre -1 est de dimension 1 ce qui suffit à montrer que la décomposition de u en une rotation r_θ et une réflexion s_H dont le plan est orthogonal à l'axe de r est unique.

dimension du sous-espace invariant	valeur du déterminant	nature de l'isométrie
3	1	Id
2	–1	réflexion s_H
1	1	rotation r_θ (y compris les retournements s_D)
0	–1	antirotation $r_\theta \circ s_H$

Considérons une rotation en dimension 3 d'axe la droite engendrée par $\boldsymbol{k}$, $\boldsymbol{i}$ et $\boldsymbol{j}$ étant une base du plan, de façon que $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ soit directe. $(\boldsymbol{i}, \boldsymbol{j})$ définit une orientation du plan. Cette orientation définit un angle de rotation θ. Si on avait pris comme vecteur directeur de la droite le vecteur $\boldsymbol{k}' = -\boldsymbol{k}$, alors une nouvelle base directe est $(\boldsymbol{j}, \boldsymbol{i}, \boldsymbol{k}')$, de sorte que l'orientation du plan est modifiée, ainsi que l'angle de la rotation, qui devient $-\theta$. Ainsi, le couple $(\boldsymbol{k}, \theta)$ est défini au signe près. Ce choix est

arbitraire ; on ne peut parler de l'angle de la rotation que si le plan ou l'axe orthogonal au plan a été orienté. Seul le cosinus de l'angle ne dépend pas de l'orientation.
Pour un demi-tour, puisque $-\pi = \pi \bmod 2\pi$, l'orientation de l'axe est sans importance.

Il existe plusieurs méthodes pour déterminer l'angle et l'axe d'une rotation r. En voici trois :

❑ On détermine la droite invariante. On choisit un vecteur v orthogonal à cette droite. On calcule $\cos(\theta) = \dfrac{<r(v), v>}{\| v \|^2}$. Si on choisit θ entre 0 et π, il suffit alors d'orienter l'axe au moyen du vecteur $v \wedge r(v)$.

❑ On détermine la droite invariante. Soit n un vecteur unitaire de cette droite. On choisit un vecteur v orthogonal à cette droite et unitaire. On calcule $w = n \wedge v$ de sorte que (v, w, n) forme une base orthonormée directe de l'espace. On calcule $r(v)$. On a alors $\cos(\theta) = <r(v), v>$ et $\sin(\theta) = <r(v), w>$.

❑ Soit M la matrice de u dans une base orthonormée. Comme $M^{-1} = M^T$, $M - M^{-1} = M - M^T$ est une matrice antisymétrique A, de la forme $\begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$. Pour tout $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, on a :

$$AX = \begin{pmatrix} -cy + bz \\ cx - az \\ -bx + ay \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \wedge \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

On a $\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2\sin(\theta)\Omega$ où Ω est un vecteur unitaire de l'axe de rotation. Pour le voir, il suffit de remarquer que A est la matrice de $r - r^{-1}$ et de vérifier le résultat dans une base orthonormée directe dont Ω est l'un des vecteurs. On obtient donc simultanément $\sin(\theta)$ et Ω (ou bien $-\theta$ et $-\Omega$). $\cos(\theta)$ est donné par la trace $\mathrm{Tr}(M) = 1 + 2\cos(\theta)$. On peut donc en déduire θ. Cette dernière méthode est la plus rapide, mais la moins facile à mémoriser. De plus, dans le cas d'un demi-tour pour lequel $\theta = \pi$, cette méthode ne permet pas de déterminer Ω.
On peut procéder de même pour une antirotation, sauf que, dans ce dernier cas, $\mathrm{Tr}(M) = -1 + 2\cos(\theta)$.

EXEMPLES :

❑ Dans une base orthonormée directe, on donne la matrice M d'une isométrie $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Les vecteurs invariants forme une droite engendrée par $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Il s'agit donc d'une rotation. Soit a de composantes $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, orthogonal à l'axe de rotation. Son image vaut $r(a) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. L'angle entre les deux vecteurs a et r(a) vaut $\dfrac{2\pi}{3}$ (le cosinus vaut $-\dfrac{1}{2}$) en orientant l'axe dans le sens du vecteur $a \wedge r(a) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Si l'on avait pris un ***K*** opposé, on aurait défini l'angle comme étant $-\dfrac{2\pi}{3}$.

On aurait également pu calculer $n = \dfrac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v = \dfrac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $w = \dfrac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $r(v) = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, donc :

$$\cos(\theta) = \langle r(v), v\rangle = -\frac{1}{2} \qquad \text{et} \qquad \sin(\theta) = \langle r(v), w\rangle = \frac{\sqrt{3}}{2}$$

On aurait pu aussi calculer $M - M^T = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ d'où $2\sin(\theta)\Omega = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Ainsi, $\Omega = \dfrac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $\sin(\theta) = \dfrac{\sqrt{3}}{2}$.

❑ Soit $M = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{pmatrix}$. On vérifiera qu'il s'agit d'une matrice de rotation (par exemple, les deux dernières colonnes sont orthonormées et leur produit vectoriel donne la première colonne). On a $M - M^T = \begin{pmatrix} 0 & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} + \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} & 0 & \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} - \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{3} & 0 \end{pmatrix}$ donc $2\sin(\theta)\Omega = \begin{pmatrix} -\frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} + \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} \end{pmatrix}$. On pourra vérifier que la norme de ce vecteur vaut $2\sin(\theta) = \dfrac{\sqrt{54 + 12\sqrt{6} + 12\sqrt{2}}}{6}$. On obtient

$$\cos(\theta) = \frac{1}{2}\left(\frac{\sqrt{6}}{3} + \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{3} - 1\right)$$

en calculant $\mathrm{Tr}(M) = 1 + 2\cos(\theta)$. Les courageux vérifieront que $\cos^2(\theta) + \sin^2(\theta) = 1$

❑ Dans $\mathbf{R}^3$, quelle est la matrice de la rotation u d'angle θ d'axe a de $\mathbf{R}^3$ dans la base canonique, où $a = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ et $\theta = \frac{\pi}{3}$?

Utiliser l'expression $u(x) = (1 - \cos(\theta)) <x, a> a + \cos(\theta)x + \sin(\theta)\, a \wedge x$. On obtient :

$$\begin{pmatrix} \frac{7}{12} & \frac{1}{6} - \frac{\sqrt{2}}{4} & \frac{1}{12} + \frac{\sqrt{2}}{2} \\ \frac{1}{6} + \frac{\sqrt{2}}{4} & \frac{5}{6} & \frac{1}{6} - \frac{\sqrt{2}}{4} \\ \frac{1}{12} - \frac{\sqrt{2}}{2} & \frac{1}{6} + \frac{\sqrt{2}}{4} & \frac{7}{12} \end{pmatrix}$$

❑ Reprenons cette expression de la rotation u d'axe le vecteur unitaire a et d'angle θ :

$$u(x) = \cos(\theta)x + (1 - \cos(\theta))<x, a>a + \sin(\theta)\, a \wedge x$$

Supposons que a et θ varient au cours du temps t, i.e. sont des fonctions d'une variable t qu'on supposera dérivables, et que θ soit nul pour $t = 0$. Cherchons la dérivée de $u(x)$ à l'instant $t = 0$, vitesse à laquelle se déplace x à l'instant initial. On a :

$$\frac{\mathrm{d}}{\mathrm{d}t} u(x) = -\frac{\mathrm{d}\theta}{\mathrm{d}t}\sin(\theta)x + \frac{\mathrm{d}\theta}{\mathrm{d}t}\sin(\theta)<x, a>a + \frac{\mathrm{d}\theta}{\mathrm{d}t}\cos(\theta)\, a \wedge x$$
$$+ (1 - \cos(\theta))<x, \frac{\mathrm{d}a}{\mathrm{d}t}>a + (1 - \cos(\theta))<x, a>\frac{\mathrm{d}a}{\mathrm{d}t} + \sin(\theta)\frac{\mathrm{d}a}{\mathrm{d}t} \wedge x$$

En $t = 0$, $\theta = 0$ et l'égalité précédente se réduit à :

$$\frac{\mathrm{d}}{\mathrm{d}t} u(x)_{|t=0} = \frac{\mathrm{d}\theta}{\mathrm{d}t}\, a \wedge x = \Omega \wedge x$$

avec $\Omega = \frac{\mathrm{d}\theta}{\mathrm{d}t} a$. Le vecteur Ω s'appelle **vecteur instantané de rotation** en $t = 0$ et joue un rôle fondamental en mécanique du solide. Il a été évoqué entre autres en Annexe (*Composition des vitesses et des accélérations en cinématique*) du cours L1/LINEF.PDF, en Annexe (*Utilisation des matrices en physique/Matrice d'inertie*) du cours L1/MATRICES.PDF, ou en Annexe (*Les torseurs*) du cours L2/GEOMEUCL.PDF.

4- Endomorphisme antisymétrique

Dans le paragraphe précédent, nous avons rencontré un exemple d'endomorphisme antisymétrique.

PROPOSITION :

Soit u un endomorphisme de E. *Il y a équivalence entre* :

(i) $\forall x, \forall y, <u(x), y> = -<x, u(y)>$

(ii) $u^* = -u$

(iii) *La matrice de u dans une base orthonormée est antisymétrique.*

(iv) $\forall x, <x, u(x)> = 0$

On dit que u est **antisymétrique**.
De plus, si u est antisymétrique :

(v) $E = \mathrm{Ker}(u) \oplus \mathrm{Im}(u)$

(vi) $\mathrm{rg}(u)$ *est pair*

<u>Démonstration</u> :

❑ (i) ⇔ (ii). En effet, par définition de u^*, $<u(x), y> = <x, u^*(y)>$, donc :

$\forall x, \forall y, <u(x), y> = -<x, u(y)>$

⇔ $\forall x, \forall y, <x, u^*(y)> = <x, -u(y)>$

⇔ $u^* = -u$

❑ (ii) ⇔ (iii). En effet, soit M la matrice de u dans une base orthonormée. Alors la matrice de u^* dans cette base est M^T et :

$u^* = -u \Leftrightarrow M^T = -M \Leftrightarrow$ M est antisymétrique.

❑ (i) ⇒ (iv) est trivial en prenant $x = y$.

❑ (iv) ⇒ (i) : On applique (iv) sur $x + y$.

$\forall x, \forall y, <x + y, u(x + y)> = 0$

⇒ $\forall x, \forall y, <x, u(x)> + <x, u(y)> + <y, u(x)> + <y, u(y)> = 0$

⇒ $\forall x, \forall y, <x, u(y)> + <y, u(x)> = 0$

⇒ $\forall x, \forall y, <x , u(y)> = -<u(x) , y>$

❑ (v) : Si $x \in \mathrm{Ker}(u) \cap \mathrm{Im}(u)$, alors il existe y tel que $x = u(y)$, et $u(x) = 0$. Donc :

$<x, u(y)> = -<u(x), y> \Rightarrow \| x \|^2 = 0 \Rightarrow x = 0$

Donc $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ sont orthogonaux. En particulier, ils sont en somme directe et le théorème du rang permet de montrer que leur somme a pour dimension dim(E).

❑ (vi) : Soit M une matrice de u dans une base orthonormée formée d'une base orthonormée de $\mathrm{Ker}(u)$ et d'une base orthonormée de $\mathrm{Im}(u)$ (ce qu'on peut faire d'après (v)). M est diagonale par bloc de la forme $\begin{pmatrix} O & O \\ O & N \end{pmatrix}$, avec N antisymétrique car M l'est. La taille de la matrice N est $r \times r$ avec $r = \dim(\mathrm{Im}(u)) = \mathrm{rg}(u)$. De plus $\mathrm{rg}(M) = r = \mathrm{rg}(N)$, donc N est inversible. On a alors :

$\det(N) = \det(N^T) = \det(-N) = (-1)^r \det(N)$

donc $(-1)^r = 1$

et r est pair.

EXEMPLE :

❑ Comme on l'a vu, en dimension 3, la matrice M est de la forme $\begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$ et si on prend Ω de composantes $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, alors, pour tout x de E, $u(x) = \Omega \wedge x$. La relation d'antisymétrie s'exprime alors en dimension 3 sous la forme :

$$\forall x, \forall y, <\Omega \wedge x, y> = -<x, \Omega \wedge y>$$

et on reconnaît le caractère alterné du produit mixte (déterminant des trois vecteurs dans une base orthonormée directe. Revoir au besoin le chapitre *Espaces euclidiens* L1/ESPEUCL.PDF) :

$$\det(\Omega, x, y) = - \det(x, \Omega, y)$$

En physique, la mécanique du solide fait un usage intensif des endomorphismes antisymétriques (voir *Torseur* dans L2/GEOMEUC.PDF).
Si Ω est non nul, $\mathrm{Ker}(u) = \mathrm{Vect}(\Omega)$, $\mathrm{Im}(u) = \mathrm{Vect}(\Omega)^{\perp}$, $\mathrm{rg}(u) = 2$.
Si $\Omega = 0$, $\mathrm{Ker}(u) = \mathrm{E}$, $\mathrm{Im}(u) = 0$, $\mathrm{rg}(u) = 0$.

❑ Une application u d'un espace euclidien E dans lui-même vérifiant la propriété (i) est nécessairement linéaire. En effet, pour tout x et y de E et tout λ de $\mathbf{R}$:

$$<u(\lambda x), y> = - <\lambda x, u(y)> = - \lambda<x, u(y)> = \lambda<u(x), y> = <\lambda u(x), y>$$

$\Rightarrow \quad \forall y, <u(\lambda x) - \lambda u(x), y> = 0$

$\Rightarrow \quad u(\lambda x) - \lambda u(x) = 0$

On procéde de même pour pour la somme.
La même remarque s'applique aux endomorphismes symétriques, objets du paragraphe suivant.

5- Endomorphisme symétrique

On peut définir de même les endomorphismes symétriques.

PROPOSITION :
Soit u un endomorphisme de E. *Il y a équivalence entre* :

(i) $\forall x, \forall y, <u(x), y> = <x, u(y)>$

(ii) $u^* = u$

(iii) *La matrice de u dans une base orthonormée est symétrique*.

On dit que u est **symétrique** *ou* **autoadjoint**.

Démonstration :
❑ (i) $\Leftrightarrow$ (ii). En effet :

$\quad \forall x, \forall y, <u(x), y> = <x, u(y)>$

$\Leftrightarrow \quad \forall x, \forall y, <x, u^*(y)> = <x, u(y)>$

$\Leftrightarrow \quad u^* = u$

❑ (ii) $\Leftrightarrow$ (iii). En effet, soit M la matrice de u dans une base orthonormée. Alors :

$$u^* = u \Leftrightarrow \mathrm{M}^{\mathrm{T}} = \mathrm{M} \Leftrightarrow \mathrm{M} \text{ est symétrique.}$$

Dans la proposition énoncée, la formulation « *une base orthonormée* » du iii) peut désigner aussi bien une base particulière qu'une base quelconque. En effet, l'implication iii) $\Rightarrow$ ii) utilise l'existence de M symétrique dans une base particulière, la réciproque ii) $\Rightarrow$ iii) montre la symétrie de M dans toute base orthonormée. Ainsi, si la matrice de u est symétrique dans une base orthonormée donnée, elle restera symétrique dans toute base orthonormée. D'ailleurs, cette dernière propriété peut se montrer indépendamment. En effet, si M est la matrice de u symétrique dans une base orthonormée, sa matrice dans une autre base orthonormée est $\mathrm{N} = \mathrm{P}^{-1}\mathrm{MP}$, avec P matrice de passage d'une base orthonormée dans une autre. Autrement dit, P est une matrice orthogonale vérifiant donc $\mathrm{P}^{-1} = \mathrm{P}^{\mathrm{T}}$. Donc $\mathrm{N} = \mathrm{P}^{\mathrm{T}}\mathrm{MP}$ et cette matrice est bien symétrique.

De même que les matrices symétriques constituent un sous-espace vectoriel de $\mathcal{M}_n(\mathbf{R})$, les endomorphismes symétriques constituent un sous-espace vectoriel de L(E).

Les mêmes remarques s'appliquent pour les endomorphismes antisymétriques.

EXEMPLES :
❑ Projecteurs orthogonaux :

p est un projecteur orthogonal $\Leftrightarrow p^2 = p$ *et p est symétrique.*

La condition est nécessaire. Si p est un projecteur orthogonal, il existe une base orthonormée telle que la matrice de p soit $M = \begin{pmatrix} I_m & O_{m,n-m} \\ O_{n-m,m} & O_{n-m,n-m} \end{pmatrix}$ où I_m est la matrice identité à m lignes et $O_{k,q}$ la matrice nulle à k lignes et q colonnes. On a alors clairement $M^2 = M$ et $M^T = M$.
Réciproquement, si $p^2 = p$ alors p est un projecteur (pas nécessairement orthogonal) sur Im(p) parallèlement à Ker(p). Il suffit de montrer que, si p est symétrique, on a Im(p) $\perp$ Ker(p). Soit $p(x)$ élément de Im(p) et y élément de Ker(p). On a alors :

$<p(x), y> = <x, p(y)>$ car p est supposée symétrique

$= <x, 0_E> = 0$

Tel est le cas par exemple dans $\mathbf{R}^3$ de l'endomorphisme p de matrice symétrique $\frac{1}{14}\begin{pmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}$. On vérifiera que p^2 = Id. p étant symétrique, p est un projecteur orthogonal. On pourra vérifier qu'il s'agit du projecteur orthogonal sur le plan $x + 2y + 3z = 0$.

❑ Symétries orthogonales :

s est une symétrie orthogonale $\Leftrightarrow s^2 = \mathrm{Id}$ *et s est symétrique.*

La condition est nécessaire. Si s est une symétrie orthogonale, il existe une base orthonormée telle que la matrice de s soit $M = \begin{pmatrix} I_m & O_{m,n-m} \\ O_{n-m,m} & -I_{n-m} \end{pmatrix}$ où I_m est la matrice identité à m lignes et $O_{k,p}$ la matrice nulle à k lignes et p colonnes. On a alors clairement $M^2 = I_n$ et $M^T = M$.
Réciproquement, si $s^2 = s$ alors s est une symétrie (pas nécessairement orthogonale) par rapport à Ker(s – Id) parallèlement à Ker(s + Id). Il suffit de montrer que, si s est symétrique, on a Ker(s – Id) $\perp$ Ker(s + Id). Soit x élément de Ker(s – Id) et y élément de Ker(s + Id). On a alors :

$<s(x),y> = <x,s(y)>$ car s est supposée symétrique

$\Rightarrow$ $<x,y> = -<x,y>$

$\Rightarrow$ $<x,y> = 0$

EXEMPLE :

❑ Dans $\mathbf{R}^3$, l'endomorphisme s de matrice symétrique $\frac{1}{9}\begin{pmatrix} 7 & -4 & 4 \\ -4 & 1 & 8 \\ 4 & 8 & 1 \end{pmatrix}$. est une symétrie oryhogonale

On vérifiera en effet que s^2 = Id. s étant symétrique, s est une symétrie orthogonale.
On pourra vérifier qu'il s'agit de la symétrie orthogonale par rapport au plan $x + 2y - 2z = 0$.

PROPOSITION :
Si u est symétrique et F *un sous-espace vectoriel de* E *stable par u, alors* $F^\perp$ *est stable par u.*

<u>Démonstration</u> :

❑ Pour tout endomorphisme u, si F est stable par u, alors $F^\perp$ est stable par u^*. Ici $u^* = u$, donc $F^\perp$ est stable par u.

6- Réduction des endomorphismes symétriques

Les endomorphismes symétriques jouissent de la propriété suivante :

PROPOSITION

Soit u un endomorphisme symétrique d'un espace euclidien. Il existe une base orthonormée de vecteurs propres.

Variantes :

Tout endomorphisme symétrique est diagonalisable dans une base orthonormée.

Toute matrice symétrique réelle M est de la forme $P^{-1}DP = P^TDP$, où D est diagonale et P orthogonale.

Démonstration :

Soit M la matrice (symétrique) de u dans une base orthonormée. Soit n la dimension de E :

❑ On montre d'abord que M possède au moins une valeur propre réelle.

Démonstration 1 :

On se place dans $\mathbf{C}^n$. Soit λ une valeur propre complexe de M, de vecteur propre X élément de $\mathbf{C}^n$, de composantes x_i. On a :

$$MX = \lambda X$$

$\Rightarrow \quad \overline{X}^T MX = \lambda \overline{X}^T X$ en multipliant à gauche par $\overline{X}^T$

On prend la transposée de cette égalité, et son conjugué :

$\Rightarrow \quad \overline{X}^T \overline{M}^T X = \overline{\lambda}\, \overline{X}^T X$

Or M est à coefficients réels et symétrique donc $\overline{M}^T = M$

$\Rightarrow \quad \overline{X}^T MX = \overline{\lambda}\, \overline{X}^T X$

$\Rightarrow \quad \lambda \overline{X}^T X = \overline{\lambda}\, \overline{X}^T X$ car $MX = \lambda X$

$\Rightarrow \quad \lambda = \overline{\lambda}$ puisque $\overline{X}^T X = \sum_{i=1}^n |x_i|^2 \neq 0$

Démonstration 2 :

On se place dans $\mathbf{C}^n$. Soit $\lambda = a + ib$ une valeur propre complexe de M, de vecteur propre $X = U + iV$, avec a et b réels et U et V éléments de $\mathbf{R}^n$ non tous deux nuls. On a :

$$MX = \lambda X$$

$\Rightarrow \quad MU + iMV = aU - bV + i(bU + aV)$

$$\Rightarrow \quad \begin{cases} MU = aU - bV \\ MV = bU + aV \end{cases}$$

Considérons $U^TMV = bU^TU + aU^TV$, obtenue en multipliant la deuxième équation par U^T. Utilisant le fait que M est symétrique, la transposée de cette dernière égalité est :

$$V^TMU = bU^TU + aV^TU$$

mais $\quad V^TMU = aV^TU - bV^TV$ obtenu en multipliant la première équation par V^T

Il en résulte que $b(U^TU + V^TV) = 0$ donc $b = 0$ puisque $U^TU + V^TV = \| U \|^2 + \| V \|^2$ est non nul. λ est donc réelle.

<u>Démonstration 3</u> :
Cette troisième démonstration utilise le fait que, dans un espace vectoriel normé de dimension finie, tout endomorphisme u est lipschitzien de rapport $C = \{\text{Sup } \{\| u(x) \|, x \in S\}$ où S est la sphère unité, et toute application continue sur un fermé borné admet un maximum (voir le chapitre L2/EVNORME.PDF). S étant fermé borné et u continue sur S, cela prouve que C est le maximum de u sur S et qu'il existe un vecteur x de S tel que $\| u(x) \| = C$. Soit y unitaire tel que $u(x) = Cy$. On a alors :

$$\begin{aligned} C &= \langle u(x), y\rangle \\ &= \langle x, u(y)\rangle && \text{car } u \text{ est symétrique} \\ &\leq \| x \| \, \| u(y) \| && \text{d'après l'inégalité de Cauchy-Schwartz} \\ &\leq C \| x \| \, \| y \| && \text{car } u \text{ est lipschitzienne de rapport C} \\ &= C && \text{car } x \text{ et } y \text{ sont unitaires} \end{aligned}$$

Les deux inégalités ci-dessus sont donc des égalités. Donc $\langle x, u(y)\rangle = \| x \| \, \| u(y) \|$ (et donc $u(y)$ est colinéaire à x et de même sens) et $\| u(y) \| = C$, donc $u(y) = Cx$. Donc :

$$u(x + y) = u(x) + u(y) = C(y + x)$$

et $\quad u(x - y) = u(x) - u(y) = C(y - x)$

L'un des deux vecteur $x + y$ ou $x - y$ étant non nul, l'un des deux est vecteur propre, de valeur propre C ou – C.

❑ On montre ensuite que E est la somme de ses sous-espaces propres. Soit $F = \bigoplus_{\lambda \in Sp(u)} E_\lambda$, somme des sous-espaces propres de u. Il s'agit de montrer que F = E. Si ce n'est pas le cas, $E = F \oplus F^\perp$ avec $F^\perp$ non réduit à {0} et stable par u car l'orthogonal d'un sous-espace vectoriel stable par un endomorphisme symétrique est lui-même stable par cet endomorphisme. Mais $u|_{F^\perp}$ est lui-même symétrique :

$$\forall x \in F^\perp, \forall y \in F^\perp, \langle u|_{F^\perp}(x), y\rangle = \langle u(x), y\rangle = \langle x, u(y)\rangle = \langle x, u|_{F^\perp}(y)\rangle$$

donc admet aussi au moins une valeur propre réelle, qui est aussi valeur propre de u. Donc $F^\perp$ contient au moins un vecteur propre de u. Ceci est absurde puisque tous les vecteurs propres de u sont éléments de F.
Donc E est la somme de ses sous-espaces propres, donc u est diagonalisable et toutes ses valeurs propres sont réelles.

❑ En outre, les sous-espaces propres sont orthogonaux deux à deux.
Si λ et μ sont deux valeurs propres distinctes et si x et y sont respectivement deux vecteurs propres, on a :

$$\langle u(x), y\rangle = \lambda\langle x, y\rangle = \langle x, u(y)\rangle = \mu\langle x, y\rangle$$

$\Rightarrow \quad \langle x, y\rangle = 0$ puisque $\lambda \neq \mu$.

En choisissant dans chaque sous-espace propre une base orthonormée, on obtient une base orthonormée de E constituée de vecteurs propres de u.

On peut remarquer que, réciproquement, tout endomorphisme ou toute matrice A diagonalisable dans une base orthonormée est nécessairement symétrique. En effet, il existe une matrice de passage

P orthogonale et une matrice diagonale D telle que $A = PDP^{-1}$. On a alors, puisque $D^T = D$ et $P^T = P^{-1}$:

$$A^T = P^{-1T}\, D^T\, P^T = PDP^{-1} = A$$

donc A est symétrique.

EXEMPLES :

❑ Soit $M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Les valeurs propres de M sont 3 associé au vecteur propre $\dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, et – 1, associé au vecteur propre $\dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. On remarque que les deux vecteurs propres sont bien orthogonaux.

❑ Soit (E, < , >) et (F, (,)) deux espaces euclidiens, dim(E) = p, dim(F) = n, et f une application linéaire de E dans F, de rang r. Montrons que :

si f est injective, l'image par f de la sphère unité de E est un ellipsoïde dans Im(f),

si f n'est pas injective, l'image par f de la sphère unité de E est le domaine borné limité par un ellipsoïde dans Im(f).

(voir le chapitre L2/CONIQUES.PDF pour la définition d'un ellipsoïde dans $\mathbb{R}^3$. Nous laissons le soin au lecteur de généraliser cette notion à toute dimension. En dimension 2, c'est une ellipse).

Posons $g = f^* \circ f \in L(E)$, où f^* est l'adjoint de f. On a :

$$g^* = (f^* \circ f)^* = f^* \circ f^{**} = f^* \circ f = g$$

donc g est un endomorphisme symétrique de E, donc diagonalisable dans une base orthonormée $(e_1, ..., e_p)$ de E. Notons $\lambda_1, ..., \lambda_p$ les valeurs propres correspondantes.

On a par ailleurs, Ker(g) = Ker(f). L'inclusion $\supset$ est triviale. Réciproquement :

$x \in \mathrm{Ker}(g)$

$\Rightarrow \quad 0 = g(x) = f^*(f(x))$

$\Rightarrow \quad 0 = \langle x, f^*(f(x))\rangle = (f(x), f(x))$

$\Rightarrow \quad f(x) = 0$

$\Rightarrow \quad x \in \mathrm{Ker}(f)$

Puisque Ker(g) = Ker(f), le théorème du rang permet d'en déduire que le rang de g est r. Quitte à permuter les vecteurs $e_1, ..., e_p$, on peut donc supposer que $\lambda_1, ..., \lambda_r$ sont non nuls et que $\lambda_{r+1}, ..., \lambda_p$ sont nuls.

Pour $1 \leq k \leq r$, on a :

$$\| f(e_k) \|^2 = (f(e_k), f(e_k)) = \langle e_k, f^*(f(e_k))\rangle = \langle e_k, g(e_k)\rangle = \lambda_k \langle e_k, e_k\rangle = \lambda_k$$

donc $\lambda_k > 0$ et $\| f(e_k) \| = \sqrt{\lambda_k}$. On peut donc définir un vecteur ε_k de F, unitaire, tel que $f(e_k) = \sqrt{\lambda_k}\, \varepsilon_k$.

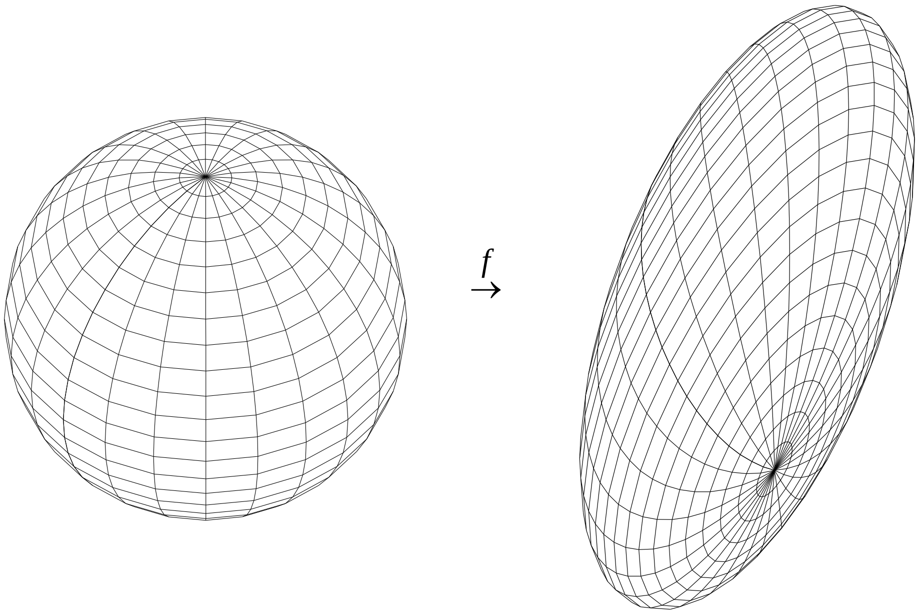
Les ε_k sont orthogonaux deux à deux, car, si $i \neq k$, on a :

$$(\varepsilon_i, \varepsilon_k) = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_k}} (f(e_i), f(e_k)) = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_k}} \langle e_i, f^*(f(e_k)\rangle = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_k}} \langle e_i, g(e_k)\rangle$$

$$= \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_k}} \lambda_k \langle e_i, e_k\rangle = 0 \qquad \text{puisque les } e_k \text{ sont orthogonaux entre eux.}$$

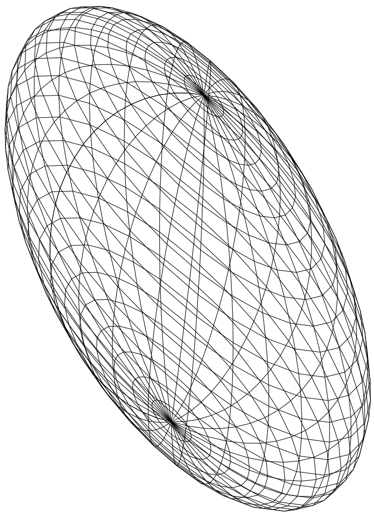
Il en résulte que $(\varepsilon_1, ..., \varepsilon_r)$ forment une base orthonormée de Im(f). Dès lors :

- Si $r = p$ (f est injective d'après le théorème du rang), soit $x = \sum_{i=1}^{p} x_i e_i$ avec $\sum_{i=1}^{p} x_i^2 = 1$. x est un élément de la sphère unité de E. On a $f(x) = \sum_{i=1}^{p} x_i\, f(e_i) = \sum_{i=1}^{p} x_i \sqrt{\lambda_i}\varepsilon_i = \sum_{i=1}^{p} y_i \varepsilon_i$, avec $y_i = x_i\sqrt{\lambda_i}$. On a donc $f(x)$ élément de l'ellipsoïde d'équation $\sum_{i=1}^{p} \dfrac{y_i^2}{\lambda_i} = 1$, et inversement, tout y élément de cet ellipsoïde peut s'écrire sous la forme $y = f(x)$ avec x élément de la sphère unité de E en prenant $x_i = \dfrac{y_i}{\sqrt{\lambda_i}}$. Ci-dessous, on a représenté la sphère unité de $\mathbf{R}^3$ et son image par l'endomorphisme f de matrice $\begin{pmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & 3 & 1 \end{pmatrix}$, de rang 3 :



- Si $r < p$ (f est non injective), le même calcul que précédemment conduit à $f(x) = \sum_{i=1}^{r} y_i \varepsilon_i$ avec $y_i = x_i\sqrt{\lambda_i}$, et on a seulement $\sum_{i=1}^{r} \dfrac{y_i^2}{\lambda_i} = \sum_{i=1}^{r} x_i^2 \leq 1$. Donc $f(x)$ appartient au domaine borné limité par l'ellipsoïde d'équation $\sum_{i=1}^{r} \dfrac{y_i^2}{\lambda_i} = 1$. Inversement, tout y de ce domaine est de la forme $y = f(x)$ avec x élément de la sphère unité, en prenant $x_i = \dfrac{y_i}{\sqrt{\lambda_i}}$ pour $1 \leq i \leq r$ et par exemple x_{r+1} tel que

${x_{r+1}}^2 = 1 - \sum_{i=1}^{r} \frac{{y_i}^2}{\lambda_i}$, puis $x_{r+2} = ... = x_p = 0$. Ci-dessous, on a représenté l'image de la sphère unité de $\mathbf{R}^3$ par l'endomorphisme f de matrice $\begin{pmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & -2 & 1 \end{pmatrix}$, de rang 2. Contrairement à l'image précédente, il ne s'agit pas d'un ellipsoïde dans $\mathbf{R}^3$, mais d'une ellipse dans le plan Im(f) d'équation $x - y - z = 0$, entièrement remplie par les tracés des images des méridiens et des parallèles de la sphère unité. On peut voir cette image comme un ellipsoïde que l'on a écrasé en une ellipse sur le plan Im(f) :



Annexe I : Utilisation d'opérateurs symétriques en physique ou en SII

1- Opérateur d'inertie

Soit P un point de masse m se déplaçant à la vitesse $\boldsymbol{V}$ par rapport à un référentiel et O un point de ce référentiel. On appelle **moment cinétique** de P par rapport à O le vecteur :

$$\boldsymbol{L}(\mathrm{O}) = \boldsymbol{OP} \wedge m\boldsymbol{V}$$

L'intérêt du moment cinétique tient dans le fait que, lorsque O est fixe dans le référentiel considéré, sa dérivée par rapport au temps, appelée **moment dynamique** de P par rapport à O, est égal, d'après le principe fondamental de la dynamique, aux moments des forces $\boldsymbol{F}$ appliquées en P par rapport à O (y compris les forces d'inertie si le référentiel n'est pas galiléen) :

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{L}(\mathrm{O}) &= \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{OP} \wedge m\boldsymbol{V} + \boldsymbol{OP} \wedge \frac{\mathrm{d}}{\mathrm{d}t} m\boldsymbol{V} \\ &= \boldsymbol{V} \wedge m\boldsymbol{V} + \boldsymbol{OP} \wedge m\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} \\ &= \boldsymbol{OP} \wedge m\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} \\ &= \boldsymbol{OP} \wedge \boldsymbol{F} \end{aligned}$$

Supposons que P effectue un mouvement de rotation autour d'un axe passant par O. La vitesse $\boldsymbol{V}$ est alors de la forme $\boldsymbol{V} = \boldsymbol{\Omega} \wedge \boldsymbol{OP}$, où $\boldsymbol{\Omega}$ est un vecteur appelé vecteur instantané de rotation. L'axe $(\mathrm{O}, \boldsymbol{\Omega})$ est l'axe instantané de rotation. On a alors :

$$\boldsymbol{L}(\mathrm{O}) = m\, \boldsymbol{OP} \wedge (\boldsymbol{\Omega} \wedge \boldsymbol{OP})$$

On introduit alors un opérateur $\boldsymbol{J}$, appelé **opérateur d'inertie**, qui, à tout vecteur $\boldsymbol{u}$, associe le vecteur :

$$\boldsymbol{J}(\boldsymbol{u}) = m\, \boldsymbol{OP} \wedge (\boldsymbol{u} \wedge \boldsymbol{OP})$$
$$= m\, \mathrm{OP}^2 \boldsymbol{u} - m <\boldsymbol{u} \mid \boldsymbol{OP}> \boldsymbol{OP}$$

en utilisant la formule du double produit vectoriel

$\boldsymbol{a} \wedge (\boldsymbol{b} \wedge \boldsymbol{c}) = <\boldsymbol{a}, \boldsymbol{c}> \boldsymbol{b} - <\boldsymbol{a}, \boldsymbol{b}> \boldsymbol{c}$

voir L1/DETERMNT.PDF

$$= m\, (\boldsymbol{OP} \wedge \boldsymbol{u}) \wedge \boldsymbol{OP}$$

de sorte que $\boldsymbol{L}(\mathrm{O}) = \boldsymbol{J}(\boldsymbol{\Omega})$.

En ce qui concerne un solide, on opère de même en sommant au moyen d'une intégrale triple sur tous les points du solide. L'opérateur d'inertie prend alors la forme :

$$\boldsymbol{J}(\boldsymbol{u}) = \iiint_S \boldsymbol{OP} \wedge (\boldsymbol{u} \wedge \boldsymbol{OP})\, \mathrm{d}m$$
$$= \iiint_S \mathrm{OP}^2 \boldsymbol{u} - <\boldsymbol{u} \mid \boldsymbol{OP}> \boldsymbol{OP}\, \mathrm{d}m$$

$\boldsymbol{J}$ est clairement linéaire et est un opérateur symétrique. En effet :

$$<\boldsymbol{J}(\boldsymbol{u}) \mid \boldsymbol{v}> = \iiint_S \mathrm{OP}^2 <\boldsymbol{u} \mid \boldsymbol{v}> - <\boldsymbol{u} \mid \boldsymbol{OP}><\boldsymbol{v} \mid \boldsymbol{OP}>\, \mathrm{d}m$$

qui est symétrique par rapport à $\boldsymbol{u}$ et $\boldsymbol{v}$, donc égal à $<\boldsymbol{u} \mid \boldsymbol{J}(\boldsymbol{v})>$. Cet opérateur est donc représenté par une matrice 3×3 symétrique dans une base orthonormée. La matrice de $\boldsymbol{J}$ s'écrit donc $\begin{pmatrix} \mathrm{A} & -\mathrm{F} & -\mathrm{E} \\ -\mathrm{F} & \mathrm{B} & -\mathrm{D} \\ -\mathrm{E} & -\mathrm{D} & \mathrm{C} \end{pmatrix}$

avec :

$\mathrm{A} = <\boldsymbol{J}(\boldsymbol{i}) \mid \boldsymbol{i}> = \iiint_S \mathrm{OP}^2 - <\boldsymbol{i} \mid \boldsymbol{OP}>^2\, \mathrm{d}m = \iiint_S (y^2 + z^2)\, \mathrm{d}m$, appelé **moment d'inertie** par rapport à l'axe $(\mathrm{O}, \boldsymbol{i})$

$\mathrm{B} = \iiint_S (x^2 + z^2)\, \mathrm{d}m$, moment d'inertie par rapport à l'axe $(\mathrm{O},\boldsymbol{j})$

$\mathrm{C} = \iiint_S (x^2+y^2)\, \mathrm{d}m$, moment d'inertie par rapport à l'axe $(\mathrm{O},\boldsymbol{k})$

$\mathrm{D} = -<\boldsymbol{J}(\boldsymbol{k}) \mid \boldsymbol{j}> = \iiint_S yz\, \mathrm{d}m$, quantité appelée **produit d'inertie**.

$$\mathrm{E} = \iiint_S xz\, \mathrm{d}m$$

$$\mathrm{F} = \iiint_S xy\, \mathrm{d}m$$

Etant symétrique, elle est diagonalisable dans un repère orthonormé, dont les axes sont appelés **axes principaux d'inertie**.

En général, bien que $\boldsymbol{L}(\mathrm{O}) = \boldsymbol{J}(\boldsymbol{\Omega})$, le moment cinétique $\boldsymbol{L}(\mathrm{O})$ n'est pas colinéaire à $\boldsymbol{\Omega}$ et n'est donc pas porté par l'axe de rotation. Cela ne se produit que si $\boldsymbol{\Omega}$ est vecteur propre de l'opérateur d'inertie. C'est le cas si $\boldsymbol{\Omega}$ est porté par un axe principal d'inertie. Plus précisément, si on prend comme base orthonormée des vecteurs principaux d'inertie $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ de sorte que la matrice de $\boldsymbol{J}$ dans cette base est diaginale, et si A, B, C sont les moments d'inertie correspondant, on appelle **ellipsoïde d'inertie** la surface d'équation $\mathrm{A}x^2 + \mathrm{B}y^2 + \mathrm{C}z^2 = 1$. Soit $\boldsymbol{\Omega}$ un vecteur non nul quelconque telle que l'axe $(\mathrm{O}, \boldsymbol{\Omega})$ coupe cette surface en un point P. Alors $\boldsymbol{J}(\boldsymbol{\Omega})$ est colinéaire au vecteur normal au plan tangent en P à l'ellipsoïde. En effet, si $\boldsymbol{\Omega}$ a pour composantes $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$, alors $\mathrm{P} = (\lambda u, \lambda v, \lambda w)$ avec λ tel que $\lambda^2(\mathrm{A}u^2 + \mathrm{B}v^2 + \mathrm{C}w^2) = 1$ (mais peu importe sa valeur). En le point P, le plan tangent est normal au gradient de $\mathrm{A}x^2 + \mathrm{B}y^2 + \mathrm{C}z^2$ en P. Ce gradient a pour composantes $\begin{pmatrix} 2\mathrm{A}x \\ 2\mathrm{B}y \\ 2\mathrm{C}z \end{pmatrix}$, donc un vecteur normal est colinéaire à $\begin{pmatrix} 2\mathrm{A}\lambda u \\ 2\mathrm{B}\lambda v \\ 2\mathrm{C}\lambda w \end{pmatrix} = 2\lambda \begin{pmatrix} \mathrm{A}u \\ \mathrm{B}v \\ \mathrm{C}w \end{pmatrix}$. Mais $\begin{pmatrix} \mathrm{A}u \\ \mathrm{B}v \\ \mathrm{C}w \end{pmatrix}$ sont précisément les composantes de $\boldsymbol{J}(\boldsymbol{\Omega})$.

A noter que, si O est un point du solide, supposé fixe dans le référentiel, l'énergie cinétique du solide est :

$$
\begin{aligned}
\mathrm{E_c} &= \frac{1}{2} \iiint_{\mathrm{S}} \boldsymbol{V}^2 \,\mathrm{d}m = \frac{1}{2} \iiint_{\mathrm{S}} (\boldsymbol{\Omega} \wedge \boldsymbol{OP})^2 \,\mathrm{d}m \\
&= \frac{1}{2} \iiint_{\mathrm{S}} [\boldsymbol{\Omega} \wedge \boldsymbol{OP}, \boldsymbol{\Omega}, \boldsymbol{OP}] \,\mathrm{d}m \qquad \text{où } [\ ,\ ,\ ] \text{ dénote le produit mixte} \\
&= \frac{1}{2} \iiint_{\mathrm{S}} [\boldsymbol{\Omega}, \boldsymbol{OP}, \boldsymbol{\Omega} \wedge \boldsymbol{OP}] \,\mathrm{d}m = \frac{1}{2} \langle \boldsymbol{\Omega}, \iiint_{\mathrm{S}} \boldsymbol{OP} \wedge (\boldsymbol{\Omega} \wedge \boldsymbol{OP}) \,\mathrm{d}m \rangle \\
&= \frac{1}{2} \langle \boldsymbol{\Omega} \mid \boldsymbol{J}(\boldsymbol{\Omega}) \rangle = \frac{1}{2} \mathrm{J}\Omega^2 \text{ où J est le moment d'inertie par rapport à l'axe porté par } \boldsymbol{\Omega}.
\end{aligned}
$$

L'opérateur $\boldsymbol{J}$ vérifie le **théorème de Koenig** : *l'opérateur d'inertie du solide par rapport à* O *est égal à la somme de l'opérateur d'inertie du solide par rapport à son centre d'inertie* G *et de l'opérateur d'inertie de* G *affecté de la masse totale* M *du solide par rapport à* O. En effet :

$$
\begin{aligned}
\boldsymbol{J}(\boldsymbol{u}) &= \iiint_{\mathrm{S}} \boldsymbol{OP} \wedge (\boldsymbol{u} \wedge \boldsymbol{OP}) \,\mathrm{d}m \\
&= \iiint_{\mathrm{S}} (\boldsymbol{OG} + \boldsymbol{GP}) \wedge (\boldsymbol{u} \wedge (\boldsymbol{OG} + \boldsymbol{GP})) \,\mathrm{d}m \\
&= \iiint_{\mathrm{S}} \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG}) \,\mathrm{d}m + \iiint_{\mathrm{S}} \boldsymbol{GP} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG}) \,\mathrm{d}m \\
&\qquad + \iiint_{\mathrm{S}} \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{GP}) \,\mathrm{d}m + \iiint_{\mathrm{S}} \boldsymbol{GP} \wedge (\boldsymbol{u} \wedge \boldsymbol{GP}) \,\mathrm{d}m
\end{aligned}
$$

Or $\iiint_S \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})\, \mathrm{d}m = \mathrm{M}\, \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})$

et $\iiint_S \boldsymbol{GP} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})\, \mathrm{d}m = \left(\iiint_S \boldsymbol{GP}\, \mathrm{d}m\right) \wedge (\boldsymbol{u} \wedge \boldsymbol{OG}) = 0$ puisque $\left(\iiint_S \boldsymbol{GP}\, \mathrm{d}m\right) = 0$ par définition du centre d'inertie. De même pour $\iiint_S \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{GP})\, \mathrm{d}m$. Il reste donc :

$$\boldsymbol{J}(\boldsymbol{u}) = \iiint_S \boldsymbol{GP} \wedge (\boldsymbol{u} \wedge \boldsymbol{GP})\, \mathrm{d}m + \mathrm{M}\, \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})$$

soit $\boldsymbol{J}_{\mathrm{O}}(\boldsymbol{u}) = \boldsymbol{J}_{\mathrm{G}}(\boldsymbol{u}) + \mathrm{M}\, \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})$

On en déduit le **théorème de Huygens** pour les moments d'inertie par rapport à un axe $(\mathrm{O},\boldsymbol{u})$, $\boldsymbol{u}$ étant unitaire :

$$\mathrm{J} = \langle \boldsymbol{u} \mid \boldsymbol{J}(\boldsymbol{u})\rangle = \langle \boldsymbol{u}, \iiint_S \boldsymbol{GP} \wedge (\boldsymbol{u} \wedge \boldsymbol{GP})\, \mathrm{d}m\rangle + \mathrm{M} \langle \boldsymbol{u} \mid \boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})\rangle$$

Le moment d'inertie d'un solide par rapport à un axe (O, $\boldsymbol{u}$) *est égal à la somme du moment d'inertie du solide par rapport à l'axe* (G, $\boldsymbol{u}$) *et du moment d'inertie de* G *affecté de la masse* M *par rapport à l'axe* (O, $\boldsymbol{u}$).

EXEMPLE 1 :

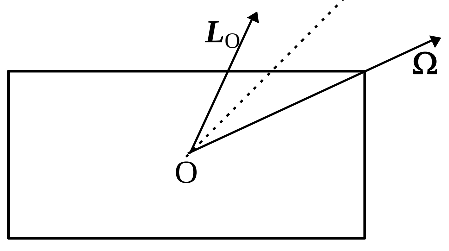
❑ Considérons une plaque de masse m dans le plan Oxy, définie par $-a \leq x \leq a$, $-b \leq y \leq b$. Dans ce cas, $\boldsymbol{J}_{\mathrm{O}}$ est définie par une intégrale double. On vérifiera que la matrice d'inertie est :

$$\begin{pmatrix} \frac{mb^2}{3} & 0 & 0 \\ 0 & \frac{ma^2}{3} & 0 \\ 0 & 0 & \frac{m(a^2+b^2)}{3} \end{pmatrix}.$$

Supposons par exemple que la plaque tourne autour d'une de ses diagonales. On a $\boldsymbol{\Omega} = \Omega\, (\cos(\theta)\, \boldsymbol{i} + \sin(\theta)\, \boldsymbol{j})$ avec $\tan(\theta) = \frac{b}{a}$. Le moment cinétique est donc :

$$\boldsymbol{L}(\mathrm{O}) = \frac{\Omega m}{3} (b^2\cos(\theta)\, \boldsymbol{i} + a^2\sin(\theta)\, \boldsymbol{j})$$

La tangente de l'angle $(\boldsymbol{L}(\mathrm{O}),\boldsymbol{i})$ vaut $\frac{a^2\sin(\theta)}{b^2\cos(\theta)} = \frac{1}{\tan(\theta)} = \tan(\frac{\pi}{2} - \theta)$, autrement dit la bissectrice des droites O$\boldsymbol{i}$ et O$\boldsymbol{j}$ est également bissectrice de $\boldsymbol{\Omega}$ et $\boldsymbol{L}$(O).



EXEMPLE 2 :

❑ Si la plaque est définie par $0 \le x \le 2a$, $0 \le y \le 2b$, on a, en utilisant le théorème de Kœnig :

$$\boldsymbol{J}_{\mathrm{O}}(\boldsymbol{u}) = \boldsymbol{J}_{\mathrm{G}}(\boldsymbol{u}) + m\,\boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG})$$

avec $m\,\boldsymbol{OG} \wedge (\boldsymbol{u} \wedge \boldsymbol{OG}) = m \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \wedge \left[\begin{pmatrix} x \\ y \\ z \end{pmatrix} \wedge \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \right] = m \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \wedge \begin{pmatrix} -bz \\ az \\ bx - ay \end{pmatrix} = m \begin{pmatrix} b^2x - aby \\ a^2y - abx \\ (a^2 + b^2)z \end{pmatrix}$

$$= m \begin{pmatrix} (a^2 + b^2)x - a(ax + by) \\ (a^2 + b^2)y - b(ax + by) \\ (a^2 + b^2)z \end{pmatrix}$$

$$= m\,(\mathrm{OG}^2\boldsymbol{u} - \langle \boldsymbol{u} \mid \boldsymbol{OG} \rangle\, \boldsymbol{OG})$$

opérateur de matrice $\begin{pmatrix} mb^2 & -mab & 0 \\ -mab & ma^2 & 0 \\ 0 & 0 & m(a^2 + b^2) \end{pmatrix}$. Il en résulte que la matrice de $\boldsymbol{J}_{\mathrm{O}}$ est :

$$\begin{pmatrix} \frac{mb^2}{3} & 0 & 0 \\ 0 & \frac{ma^2}{3} & 0 \\ 0 & 0 & \frac{m(a^2 + b^2)}{3} \end{pmatrix} + \begin{pmatrix} mb^2 & -mab & 0 \\ -mab & ma^2 & 0 \\ 0 & 0 & m(a^2 + b^2) \end{pmatrix} = \begin{pmatrix} \frac{4mb^2}{3} & -mab & 0 \\ -mab & \frac{4ma^2}{3} & 0 \\ 0 & 0 & \frac{4m(a^2 + b^2)}{3} \end{pmatrix}$$

On peut aussi faire un calcul direct.

EXEMPLE 3 :

❑ Considérons une surface cônique S de révolution de hauteur h, le rayon de la base circulaire étant R. On place le sommet en O et l'axe selon Oz. La matrice d'inertie se calcule au moyen d'intégrale double. Soit θ l'angle polaire dans le plan Oxy repérant le projeté (x, y) sur ce plan d'un point (x, y, z) de S. On vérifiera que, en prenant une densité surfacique égale à 1 :

$$\mathrm{A} = \iint_{\mathrm{S}} (y^2 + z^2)\,\mathrm{d}m = \iint_{\mathrm{S}} (x^2 + z^2)\,\mathrm{d}m = \mathrm{B}$$

$$\text{avec } x = \frac{\mathrm{R}}{h} z \cos(\theta) \text{ et } \mathrm{d}m = \frac{\mathrm{R}}{h} z\,\mathrm{d}\theta\,\frac{\sqrt{\mathrm{R}^2 + h^2}}{h}\,\mathrm{d}z,\ 0 \le \theta \le 2\pi,\ 0 \le z \le h$$

$$= m\left(\frac{\mathrm{R}^2}{4} + \frac{h^2}{2}\right)$$

$$\mathrm{C} = \iint_{\mathrm{S}} (x^2 + y^2)\,\mathrm{d}m \text{ avec } x^2 + y^2 = \frac{\mathrm{R}^2}{h^2} z^2 \text{ et } \mathrm{d}m = \frac{\mathrm{R}}{h} z\,\mathrm{d}\theta\,\frac{\sqrt{\mathrm{R}^2 + h^2}}{h}\,\mathrm{d}z$$

$$= \frac{1}{2} m\mathrm{R}^2$$

Les autres coefficients sont nuls donc la matrice est $\begin{pmatrix} m\left(\frac{\mathrm{R}^2}{4} + \frac{h^2}{2}\right) & 0 & 0 \\ 0 & m\left(\frac{\mathrm{R}^2}{4} + \frac{h^2}{2}\right) & 0 \\ 0 & 0 & \frac{1}{2} m\mathrm{R}^2 \end{pmatrix}$. Comme cas particulier limite, on obtient pour $h = 0$ la matrice d'inertie d'un disque relativement à ses axes principaux $\begin{pmatrix} \frac{m\mathrm{R}^2}{4} & 0 & 0 \\ 0 & \frac{m\mathrm{R}^2}{4} & 0 \\ 0 & 0 & \frac{m\mathrm{R}^2}{2} \end{pmatrix}$.

EXEMPLE 4 :

❑ Si le cône est plein, on vérifiera qu'un calcul mené avec des intégrales triples donne :

$$\frac{3m}{5}\begin{pmatrix} \frac{R^2}{4} + h^2 & 0 & 0 \\ 0 & \frac{R^2}{4} + h^2 & 0 \\ 0 & 0 & \frac{R^2}{2} \end{pmatrix}$$

Si le vecteur de rotation instantané $\mathbf{\Omega}$ est porté par une génératrice du cône (cas où le cône roule sans glisser sur un plan par exemple), colinéaire à $\begin{pmatrix} 0 \\ R \\ h \end{pmatrix}$, alors $\boldsymbol{L}$(O) est colinéaire à $\begin{pmatrix} 0 \\ \frac{R^2}{4} + h^2 \\ \frac{hR}{2} \end{pmatrix}$. En général, $\boldsymbol{L}$(O) n'est pas colinéaire à $\mathbf{\Omega}$ sauf si R = 2h.

EXEMPLE 5 :

❑ Pour une boule pleine de masse m de rayon R, la matrice d'inertie vaut $\frac{2}{5}mR^2\, I_3$.

2- Polarisation

Considérons un corps diélectrique (non conducteur), plongé dans un champ électrique $\boldsymbol{E}$. Les électrons et les noyaux de chaque atome subissent des forces électrostatiques opposées qui tendent à créer, dans chaque élément de volume, un dipôle électrique. Il en résulte une polarisation globale par unité de volume $\boldsymbol{P}$. Dans un milieu homogène et isotrope et pour des valeurs de $\boldsymbol{E}$ pas trop grande, $\boldsymbol{P}$ est simplement proportionnelle à $\boldsymbol{E}$: $\boldsymbol{P} = \varepsilon_0\chi\boldsymbol{E}$ où ε_0 est la permittivité du vide et χ un facteur appelé **susceptibilité électrique**. Si le milieu n'est pas isotrope, $\boldsymbol{P}$ est une fonction linéaire de $\boldsymbol{E}$. Il existe une transformation linéaire M telle que :

$$\boldsymbol{P} = \mathrm{M}\boldsymbol{E}$$

qu'on peut encore écrire sous forme de composantes, dans unee base orthonormée donnée :

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \begin{pmatrix} M_{xx} & M_{xy} & M_{xz} \\ M_{yx} & M_{yy} & M_{yz} \\ M_{zx} & M_{zy} & M_{zz} \end{pmatrix}\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}$$

M est symétrique. En effet, l'énergie nécessaire à une polarisation est, par unité de volume, $\int \langle\boldsymbol{E} \mid d\boldsymbol{P}\rangle$ où < , > désigne le produit scalaire. Commençons par appliquer un champ électrique selon x, à savoir $\boldsymbol{E} = E\boldsymbol{i}$, $0 \le E \le E_x$, ce qui nécessite une énergie $\int_0^{E_x} E\, M_{xx}\, dE = \frac{1}{2} M_{xx}\, E_x^2$. Puis appliquons un champ électrique selon y, soit $\boldsymbol{E} = E_x\boldsymbol{i} + E\boldsymbol{j}$, $0 \le E \le E_y$. On a alors $\boldsymbol{P}$ de composantes $\begin{pmatrix} M_{xx}E_x + M_{xy}E \\ M_{yx}E_x + M_{yy}E \\ 0 \end{pmatrix}$, donc d$\boldsymbol{P}$ a pour composantes $dE\begin{pmatrix} M_{xy} \\ M_{yy} \\ 0 \end{pmatrix}$. On a donc $\langle\boldsymbol{E},d\boldsymbol{P}\rangle = (E_xM_{xy} + EM_{yy})\, dE$, donc $\int \langle\boldsymbol{E},d\boldsymbol{P}\rangle = M_{xy}E_xE_y + \frac{1}{2}M_{yy}E_y^2$. Le travail total est donc $\frac{1}{2}M_{xx}\, E_x^2 + M_{xy}E_xE_y + \frac{1}{2}M_{yy}E_y^2$. Si on avait commencé par appliquer un champ suivant y puis suivant x, les indices x et y auraient été inversés et le travail aurait été $\frac{1}{2}$ M_{xx} E_x^2 + $M_{yx}E_xE_y$ + $\frac{1}{2}$ $M_{yy}E_y^2$. Les états initiaux et finaux étant identiques (ce qu'on vérifie expérimentalement), le travail est le même selon les deux méthodes (principe de conservation de l'énergie). On en déduit que $M_{xy} = M_{yx}$. On raisonne de même pour les autres indices. On pourra également vérifier que le travail total pour appliquer un champ quelconque est W = $\frac{1}{2}$ <***E***, M***E***>, ou bien, si on note E les composantes ***E***, W = $\frac{1}{2}$ E^TME, norme du produit scalaire défini par la matrice symétrique $\frac{1}{2}$ M.

Le fait qu'une matrice symétrique soit diagonalisable dans une base orthonormée signifie qu'il existe trois directions orthogonales de polarisation privilégiée, pour lesquelles ***P*** est colinéaire à ***E***, mais les coefficients de colinéarité ne sont pas forcément identiques. Ce sont les valeurs propres de la matrice M. Les valeurs propres sont identiques si le cristal possède les mêmes propriétés dans les trois directions, par exemple, s'il possède une structure cubique. Dans ce cas, la matrice M est non seulement diagonale, mais scalaire, et l'on retrouve alors la formule $\boldsymbol{P} = \varepsilon_0\chi\boldsymbol{E}$.

Annexe II : produit vectoriel en dimension n

1- Définition et propriétés

Soit E un espace vectoriel euclidien orienté de dimension *n*. On se propose de généraliser à E la notion de produit vectoriel. Il ne s'agira pas de définir ce qu'est le produit vectoriel de deux vecteurs, mais de prendre $n - 1$ vecteurs $y_2, ..., y_n$ est de définir ce qu'est, en bloc, le produit vectoriel de ces $n - 1$ vecteurs. Pour cela, on procède comme suit. Pour tout vecteur $y_2, ..., y_n$, considérons la forme linéaire :

$$\varphi : x \to \det(x, y_2, ..., y_n) \in \mathbb{R}$$

où le déterminant est calculé dans une **base orthonormée directe** quelconque de E (produit mixte en dimension *n*). Le fait que le produit mixte ne dépende pas de la base orthonormée directe choisie se montre en dimension *n* comme on l'a fait en dimension 3 dans le chapitre L1/ESPEUCL.PDF.

Comme pour toute forme linéaire dans un espace vectoriel euclidien, $\varphi(x)$ peut s'exprimer sous la forme du produit scalaire de *x* par un vecteur de E, dépendant ici de $y_2, ..., y_n$ et que l'on note $y_2 \times y_3 \times ... \times y_n$, appelé **produit vectoriel** de $(y_2, ..., y_n)$:

$$\forall x \in E, \det(x, y_2, ..., y_n) = \langle x, y_2 \times y_3 \times ... \times y_n \rangle$$

Cette définition généralise le produit vectoriel en dimension *n*. En effet, en dimension 3, on reconnaît le produit vectoriel usuel puisque le produit mixte vérifie :

$$\forall x, \det(x, y_2, y_3) = \langle x, y_2 \wedge y_3 \rangle$$

PROPRIETES

(i) *L'application* $(y_2, ..., y_n) \in E^{n-1} \to y_2 \times y_3 \times ... \times y_n$ *est multilinéaire alternée.*

(ii) *Si* $(e_1, ..., e_n)$ *est une base orthonormée directe de* E, *les composantes de* $y_2 \times ... \times y_n$ *dans cette base sont les* $\det(e_k, y_2, ..., y_n)$, $1 \le k \le n$, *où le déterminant est calculé dans une base orthonormée directe de* E.

(iii) *Les* y_i *sont liés si et seulement si* $y_2 \times ... \times y_n = 0$.

(iv) $y_2 \times ... \times y_n$ *est orthogonal au sous-espace vectoriel engendré par les* y_i.

(v) *Si* $y_2, ..., y_n$ *sont indépendants,* $(y_2 \times ... \times y_n, y_2, ..., y_n)$ *est une base directe de* E

(vi) $\| y_2 \times ... \times y_n \| = \left|\det(y_2, ..., y_n)\right|$, *où ce dernier déterminant est calculé dans une base orthonormée d'un sous-espace vectoriel de dimension* $n - 1$ *contenant* $y_2, ..., y_n$ (*défini au signe près selon l'orientation de cette base*).

(vii) *Pour tout* $x_2, ..., x_n, y_2, ..., y_n$, $<x_2 \times ... \times x_n, y_2 \times ... \times y_n> = \det(<x_i, y_j>_{2\le i,j\le n})$.

Démonstration :

❑ (i) : Le caractère multilinéaire alterné de $(y_2, ..., y_n)$ se déduit du fait que le déterminant possède le même caractère. Ainsi, pour tout $x, y_2, ..., z_n, z$ de E, tout indice i entre 2 et n et tout réel α et β, on a :

$$<x, y_2 \times .... \times y_{i-1} \times (\alpha y_i + \beta z) \times y_{i+1} \times ... \times y_n> = \det(x, y_2, ..., y_{i-1}, \alpha y_i + \beta z, y_{i+1}, ..., y_n)$$
$$= \alpha \det(x, y_2, ..., y_{i-1}, y_i, y_{i+1}, ..., y_n) + \beta \det(x, y_2, ..., y_{i-1}, z, y_{i+1}, ..., y_n)$$

en utilisant la linéarité du déterminant sur le i-ème vecteur

$$= \alpha <x, y_2 \times ... \times y_{i-1} \times y_i \times y_{i+1} \times ... \times y_n> + \beta <x, y_2 \times ... \times y_{i-1} \times z \times y_{i+1} \times ... \times y_n>$$
$$= <x, \alpha(y_2 \times ... \times y_{i-1} \times y_i \times y_{i+1} \times ... \times y_n) + \beta(y_2 \times ... \times y_{i-1} \times z \times y_{i+1} \times ... \times y_n)>$$

Cette égalité étant vraie pour tout x, on a :

$$y_2 \times .... \times y_{i-1} \times (\alpha y_i + \beta z) \times y_{i+1} \times ... \times y_n = \alpha(y_2 \times ... \times y_n) + \beta(y_2 \times ... \times y_{i-1} \times z \times y_{i+1} \times ... \times y_n)$$

ce qui prouve la linéarité du produit vectoriel par rapport au $(i - 1)$-ème vecteur.

Le caractère alterné se prouve de même. Si on permute deux vecteurs y_i et y_j, le signe de $\det(x, y_2, ..., y_n)$ est changé, donc celui de $<x, y_2 \times y_3 \times ... \times y_n>$ aussi, et ceci étant vrai pour tout vecteur x, c'est $y_2 \times y_3 \times ... \times y_n$ qui voit son signe changé.

❑ (ii) : la composante de $y_2 \times ... \times y_n$ selon e_k est :

$$<e_k, y_2 \times ... \times y_n> = \det(e_k, y_2, ..., y_n)$$

❑ (iii) : si les y_i sont liés, alors, pour tout x, il en est a fortiori de même de x et des y_i, donc :

$$\det(x, y_2, ..., y_n) = 0$$

donc :

$$\forall\, x, <x, y_2 \times ... \times y_n> = 0$$

donc $\quad y_2 \times ... \times y_n = 0$.

Si les y_i sont indépendants, on peut choisir x tel que $(x, y_2, ..., y_n)$ est une base de E, donc :

$$\det(x, y_2, ..., y_n) \neq 0$$

donc $\quad <x, y_2 \times ... y_n> \neq 0$

donc $\quad y_2 \times ... \times y_n \neq 0$.

❑ (iv) : pour tout i variant de 2 à n, on a :

$$<y_i, y_2 \times ... \times y_n> = \det(y_i, y_2, ..., y_n)$$
$$= 0$$

car y_i apparaît deux fois dans le déterminant. Ainsi, $y_2 \times ... \times y_n$ est orthogonal à tout y_i donc également au sous-espace vectoriel engendré.

❑ (v) : $\det(y_2 \times ... \times y_n, y_2, ..., y_n) = <y_2 \times ... \times y_n, y_2 \times ... \times y_n>$
$$= \| y_2 \times ... \times y_n \|^2 > 0$$

❑ (vi) : Si $y_2, ..., y_n$ sont liés, on a $y_2 \times ... \times y_n = 0 = \det(y_2, ..., y_n)$.

Sinon, calculons $\det(y_2 \times ... \times y_n, y_2, ..., y_n)$ dans une base orthonormée directe $(e_1, ..., e_n)$ dont le premier vecteur e_1 est colinéaire et de même sens que $y_2 \times ... \times y_n$, de sorte que :

$$y_2 \times ... \times y_n = \| y_2 \times ... \times y_n \|\, e_1$$

Puisque $y_2 \times ... \times y_n$ et donc e_1 est orthogonal à $y_2, ..., y_n$, les composantes de $y_2, ..., y_n$ selon e_1 sont nulles, donc la première ligne de $\det(y_2 \times ... \times y_n, y_2, ..., y_n)$ est constituée du terme $\| y_2 \times ... \times y_n \|$ suivi de termes nuls. Développons le déterminant par rapport à la première ligne. Le mineur du terme (1,1) du déterminant $n \times n$ est formé des composantes de $y_2, ..., y_n$ dans la base orthonormée $(e_2, ..., e_n)$. C'est, au signe près, ce que nous avons appelé $\det(y_2, ..., y_n)$. On obtient :

$$\det(y_2 \times ... \times y_n, y_2, ..., y_n) = \pm \| y_2 \times ... \times y_n \| \det(y_2, ..., y_n)$$

Comme on a vu ci-dessus que $\det(y_2 \times ... \times y_n, y_2, ..., y_n) = \| y_2 \times ... \times y_n \|^2$, on obtient bien :

$$\left|\det(y_2, ..., y_n)\right| = \| y_2 \times ... \times y_n \|$$

avec la possibilité de supprimer la valeur absolue si le calcul de $\det(y_2, ..., y_n)$ se fait dans la base $(e_2, ..., e_n)$ dont l'orientation est liée à celle de e_1, i.e. à la direction du vecteur $y_2 \times ... \times y_n$.

❑ (vii) : Pour tout $x_1, ..., x_n, y_1, ..., y_n$ vecteurs de E, notons $X_1, ..., X_n, Y_1, ..., Y_n$ les colonnes constituées des composantes de ces vecteurs dans une base orthonormée directe $(e_1, ..., e_n)$ de E. Notons X la matrice constituée des colonnes $X_1, ..., X_n$ et Y la matrice constituée des colonnes $Y_1, ..., Y_n$. $<x_i, y_j>$ n'est autre que le terme général de la matrice X^TY, où X^T est la transposée de X, donc :

$$\begin{aligned}\det(<x_i, y_j>_{1\leq i,j\leq n}) &= \det(X^TY)\\ &= \det(X^T)\det(Y)\\ &= \det(X)\det(Y)\\ &= \det(x_1, ..., x_n)\det(y_1, ..., y_n)\end{aligned}$$

Prenons x_1 égal à $y_2 \times ... \times y_n$ de sorte que, d'après (iv), $<x_1, y_i> = 0$ pour $i \geq 2$. La première ligne de la matrice $(<x_i, y_j>_{1\leq i,j\leq n})$ contient le terme $<x_1, y_1>$ suivi de termes tous nuls. Si on développe ce déterminant par rapport à la première ligne, on obtient d'une part, pour le membre de gauche :

$$\begin{aligned}\det(<x_i, y_j>_{1\leq i,j\leq n}) &= <x_1, y_1> \det(<x_i, y_j>_{2\leq i,j\leq n})\\ &= <y_2 \times ... \times y_n, y_1> \det(<x_i, y_j>_{2\leq i,j\leq n})\end{aligned}$$

D'autre part, pour le membre de droite :

$$\begin{aligned}\det(x_1, ..., x_n)\det(y_1, ..., y_n) &= \det(y_2 \times ... \times y_n, x_2, ..., x_n)\det(y_1, ..., y_n)\\ &\qquad \text{en remplaçant } x_1 \text{ par sa valeur}\\ &= <y_2 \times ... \times y_n, x_2 \times ... \times x_n><y_1, y_2 \times ... \times y_n>\\ &\qquad \text{par définition de } x_2 \times ... \times x_n \text{ et de } y_2 \times ... \times y_n\end{aligned}$$

Les deux membres étant égaux :

$$<y_2 \times ... \times y_n, y_1> \det(<x_i, y_j>_{2\leq i,j\leq n}) = <y_2 \times ... \times y_n, x_2 \times ... \times x_n><y_1, y_2 \times ... \times y_n>$$

Si $y_2 \times ... \times y_n \neq 0$, prenons $y_1 = y_2 \times ... \times y_n$ et simplifions l'égalité par $<y_1, y_2 \times ... \times y_n>$ qui est non nul, conduisant à l'égalité demandée $\det(<x_i, y_j>_{2\leq i,j\leq n}) = <y_2 \times ... \times y_n, x_2 \times ... \times x_n>$.
Si $y_2 \times ... \times y_n = 0$, alors les y_j sont liés, donc l'un des y_j est combinaison linéaire des autres y_k. Mais dans ce cas, la colonne $(<x_i, y_j>_{2\leq i\leq n})$ est aussi combinaison linéaire des autres colonnes $(<x_i, y_k>_{2\leq i\leq n})$ et $\det(<x_i, y_j>_{2\leq i,j\leq n}) = 0 = <y_2 \times ... \times y_n, x_2 \times ... \times x_n>$.

2- Mesure d'un parallélotope en dimension *n*

Nous avons vu, dans le chapitre sur les déterminants L1/DETERMNT.PDF, que, en dimension 2, $\det(y_1, y_2)$, calculé dans une base (e_1, e_2) donnée, est l'aire algébrique du parallélogramme de côtés y_1 et y_2, l'unité de mesure étant donnée par l'aire du parallélogramme construit sur les côtés e_1 et e_2. De même, en dimension 3, $\det(y_1, y_2, y_3)$, calculé dans une base (e_1, e_2, e_3) donnée, est le volume algébrique du parallélépipède de côtés y_1, y_2 et y_3, l'unité de mesure étant donnée par le volume du parallélépipède construit sur les côtés e_1, e_2 et e_3. Les aires et volumes géométriques sont pris en prenant les valeurs absolues des aires et volumes algébriques. Dans un espace euclidien, il est naturel de travailler dans une base orthonormée, l'aire ou le volume unité étant celle d'un carré ou d'un cube de côté 1. L'aire ou le volume algébrique dépendent de l'orientation de la base, mais non l'aire ou le volume géométrique qui en sont la valeur absolue.

Nous souhaitons étendre ces notions en dimension n, sans recours au calcul intégral. Soient $y_1, ..., y_n$ n vecteurs d'un espace vectoriel euclidien orienté E. Ces vecteurs engendrent un **parallélotope** de dimension n, défini comme étant l'ensemble des $\sum_{i=1}^{n} \lambda_i y_i$, $(\lambda_1, ..., \lambda_n) \in [0, 1]^n$. On appelle **mesure algébrique** de ce parallélotope le produit mixte de $y_1, ..., y_n$, i.e. $\det(y_1, ..., y_n)$ calculé dans une base orthonormée directe. Sa **mesure géométrique** est égale à $|\det(y_1, ..., y_n)|$. Elle est nulle si et seulement si les vecteurs sont liés, i.e. contenus dans un même hyperplan.
En ce sens, les endomorphismes u de E de déterminant ± 1 jouent un rôle particulier. Ce sont ceux qui conservent les mesures géométriques des parallélotopes au cours de leur transformation, puisque, pour un tel endomorphisme :

$$|\det(u(y_1), ..., u(y_2))| = |\det(u) \det(y_1, ..., y_n)| = |\det(y_1, ..., y_n)|$$

Les endomorphismes de déterminant 1 conservent les mesures algébriques.
Plus généralement, pour un endomorphisme quelconque, $\det(u)$ donne le facteur de dilatation ou de contraction de la mesure d'une figure lors de sa transformation par u.

Soit H un hyperplan contenant $y_2, ..., y_n$ ($\mathrm{H} = \mathrm{Vect}(y_2, ..., y_n)$ si ces vecteurs sont linéairement indépendants). On peut définir le parallélotope de dimension $n - 1$ engendré par $y_2, ..., y_n$ dans H. Ce parallélotope constitue une **hyperface** du parallélotope initial. La formule (vi) du paragraphe précédent exprime que, dans le sous-espace vectoriel H de dimension $n - 1$, cette hyperface aura pour mesure géométrique $\| y_2 \times ... \times y_n \| = |\det(y_2, ..., y_n)|$. Le signe de sa mesure algébrique dépendra de l'orientation choisie pour H. Projetons y_1 sur la droite orthogonale à H, et notons z_1 son projeté. $\| z_1 \|$ est une **hauteur** du parallélotope initial, relativement à l'hyperface $(y_2, ..., y_n)$. y_1 est égal à la somme de z_1 et d'une combinaison linéaire de $y_2, ..., y_n$. Donc :

$$|\det(y_1, y_2, ..., y_n)| = |\det(z_1, y_2, ..., y_n)|$$

Si on calcule ce dernier déterminant dans une base orthonormée dont le premier vecteur est colinéaire et de même sens que z_1, la première ligne est constituée des termes $\| z_1 \|, 0, ..., 0$ car $y_2, ..., y_n$ sont orthogonaux à z_1. On peut donc développer le déterminant par rapport à la première ligne, le mineur du terme (1, 1) du déterminant initial n'étant autre que $\det(y_2, ..., y_n)$, calculé dans une base orthonormée de H. Autrement dit, ce mineur a pour valeur absolue la mesure géométrique du parallélotope $(y_2, ..., y_n)$ dans H. On a donc :

$$|\det(y_1, ..., y_n)| = \| z_1 \| \times \| y_2 \times ... \times y_n \|$$

Ainsi, la mesure du parallélotope est le produit de la norme de la hauteur par la mesure de l'hyperface associée, comme on l'a en dimension 2 et 3 pour le parallélogramme et le parallélépipède.

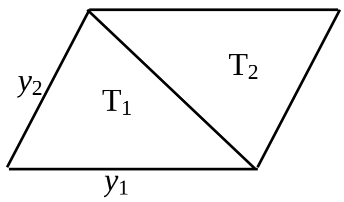
3- Mesure d'un simplexe en dimension *n*

On peut également généraliser la notion de triangle du plan et de tétraèdre de l'espace tridimensionel. En dimension n, le **simplexe** engendré par $y_1, ..., y_n$ est l'ensemble des $\sum_{i=1}^{n} \lambda_i y_i$, avec les $\lambda_i \geq 0$ vérifiant $\sum_{i=1}^{n} \lambda_i \leq 1$. Il sera utile, dans la suite, de pouvoir aussi considérer E comme un espace affine, un élément de E pouvant être vu, selon l'usage qui en est fait, comme un vecteur ou comme un point (voir le chapitre L1/GEOMAFF.PDF). Posons $y_0 = 0$. Le simplexe précédent est alors l'enveloppe convexe des $n + 1$ points $y_0, y_1, ..., y_n$, puisque :

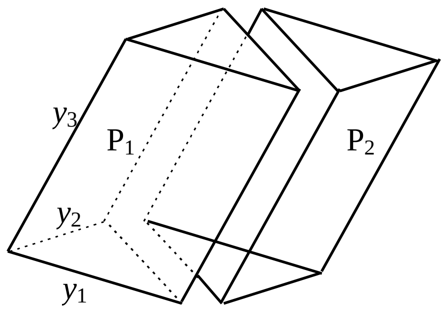
$$\sum_{i=1}^{n} \lambda_i y_i = (1 - \sum_{i=1}^{n} \lambda_i) y_0 + \sum_{i=1}^{n} \lambda_i y_i$$

tous les coefficients étant positifs ou nuls et de somme égale à 1.
Plus généralement, le simplexe engendré par $n + 1$ points $s_0, ..., s_n$ de E est l'enveloppe convexe de ces $n + 1$ points. Si on pose $y_i = s_i - s_0$, $1 \leq i \leq n$, c'est simplement le translaté du simplexe précédent par le vecteur s_0.

On rappelle que, dans le plan, l'aire géométrique d'un triangle de côtés y_1 et y_2 est $\frac{1}{2} |\det(y_1, y_2)|$ puisque le parallélogramme de côtés y_1 et y_2 peut être découpé en les deux triangles de même aire :

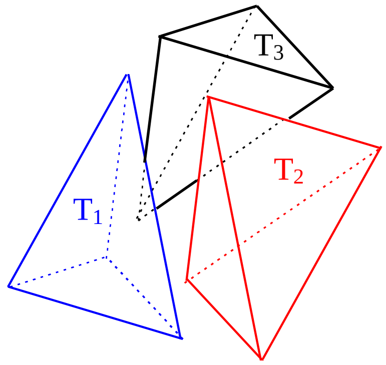


En dimension 3, le volume du tétraèdre construit sur les côtés y_1, y_2 et y_3 est le $\frac{1}{6}$ du volume du parallélépipède correspondant. Pour le voir, il suffit de découper le parallélépipède en 6 tétraèdres de même volume. On peut déjà diviser le parallélépipède en deux prismes P_1 et P_2 de même volume :



découpage d'un parallélépipède en deux prismes de même volume, légérement séparés pour plus de visibilité

On découpe ensuite chaque prisme en trois tétraèdres de même volume.



découpage du prisme P_1 en trois tétraèdres de même volume

Le tétraèdre T_1 est celui dont trois côtés sont y_1, y_2, y_3. Son volume est donc bien le tiers de celui de P_1 et le sixième du parallélotope complet.

S'inspirant des démarches précédentes, on se propose de montrer que, en dimension n, la mesure géométrique d'un simplexe engendré par $(y_1, ..., y_n)$ vaut $\frac{1}{n!}\,|\det(y_1, ..., y_n)|$. Il suffit pour cela de montrer que le parallélotope V_n engendré par $(y_1, ..., y_n)$ peut être découpé en $n!$ simplexes de même mesure, deux simplexes différents ayant éventuellement une intersection non vide, mais de mesure nulle. Pour cela, on procède par récurrence. Nous supposerons dans la suite que y_1, ..., y_n sont linéairement indépendants, sinon, toutes les mesures sont nulles. On commence par se placer dans l'hyperplan $H = \mathrm{Vect}(y_1, ..., y_{n-1})$. D'après l'hypothèse de récurrence, le parallélotope V_{n-1} de dimension $n-1$ engendré par $(y_1, ..., y_{n-1})$, peut être découpé en $(n-1)!$ simplexes de dimension $n-1$ et de même mesure. Le parallélotope V_n engendré par $(y_1, ..., y_n)$ est la réunion des translatés $V_{n-1} + \lambda_n y_n$, $0 \le \lambda_n \le 1$. La réunion des translatés de chaque simplexe de dimension $n-1$ est un prisme. On obtient ainsi un découpage de V_n en $(n-1)!$ prismes de même mesure, et pour conclure, il suffit de découper chaque prisme en n simplexes de dimension n et de même mesure. Faisons-le pour le prisme P_1 défini par :

$$P_1 = \{\sum_{i=1}^{n} \lambda_i y_i,\ \lambda_i \ge 0,\ \lambda_1 + ... + \lambda_{n-1} \le 1,\ 0 \le \lambda_n \le 1\}$$

Il suffit de prendre les n simplexes S_i dont les $n+1$ sommets sont respectivement :

$S_0 : (0, y_1, y_2, y_3, ..., y_n)$,
$S_1 : (y_1, y_2, y_3, ..., y_n, y_1 + y_n)$,
$S_2 : (y_2, y_3, ..., y_n, y_1 + y_n, y_2 + y_n)$,
$S_3 : (y_3, ..., y_n, y_1 + y_4, y_2 + y_n, y_3 + y_n)$
...
$S_k : (y_k, ..., y_n, y_1 + y_n, y_2 + y_n, ..., y_k + y_n)$
...
$S_{n-1} : (y_{n-1}, y_n, y_1 + y_4, y_2 + y_n, ..., y_{n-1} + y_n)$

❑ Vérifions que les S_k, ont même mesure. Le simplexe S_k a pour sommet y_k ($y_0 = 0$ pour $k = 0$) et pour côtés $y_{k+1} - y_k$, ..., $y_n - y_k$, $y_1 + y_n - y_k$, ..., $y_{k-1} + y_n - y_k$, y_n. Considérons l'application linéaire φ qui transforme les côtés $(y_1, ..., y_n)$ de S_0 en les côtés de S_k, et vérifions que $\det(\varphi) = \pm 1$. On a :

$$|\det(\varphi(y_1), ... \varphi(y_n))| = |\det(y_{k+1} - y_k, ..., y_{n-1} - y_k, y_n - y_k, y_1 + y_n - y_k, ..., y_{k-1} + y_n - y_k, y_n)|$$
$$= |\det(y_{k+1} - y_k, ..., y_{n-1} - y_k, - y_k, y_1 - y_k, ..., y_{k-1} - y_k, y_n)|$$

en retranchant la dernière colonne y_n aux autres qui contiennent ce terme

$= |\det(y_{k+1}, ..., y_{n-1}, -y_k, y_1, ..., y_{k-1}, y_n)|$

en retranchant la colonne $-y_k$ aux autres qui contiennent ce terme

$= |\det(y_1, ..., y_n)|$

en remettant les indices dans l'ordre

Comme par ailleurs, $|\det(\varphi(y_1), ... \varphi(y_n))| = |\det(\varphi)|\ |\det(y_1, ..., y_n)|$, on en conclut que $|\det(\varphi)| = 1$. Donc S_k et S_0 ont même mesure.

❑ Soient deux simplexes différents S_i et S_k, $0 \le i < k < n$:

$S_i : (y_i, ..., y_k, ..., y_n, y_1 + y_n, y_2 + y_n, ..., y_i + y_n)$

$S_k : (y_k, ..., y_n, y_1 + y_n, y_2 + y_n, ..., y_i + y_n, ..., y_k + y_n)$

Montrons que $S_i \cap S_k$ est l'enveloppe convexe C des points $y_k, ..., y_n, y_1 + y_n, ..., y_i + y_n$. C sera alors inclus dans le sous-espace affine engendré par les mêmes points, qui est de dimension $n - k + i < n$. Donc C sera de mesure nulle.

On a trivialement $C \subset S_i \cap S_k$. Montrons la réciproque. Un point M de l'intersection est de la forme :

$$M = \lambda_i y_i + ... + \lambda_k y_k + ... + \lambda_n y_n + \mu_1(y_1 + y_n) + ... + \mu_i(y_i + y_n) \in S_i$$
$$= \alpha_k y_k + ... + \alpha_n y_n + \beta_1(y_1 + y_n) + ... + \beta_i(y_i + y_n) + ... + \beta_k(y_k + y_n) \in S_k$$

tous les coefficients étant positifs ou nuls et $\sum_{p=i}^{n} \lambda_p + \sum_{p=1}^{i} \mu_p = 1$, $\sum_{p=k}^{n} \alpha_p + \sum_{p=1}^{k} \beta_p = 1$.

Les coefficients selon $y_1, ..., y_{i-1}$ devant être égaux dans les deux expressions de M, on a nécessairement $\mu_1 = \beta_1, ..., \mu_{i-1} = \beta_{i-1}$. De même, en comparant les coefficients de $y_{i+1}, ..., y_{k-1}$, on a $\lambda_{i+1} = \beta_{i+1}, ..., \lambda_{k-1} = \beta_{k-1}$. Enfin, en comparant les coefficients de $y_{k+1}, ..., y_{n-1}$, on a $\lambda_{k+1} = \alpha_{k+1}, ..., \lambda_{n-1} = \alpha_{n-1}$. L'élimination des termes communs dans l'égalité $\sum_{p=i}^{n} \lambda_p + \sum_{p=1}^{i} \mu_p = \sum_{p=k}^{n} \alpha_p + \sum_{p=1}^{k} \beta_p$ et dans les deux expressions de m donnent alors :

$$\lambda_i + \lambda_k + \lambda_n + \mu_i = \alpha_k + \alpha_n + \beta_i + \beta_k$$

et $\lambda_i y_i + \lambda_k y_k + \lambda_n y_n + \mu_i(y_i + y_n) = \alpha_k y_k + \alpha_n y_n + \beta_i(y_i + y_n) + \beta_{i+1} y_n + ... + \beta_{k-1} y_n + \beta_k(y_k + y_n)$

d'où l'on tire, en comparant les coefficients de y_i, y_k et y_n :

$$\left[\begin{array}{l} \lambda_i + \lambda_k + \lambda_n + \mu_i = \alpha_k + \alpha_n + \beta_i + \beta_k \\ \lambda_i + \mu_i = \beta_i \\ \lambda_k = \alpha_k + \beta_k \\ \lambda_n + \mu_i = \alpha_n + \beta_i + \beta_{i+1} + ... + \beta_{k-1} + \beta_k \end{array}\right.$$

$$\Leftrightarrow \left[\begin{array}{l} \lambda_n = \alpha_n \\ \lambda_i + \mu_i = \beta_i \\ \lambda_k = \alpha_k + \beta_k \\ \lambda_n = \alpha_n + \lambda_i + \lambda_{i+1} + ... + \lambda_{k-1} + \lambda_k - \alpha_k \end{array}\right.$$

en remplaçant dans la première et la dernière équation β_i par $\lambda_i + \mu_i$, β_k par $\lambda_k - \alpha_k$ et $\beta_{i+1}, ..., \beta_{k-1}$ par $\lambda_{i+1}, ..., \lambda_{k-1}$.

$$\Leftrightarrow \left[\begin{array}{l} \lambda_n = \alpha_n \\ \lambda_i + \mu_i = \beta_i \\ -\lambda_i - ... - \lambda_{k-1} = \beta_k \\ \alpha_k = \lambda_i + \lambda_{i+1} + ... + \lambda_{k-1} + \lambda_k \end{array}\right.$$

Comme les λ et les β sont positifs ou nuls, la troisième équation impose que :

$$\beta_k = 0 = \lambda_i = ... = \lambda_{k-1}$$

donc $\beta_i = \mu_i$, $\beta_{i+1} = ... = \beta_{k-1} = 0$ et $\alpha_k = \lambda_k$

Il reste donc dans l'expression de M :

$$M = \lambda_k y_k + ... + \lambda_n y_n + \mu_1(y_1 + y_n) + ... + \mu_i(y_i + y_n)$$

avec tous les coefficients positifs ou nuls et $\sum_{p=k}^{n} \lambda_p + \sum_{p=1}^{i} \mu_p = 1$. Donc $M \in C$.

On a ainsi montré que $S_i \cap S_k = C$.

❑ Tous les simplexes S_i sont inclus dans P_1. En effet, un point M de S_i s'écrit :

$$\begin{aligned} M &= \lambda_i y_i + ... + \lambda_n y_n + \mu_1(y_1 + y_n) + ... + \mu_i(y_i + y_n) \\ &= \mu_1 y_1 + ... + \mu_{i-1} y_{i-1} + (\lambda_i + \mu_i) y_i + \lambda_{i+1} y_{i+1} + ... + \lambda_{n-1} y_{n-1} + (\lambda_n + \mu_1 + ... + \mu_i) y_n \end{aligned}$$

tous les coefficients étant positifs ou nuls et $\sum_{p=i}^{n} \lambda_p + \sum_{p=1}^{i} \mu_p = 1$.

M est le translaté du point $N = \mu_1 y_1 + ... + \mu_{i-1} y_1 + (\lambda_i + \mu_i) y_i + \lambda_{i+1} y_{i+1} + ... + \lambda_{n-1} y_{n-1}$ par le vecteur $(\lambda_n + \mu_1 + ... + \mu_i) y_n$. Comme les coefficients de N sont positifs ou nuls et de somme inférieure ou égale à 1, N est élément du simplexe engendré par $(0, y_1, ..., y_{n-1})$ inclus dans le parallélotope V_{n-1}. Comme le vecteur de translation de N vérifie $0 \leq \lambda_n + \mu_1 + ... + \mu_i \leq 1$, M est bien dans P_1.

❑ Tout élément de P_1 est dans l'un des S_i. Un élément de P_1 s'écrit :

$$M = \sum_{i=1}^{n} \lambda_i y_i,\ \lambda_i \geq 0,\ \lambda_1 + ... + \lambda_{n-1} \leq 1,\ 0 \leq \lambda_n \leq 1$$

Par ordre croissant, on a :

$$0 \leq 1 - \sum_{i=1}^{n-1} \lambda_i \leq 1 - \sum_{i=2}^{n-1} \lambda_i \leq ... \leq 1 - \sum_{i=k}^{n-1} \lambda_i \leq 1 - \sum_{i=k+1}^{n-1} \lambda_i \leq ... \leq 1 - \lambda_{n-1} \leq 1$$

λ_n, élément de $[0, 1]$, est dans l'un des intervalles limités par deux de ces valeurs successives. Soit $k \in [\![0, n-1]\!]$ tel que :

$$1 - \sum_{i=k}^{n-1} \lambda_i \leq \lambda_n \leq 1 - \sum_{i=k+1}^{n-1} \lambda_i$$

(en convenant que $1 - \sum_{i=k}^{n-1} \lambda_i = 0$ si $k = 0$, autrement dit $0 \leq \lambda_n \leq 1 - \sum_{i=1}^{n-1} \lambda_i$ dans ce cas).

Montrons alors que $M \in S_k$. Il suffit pour cela de constater que :

$$\begin{aligned} M &= \lambda_1 y_1 + ... + \lambda_n y_n \\ &= \lambda_1(y_1 + y_n) + ... + \lambda_{k-1}(y_{k-1} + y_n) + \left(\sum_{i=k}^{n} \lambda_i - 1\right)(y_k + y_n) \\ &\quad + \left(1 - \sum_{i=k+1}^{n} \lambda_i\right) y_k + \lambda_{k+1} y_{k+1} + ... + \lambda_{n-1} y_{n-1} + \left(1 - \sum_{i=1}^{n-1} \lambda_i\right) y_n \end{aligned}$$

On vérifiera que la somme des coefficients dans le terme ci-dessus est égal à 1, et qu'ils sont tous positifs ou nuls. Par exemple :

$$\sum_{i=k}^{n} \lambda_i - 1 \geq 0 \qquad \text{car } 1 - \sum_{i=k}^{n-1} \lambda_i \leq \lambda_n$$

$$1 - \sum_{i=k+1}^{n} \lambda_i \geq 0 \qquad \text{car } \lambda_n \leq 1 - \sum_{i=k+1}^{n-1} \lambda_i$$

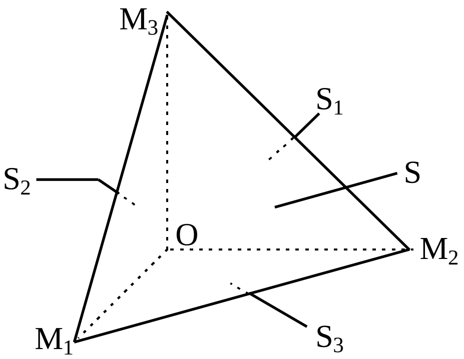
$$1 - \sum_{i=1}^{n-1} \lambda_i \geq 0 \qquad \text{car } \sum_{i=1}^{n-1} \lambda_i \leq 1 \text{ par hypothèse}$$

Il en résulte que M est dans l'enveloppe convexe de $(y_k, ..., y_n, y_1 + y_n, y_2 + y_n, ..., y_k + y_n)$, donc dans S_k.

4- Un exemple d'utilisation, une généralisation du théorème de Pythagore.

Dans le plan euclidien, on dispose du théorème de Pythagore s'appliquant à un triangle ayant deux côtés perpendiculaires. On se propose d'étendre ce théorème aux tétraèdres de l'espace de dimension 3 ayant trois faces perpendiculaires, avant de passer à la dimension n.

Soient quatre points d'un espace affine euclidien orienté de dimension 3, O, M_1, M_2 et M_3 tels que les trois vecteurs $a_i = \boldsymbol{OM}_i$ soient orthogonaux. Pour $1 \leq i \leq 3$, soit S_i l'aire du triangle formé par O et les deux points M_j pour j différent de i. Soit S l'aire du triangle formé par les trois points M_i.



Alors on a : $S^2 = S_1^2 + S_2^2 + S_3^2$.

En effet, pour $1 \leq i \leq 3$, on peut définir un vecteur $\boldsymbol{S}_i$ orthogonal à la i-ème surface triangulaire et de norme S_i, dirigé vers l'extérieur du tétraèdre $OM_1M_2M_3$, à savoir :

$$\boldsymbol{S}_1 = \frac{1}{2} a_3 \wedge a_2$$

$$\boldsymbol{S}_2 = \frac{1}{2} a_1 \wedge a_3$$

$$\boldsymbol{S}_3 = \frac{1}{2} a_2 \wedge a_1$$

et de même :

$$\boldsymbol{S} = \frac{1}{2} \boldsymbol{M}_1\boldsymbol{M}_2 \wedge \boldsymbol{M}_1\boldsymbol{M}_3 = \frac{1}{2} (a_2 - a_1) \wedge (a_3 - a_1)$$

$$= \frac{1}{2} (- a_3 \wedge a_2 - a_2 \wedge a_1 - a_1 \wedge a_3)$$

$$= - \boldsymbol{S}_1 - \boldsymbol{S}_2 - \boldsymbol{S}_3$$

Or $\boldsymbol{S}_1$, $\boldsymbol{S}_2$ et $\boldsymbol{S}_3$ sont orthogonaux deux à deux. On a donc, en prenant la norme au carré :

$$S^2 = S_1^2 + S_2^2 + S_3^2$$

Cette relation généralise à un tétraèdre dont trois faces sont perpendiculaires le théorème de Pythagore dans le triangle.

REMARQUES :

❑ La relation $\mathbf{S} + \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 = 0$ est un cas particulier de la relation plus générale $\iint_\Sigma d\mathbf{S} = 0$ pour toute surface Σ englobant un volume, elle-même cas particulier de la **formule du gradient**

$$\iint_\Sigma f\, d\mathbf{S} = \iiint_V \mathbf{grad}(f)\, dxdydz$$

énoncée dans chapitre *Gradient, divergence, rotationnel* L2/GRADIVRT.PDF.

❑ La relation $\mathbf{S} + \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 = 0$ sert à montrer en sciences physiques que le tenseur des contraintes dans la théorie des milieux continus est un morphisme additif. En effet, si une contrainte $F(\mathbf{S}_i)$ s'applique sur chaque face du tétraèdre, on a :

$F(\mathbf{S}) + F(\mathbf{S}_1) + F(\mathbf{S}_2) + F(\mathbf{S}_3) = 0$ équilibre des forces

et $F(-\mathbf{S}) = -F(\mathbf{S})$ si on change le signe de $\mathbf{S}$, on change le signe de la contrainte s'exerçant dans la direction de $\mathbf{S}$

donc $F(\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3) = F(-\mathbf{S}) = -F(\mathbf{S}) = F(\mathbf{S}_1) + F(\mathbf{S}_2) + F(\mathbf{S}_3)$

Passons maintenant à la dimension n. Soient $n + 1$ points O, M_1, ..., M_n dans un espace euclidien orienté E de dimension n, tels que les n vecteurs $a_i = \boldsymbol{OM}_i$ soient orthogonaux deux à deux. On note S_i la mesure $(n-1)$-dimensionnelle de l'enveloppe convexe de O et des M_j pour j différent de i, et S celle de l'enveloppe convexe des M_i. On a :

$$S_i = \frac{1}{(n-1)!} \left| \det(a_1, ..., a_{i-1}, a_{i+1}, ..., a_n) \right|$$

$$= \frac{1}{(n-1)!} \| a_1 \times ... \times a_{i-1} \times a_{i+1} \times ... \times a_n \|$$

$$= \| \boldsymbol{S}_i \| \qquad \text{en posant } \boldsymbol{S}_i = \frac{1}{(n-1)!} a_1 \times ... \times a_{i-1} \times a_{i+1} \times ... \times a_n$$

et

$$S = \frac{1}{(n-1)!} \left| \det(\boldsymbol{M}_1\boldsymbol{M}_2, \boldsymbol{M}_1\boldsymbol{M}_3, ..., \boldsymbol{M}_1\boldsymbol{M}_n) \right|$$

$$= \frac{1}{(n-1)!} \| \boldsymbol{M}_1\boldsymbol{M}_2 \times ... \times \boldsymbol{M}_1\boldsymbol{M}_n \|$$

$$= \frac{1}{(n-1)!} \| (a_2 - a_1) \times ... \times (a_n - a_1) \|$$

$$= \frac{1}{(n-1)!} \| a_2 \times ... \times a_n - a_1 \times a_3 \times ... \times a_n - a_2 \times a_1 \times a_4 \times ... \times a_n - ... - a_2 \times ... \times a_{n-1} \times a_1 \|$$

le premier terme de la somme ne possède aucun a_1, et les autres un seul a_1 issu successivement de chaque facteur. Il ne peut y avoir deux a_1 dans le même terme en raison du caractère alterné de $\times$.

$$= \| \boldsymbol{S}_1 - \boldsymbol{S}_2 + \boldsymbol{S}_3 - ... + (-1)^{n-1} \boldsymbol{S}_n \|$$

(on aurait pu se dispenser des changements de signe en orientant convenablement les $\boldsymbol{S}_i$)

Or les $\boldsymbol{S}_i$ sont orthogonaux deux à deux. En effet, pour $i \neq j$:

$$<\boldsymbol{S}_i, \boldsymbol{S}_j> = \frac{1}{(n-1)!^2} <a_1 \times ... \times a_{i-1} \times a_{i+1} \times ... \times a_n, a_1 \times ... \times a_{j-1} \times a_{j+1} \times ... \times a_n>$$

$$= \frac{1}{(n-1)!^2} \det(<a_k, a_l>_{k \neq i, l \neq j})$$

en utilisant la propriété (vii) du produit vectoriel

$$= 0$$

car tous les termes $<a_j, a_l>$, $l \neq j$, de la j-ème ligne du déterminant sont nuls (il en est de même de tous les termes de la i-ème colonne.

En élevant les normes au carré, on obtient $S^2 = \sum_{i=1}^{n} S_i^2$.

Exercices

1- Enoncés

Exo.1) Soient F et G deux sous-espaces vectoriels d'un espace vectoriel préhilbertien E.

a) Montrer que $(F + G)^\perp = F^\perp \cap G^\perp$.

b) Montrer que, si E est de dimension finie, alors $(F \cap G)^\perp = F^\perp + G^\perp$

Exo.2) Soit E un espace préhilbertien et trois éléments a, b, c tels que $b \neq c$. Montrer que :
$\| a - b \| = \| a - c \| \Rightarrow (b - a \mid b - c) > 0$
où (|) désigne le produit scalaire et $\| \ \|$ la norme euclidienne associée.

Exo.3) a) Quelle est la borne inférieure, pour a, b, c réels de $\displaystyle\int_{-1}^{1} (t^3 - at^2 - bt - c)^2 \, dt$?

b) Quelle est la borne inférieure, pour a, b, c réels de $\displaystyle\int_{0}^{\infty} (t^3 - at^2 - bt - c)^2 \, e^{-2t} \, dt$?

c) Soit E = $C^0([-1,1], \mathbf{R})$ et f élément de E. Que vaut $d = \text{Inf} \{ \displaystyle\int_{-1}^{1} (f(t) - g(t))^2 \, dt, g \text{ paire} \}$?

Calculer d lorsque $f(t) = e^t$.

Exo.4) Soit E un espace vectoriel euclidien orienté de dimension 3, et u un endomorphisme inversible de E.

a) Montrer que, pour tous vecteurs a et b de E, $u(a) \wedge u(b) = \det(u)\,(u^{-1})^*(a \wedge b)$, où $(u^{-1})^*$ est l'adjoint de u^{-1}. (On pourra considérer un produit scalaire $<u(a) \wedge u(b), c>$ avec c quelconque).

b) Comparer $f(a \wedge b)$ et $f(a) \wedge f(b)$ dans le cas où f est une isométrie vectorielle.

Exo.5) Soit E un espace vectoriel euclidien orienté de dimension 3, u un vecteur de E et f un endomorphisme de E. Quel est l'adjoint de l'endomorphisme $\varphi : x \rightarrow u \wedge f(x)$?

Exo.6) Vérifier que les matrices M suivantes sont orthogonales. Quelle est la nature de l'opérateur associé ?

a) $\dfrac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}$ b) $\dfrac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$

c) $\frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ 2 & -2 & -1 \\ -1 & -2 & 2 \end{pmatrix}$ d) $\frac{1}{7}\begin{pmatrix} -6 & -2 & 3 \\ -2 & -3 & -6 \\ 3 & -6 & 2 \end{pmatrix}$

Exo.7) Montrer que la matrice $\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ est dans $SO_3(\mathbf{R})$ si et seulement si a, b et c sont racines de l'équation $x^3 - x^2 + k = 0$, avec $0 \le k \le \frac{4}{27}$.

Exo.8) Paramétrisation des matrices de $SO_n(\mathbb{R})$:

a) Montrer que, si A est une matrice antisymétrique de $\mathcal{M}_n(\mathbb{R})$ alors $I_n - A$ est inversible et $M = (I_n - A)^{-1}(I_n + A)$ est une matrice orthogonale. Montrer en outre que M ne possède pas -1 comme valeur propre et que $\det(M) = 1$.

b) Réciproquement, montrer que toutes les matrices de $O_n(\mathbb{R})$ n'ayant pas -1 pour valeur propre s'obtiennent de cette façon.

c) Exemple dans $\mathbb{R}^2$: soit t un réel et $A = \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}$. Calculer $M = (I_2 - A)^{-1}(I_2 + A)$. Reconnaître une matrice de rotation. Quand t varie dans $\mathbb{R}$, obtient-on toute les matrices de rotations du plan $\mathbb{R}^2$?

d) Exemple dans $\mathbb{R}^3$: A antisymétrique est de la forme $\begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$. Soit $\Omega = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Alors, pour tout X de $\mathbb{R}^3$, $AX = \Omega \wedge X$. Toujours avec $M = (I_3 - A)^{-1}(I_3 + A)$, matrice de rotation de $\mathbb{R}^3$ d'après a), quels sont les vecteurs invariants par M et quel rapport ont-ils avec Ω ? Quel est l'angle de la rotation ? Quel rapport y a-t-il entre cet angle et $\| \Omega \|$? Quand A varie, obtient-on toutes les matrices de rotation de $\mathbb{R}^3$?

e) En 1770, Euler propose de paramétrer les rotations de l'espace tridimensionnel par la formule $\frac{1}{\alpha^2+\beta^2+\gamma^2+\delta^2}\begin{pmatrix} \alpha^2+\beta^2-\gamma^2-\delta^2 & -2\alpha\delta+2\beta\gamma & 2\alpha\gamma+2\beta\delta \\ 2\alpha\delta+2\beta\gamma & \alpha^2-\beta^2+\gamma^2-\delta^2 & -2\alpha\beta+2\gamma\delta \\ -2\alpha\gamma+2\beta\delta & 2\alpha\beta+2\gamma\delta & \alpha^2-\beta^2-\gamma^2+\delta^2 \end{pmatrix}$ avec $(\alpha, \beta, \gamma, \delta)$ non tous nuls. Y a-t-il un rapport avec la question précédente ? La paramétrisation d'Euler permet-elle d'obtenir toutes les matrices de rotations de $\mathbb{R}^3$?

Exo.9) a) Soit E espace vectoriel euclidien orienté de dimension 3, et f un endomorphisme de E. Montrer que le vecteur $Z = u \wedge f(u) + v \wedge f(v) + w \wedge f(w)$ ne dépend pas de la base orthonormée (u, v, w), directe ou non, choisie. Pour cela, on pourra calculer le produit vectoriel de Z par un vecteur x quelconque.

b) Que vaut Z pour une réflexion ? une projection orthogonale ? une rotation ?

c) Montrer que, pour tout endomorphisme f de E, il existe g autoadjoint et un vecteur Ω tel que :
$\forall\, x,\ f(x) = g(x) + \Omega \wedge x$
En déduire $u \wedge f(u) + v \wedge f(v) + w \wedge f(w)$ en fonction de Ω. Retrouver les résultats du b).

Exo.10) Soit A = (a_{ij}) une matrice orthogonale, élément de $\mathcal{O}_n(\mathbb{R})$. Montrer que $\left|\sum_{i,j} a_{ij}\right| \leq n$.

Exo.11) Soit $n \geq 2$. On note $(e_1, ..., e_n)$ la base orthonormée canonique de $\mathbb{R}^n$. Pour tout $k \in [\![1, n-1]\!]$, on note $R_k(\theta)$ l'élément de $SO_n(\mathbb{R})$ qui laisse $e_1, e_2, ..., e_{k-1}, e_{k+2}, ..., e_n$ invariants et dont la restriction au plan engendré par (e_k, e_{k+1}) et orienté par cette base est une rotation d'angle θ de ce plan. Soit G_n le groupe multiplicatif engendré par les $R_k(\theta)$, $k \in [\![1, n-1]\!]$, $\theta \in \mathbb{R}$.

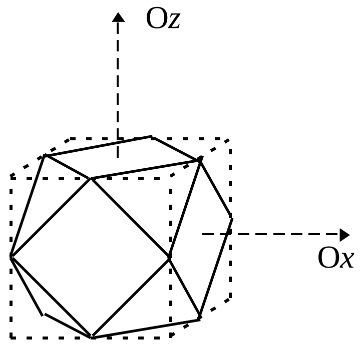
a) Soit u un vecteur unitaire de $\mathbb{R}^n$. Montrer que :

$$\exists\ \theta_1, ..., \theta_{n-1},\ R_{n-1}(\theta_{n-1})R_{n-2}(\theta_{n-2})...R_1(\theta_1)e_1 = u$$

b) Montrer que $G_n = SO_n(\mathbb{R})$.

c) Dans $\mathbb{R}^3$, on souhaite faire pivoter le plan $x = 0$ pour le faire coïncider avec le plan $x + y + z = 0$. Pour cela, on ne peut utiliser que deux rotations d'axe Ox ou Oz. Comment faire ?

d) Voici ci-dessous un cuboctaèdre. Il s'agit d'un polyèdre semi-régulier dont les sommets se situent au milieu des arêtes d'un cube.



La face inférieure, perpendiculaire à Oz est un carré. On souhaite effectuer une rotation du cuboctaèdre de façon à ce que la face inférieure soit un triangle. Pour cela, on ne peut effectuer que des rotations autour de Ox et de Oz. Expliquer comment procéder. Donner la matrice de la rotation finale.

e) On considère un tétraèdre régulier ABCD centré en (0,0,0). D est sur l'axe Oy avec $y > 0$, C est dans le plan Oxy avec $x > 0$, A et B sont symétriques par rapport à ce plan, la composante z de A étant négative. On considère l'axe passant par le milieu de [AD] et de [BC]. On souhaite effectuer une rotation d'angle θ selon cet axe, mais on ne peut procéder qu'à trois rotations d'axe Ox ou Oz. Comment faire ?

Exo.12) Pour A matrice orthogonale réelle, étudier la limite de la suite de terme général $\dfrac{I + A + ... + A^n}{n+1}$, où I désigne la matrice identité.

Exo.13) Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. Montrer qu'il existe Ω, élément du groupe orthogonal $\mathcal{O}_n(\mathbb{R})$ tel que $AA^T = \Omega^T A^T A \Omega$, où A^T désigne la matrice transposée de A.

Exo.14) Soit E euclidien de dimension n, et u un endomorphisme orthogonal.

a) Montrer que $u + u^{-1}$ est un endomorphisme symétrique.

b) En déduire qu'il existe une droite ou un plan de E stable par u.

c) Montrer qu'il existe une base orthonormée de E dans laquelle la matrice de u est de la forme $\begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ 0 & 0 & A_3 & \dots & 0 \\ \dots & & \dots & & \dots \\ 0 & 0 & \dots & 0 & A_p \end{pmatrix}$ où les A_i sont des blocs de taille 1×1 et valant ± 1, ou de taille 2×2 et de la forme $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$

d) Application : réduire $\frac{1}{3}\begin{pmatrix} 0 & -1 & -2 & 2 \\ -2 & -2 & 1 & 0 \\ 2 & -2 & 0 & -1 \\ 1 & 0 & 2 & 2 \end{pmatrix}$

Exo.15) Soit (E, $< | >$) un espace vectoriel euclidien, a un vecteur unitaire de E et α un réel non nul. On définit l'application f de E dans E par : $x \rightarrow f(x) = x + \alpha<x \mid a>a$.

a) f est-elle bijective ? Si oui, déterminer f^{-1}.

b) Montrer que f est symétrique.

c) Trouver les valeurs propres et les vecteurs propres de f. f est-elle diagonalisable ?

d) Trouver une condition nécessaire et suffisante pour f soit une isométrie. Quelle est la nature de f dans ce cas ?

Exo.16) a) Soit u un endomorphisme symétrique d'un espace vectoriel euclidien E. On suppose que : $\forall x \in E, <u(x), x> \geq 0$. Un tel endomorphisme est dit positif. Montrer qu'il existe un unique endomorphisme symétrique positif v tel que $v^2 = u$.

b) Soit f un automorphisme de E. Montrer qu'il existe un unique couple (g, h) avec g automorphisme orthogonal et h symétrique positif telle que $f = g \circ h$. Cette décomposition s'appelle **décomposition polaire d'un endomorphisme**.

c) Dans $\mathbb{R}^2$, donner la décomposition polaire de f de matrice $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$.

d) Montrer qu'une telle décomposition existe pour f endomorphisme de E non bijectif.

e) Dans $\mathbb{R}^2$, donner une décomposition de f de matrice $\begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$.

Exo.17) Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. Montrer que $|\det(A)| \leq (\frac{\mathrm{Tr}(A^{T}A)}{n})^{n/2}$ où Tr désigne la trace.

2- Solutions

Sol.1) a) Si x est orthogonal à F + G, alors x est orthogonal à tout vecteur $y + z$, $y \in F$, $z \in G$, donc x est orthogonal à tout y de F (en prenant $z = 0$) et à tout z de G (en prenant $y = 0$). Donc x est orthogonal à F et x est orthogonal à G, donc $x \in F^{\perp} \cap G^{\perp}$.

Réciproquement, si $x \in F^{\perp} \cap G^{\perp}$, alors x est orthogonal à tout y de F et à tout z de G, donc à tout $y + z$, $y \in F$, $z \in G$. Donc x est orthogonal à F + G.

b) En appliquant le a) à $F^{\perp}$ et à $G^{\perp}$, on a également :

$$(F^{\perp} + G^{\perp})^{\perp} = (F^{\perp})^{\perp} \cap (G^{\perp})^{\perp}$$

mais en dimension finie, $(F^{\perp})^{\perp} = F$, donc :

$$(F^{\perp} + G^{\perp})^{\perp} = F \cap G$$

Si on prend l'orthogonal, on obtient :

$$F^\perp + G^\perp = (F \cap G)^\perp$$

On peut aussi montrer que $F^\perp + G^\perp \subset (F \cap G)^\perp$ puis comparer les dimensions :

$$x \in F^\perp + G^\perp \Rightarrow \exists\, y \in F^\perp, \exists\, z \in G^\perp, x = y + z$$

donc, pour tout t de $F \cap G$:

$$<x, t> = <y + z, t> = <y, t> + <z, t> = 0 + 0 = 0$$

Donc $x \in (F \cap G)^\perp$.

Puis :

$$\begin{aligned}
\dim(F^\perp + G^\perp) &= \dim(F^\perp) + \dim(G^\perp) - \dim(F^\perp \cap G^\perp) \\
&= \dim(E) - \dim(F) + \dim(E) - \dim(G) - \dim((F + G)^\perp) \text{ d'après a)} \\
&= 2\dim(E) - \dim(F) - \dim(G) - (\dim(E) - \dim(F + G)) \\
&= \dim(E) - \dim(F) - \dim(G) + \dim(F) + \dim(G) - \dim(F \cap G) \\
&= \dim(E) - \dim(F \cap G) \\
&= \dim((F \cap G)^\perp)
\end{aligned}$$

REMARQUE :

Le résultat du b) peut être faux en dimension infinie. Dans le cours, on a donné des exemples d'espaces préhilbertiens E et d'un sous-espace vectoriel $F \neq E$ tel que $F^\perp = \{0\}$. Soit G un supplémentaire de F dans E. Comme $F \neq E$, $G \neq \{0\}$. Donc $G^\perp \neq E$, car si $G^\perp = E$, alors on aurait $G \subset (G^\perp)^\perp = E^\perp = \{0\}$. On observe alors que :

$$(F \cap G)^\perp = \{0\}^\perp = E$$

mais $F^\perp + G^\perp = G^\perp \neq E$

donc $(F \cap G)^\perp \neq F^\perp + G^\perp$.

Sol.2) $\| a - b \| = \| a - c \|$

$\Leftrightarrow \quad a^2 - 2(a \mid b) + b^2 = a^2 - 2(a \mid c) + c^2$ en élevant au carré

$\Leftrightarrow \quad 2(a \mid c) - 2(a \mid b) + b^2 - c^2 = 0$

$\Leftrightarrow \quad 2(a \mid c - b) - (b + c, c - b) = 0$

$\Leftrightarrow \quad (a - \dfrac{b + c}{2} \mid c - b) = 0$

Donc :

$$(b - a \mid b - c) = (\frac{b - c}{2} + \frac{b + c}{2} - a \mid b - c) = \frac{1}{2} \| b - c \|^2 > 0$$

On peut donner des interprétations géométriques de ce résultat.

<u>Interprétation 1</u> : on fait varier a.

L'ensemble des a tels que $\| a - b \| = \| a - c \|$ est l'hyperplan médiateur H_1 de $[b, c]$, ce qui exprime simplement le fait que $(a - \dfrac{b + c}{2} \mid c - b) = 0$. H_1 est un hyperplan affine passant par $\dfrac{b + c}{2}$ de direction l'hyperplan vectoriel d'équation $(a \mid c - b) = 0$, a étant l'inconnue de cette équation.

L'ensemble des a tels que $(b - a \mid b - c) > 0$ est le demi-espace ouvert U contenant le point c et limité par l'hyperplan affine H_2 d'équation $(b - a \mid b - c) = 0$. Cet hyperplan affine a pour direction l'hyperplan vectoriel d'équation $(a \mid b - c) = 0$. H_1 et H_2 sont donc parallèles puisqu'ils ont même direction vectorielle. Ils sont disjoints car H_2 passe par b ce que ne fait pas H_1. Or $\dfrac{b + c}{2}$ élément de H_1 appartient à U. Donc H_1 est inclus dans U.

<u>Interprétation 2</u> : on fait varier b.

L'ensemble des b différents de c tels que $\| a - b \| = \| a - c \|$ est la sphère S de centre a et de rayon $\| a - c \|$, privée du point c.

L'ensemble des b tels que $(b - a \mid b - c) > 0$, inégalité qui équivaut à :

$$(b - \frac{a+c}{2} - \frac{a-c}{2} \mid b - \frac{a+c}{2} + \frac{a-c}{2}) > 0$$

ou encore à $\| b - \frac{a+c}{2} \| > \| \frac{a-c}{2} \|$

est l'extérieur ouvert U de la boule B de centre $\frac{a+c}{2}$ et de rayon $\| \frac{a-c}{2} \|$, ou encore de diamètre $[a, c]$. La sphère S est tangente à la boule B en c, mais est de rayon plus grand. Donc S privée de c est incluse dans U.

Interprétation 3 : on fait varier c.

L'ensemble des $c \neq b$ tels que $\| a - b \| = \| a - c \|$ est la sphère S de centre a et de rayon $\| a - b \|$, privée du point b.

L'ensemble des c tels que $(b - a \mid b - c) > 0$ est le demi-espace U limité par l'hyperplan affine d'équation $(b - a \mid b - c) = 0$, c étant l'inconnue de l'équation. Cet hyperplan passe par b et est normal à $b - a$. La droite (ab) est donc perpendiculaire à l'hyperplan en b. Donc la sphère est tangente à l'hyperplan en b. Comme a appartient à U, la sphère S de centre a privée de b est incluse dans U.

Sol.3) a) Soit $<P,Q> = \int_{-1}^{1} P(t)Q(t)dt$ produit scalaire sur les polynômes. Pour la norme induite par ce produit scalaire, l'intégrale demandée est le carré de la distance du polynôme X^3 au sous-espace engendré par 1, X et X^2. Cette distance est obtenue en projetant orthogonalement X sur $\mathbf{R}_2[X]$. On cherche donc a, b et c tels que $X^3 - aX^2 - bX - c$ soit orthogonal à 1, X et X^2. On résout donc :

$$\int_{-1}^{1} (x^3 - ax^2 - bx - c)\, dx = 0$$

$$\int_{-1}^{1} (x^3 - ax^2 - bx - c)\, x\, dx = 0$$

$$\int_{-1}^{1} (x^3 - ax^2 - bx - c)\, x^2\, dx = 0$$

ce qui conduit après calcul des intégrales et résolution du système à $a = 0$, $b = \frac{3}{5}$ et $c = 0$.

b) On procède de même avec le produit scalaire $<P,Q> = \int_{0}^{\infty} P(t)Q(t)e^{-2t}\, dt$. On est amené à résoudre :

$$\int_{0}^{\infty} (x^3 - ax^2 - bx - c)\, e^{-2x}\, dx = 0$$

$$\int_{0}^{\infty} (x^3 - ax^2 - bx - c)\, x\, e^{-2x}\, dx = 0$$

$$\int_{0}^{\infty} (x^3 - ax^2 - bx - c)\, x^2\, e^{-2x}\, dx = 0$$

D'où $a = \frac{9}{2}$, $b = -\frac{9}{2}$ et $c = \frac{3}{4}$.

c) On munit E du produit scalaire $<f, g> = \int_{-1}^{1} f(t)g(t)\,dt$. Pour ce produit scalaire, le sous-espace vectoriel F des fonctions paires est orthogonal au sous-espace vectoriel G des fonctions impaires. d est le carré de la distance de f au sous-espace vectoriel des fonctions paires. Il suffit donc de décomposer f selon la somme $E = F \oplus G$, à savoir :

$$\forall\, t,\ f(t) = \frac{f(t) + f(-t)}{2} + \frac{f(t) - f(-t)}{2}$$

et de prendre pour g la composante élément de F, à savoir $\frac{f(t) + f(-t)}{2}$. Le minimum cherché est $\int_{-1}^{1} (\frac{f(t) - f(-t)}{2})^2\,dt$.

Pour $f(t) = e^t$, $f(t) = \mathrm{ch}(t) + \mathrm{sh}(t)$.

$$d = \int_{-1}^{1} \mathrm{sh}^2(t)\,dt = \int_{-1}^{1} \frac{\mathrm{ch}(2t) - 1}{2}\,dt = \int_{0}^{1} \mathrm{ch}(2t) - 1\,dt = \frac{\mathrm{sh}(2)}{2} - 1$$

Sol.4) a) Pour tout vecteur c, on a :

$$<u(a) \wedge u(b), c> = [(u(a), u(b), c] \qquad \text{(produit mixte, voir L1/DETERMNT.PDF)}$$
$$= [(u(a), u(b), u(u^{-1}(c))]$$
$$= \det(u)\,[a, b, u^{-1}(c)]$$
$$= \det(u) <a \wedge b, u^{-1}(c)>$$
$$= \det(u) <(u^{-1})^*(a \wedge b), c>$$

d'où le résultat.

b) Si f est une isométrie, f^{-1} aussi et donc $(f^{-1})^* = f$. Donc, d'après le a) :

$$f(a) \wedge f(b) = \det(f)\, f(a \wedge b) = \pm\, f(a) \wedge f(b)$$

le signe dépendant du fait que f est directe ou non. Pour une isométrie directe, le signe est +. On retrouve le fait qu'une telle isométrie conserve le produit vectoriel (i.e. l'image d'un produit vectoriel par une rotation vectorielle est le produit vectoriel des images). Par contre, une isométrie indirecte en change le signe.

Sol.5) Pour tout x, y, on a :

$$<\varphi(x), y> = <u \wedge f(x), y> = [u, f(x), y] \qquad \text{(produit mixte)}$$
$$= [y, u, f(x)]$$
$$= <y \wedge u, f(x)>$$
$$= <f^*(y \wedge u), x>$$

Par ailleurs :

$$<\varphi(x), y> = <x, \varphi^*(y)>$$

donc, pour tout x :

$$<\varphi^*(y), x> = <f^*(y \wedge u), x>$$

donc :

$$\varphi^*(y) = f^*(y \wedge u)$$

L'adjoint φ^* cherché est donc l'application $y \to f^*(y \wedge u)$.

On peut aussi chercher l'adjoint de $\psi : x \to u \wedge x$ qui est $\psi^* : y \to y \wedge u$ puis composer avec f. On a alors, puisque $\varphi = \psi \circ f$:

$$\varphi^* = (\psi \circ f)^* = f^* \circ \psi^*$$

ce qui donne le même résultat.

Sol.6) On montrera que chaque matrice est orthogonale en vérifiant que leurs colonnes forment un système orthonormé. On calculera donc la norme de chaque colonne (qui doit valoir 1), et les produits scalaires de deux colonnes différentes (qui doit valoir 0). Cela nécessite 6 calculs, alors que la vérification de la propriété $M^TM = I_3$ nécessite le calcul de 9 coefficients.
Pour déterminer la nature de chaque isométrie, on commencera par chercher l'ensemble des vecteurs invariants.

a) Les vecteurs invariants forment une droite engendrée par le vecteur unitaire $n = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}$. Il s'agit donc d'une rotation. Le vecteur unitaire $v = \begin{pmatrix}0\\1\\0\end{pmatrix}$ est orthogonal à n, de même que $w = n \wedge v = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}$. La base (v, w, n) est directe et l'angle θ de la rotation est le même que celui de la rotation plane, restriction de la rotation au plan $\text{Vect}(v, w)$. L'image de v est $Mv = \frac{1}{3}\begin{pmatrix}2\\1\\-2\end{pmatrix}$. On a alors :

$$\cos(\theta) = \langle Mv, v\rangle = \frac{1}{3}$$

$$\sin(\theta) = \langle Mv, w\rangle = -\frac{4}{3\sqrt{2}} = -\frac{2\sqrt{2}}{3}$$

θ est donc élément de $]-\frac{\pi}{2}, 0[$, et vaut $-\arcsin(\frac{2\sqrt{2}}{3})$.

On aurait pu aussi calculer $M - M^T = \frac{4}{3}\begin{pmatrix}0 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0\end{pmatrix}$, de la forme $\begin{pmatrix}0 & -c & b\\ c & 0 & -a\\ -b & a & 0\end{pmatrix}$, avec :

$$2\sin(\theta)n = \begin{pmatrix}a\\b\\c\end{pmatrix} = -\frac{4}{3}\begin{pmatrix}1\\0\\1\end{pmatrix} \text{ et } \cos(\theta) = \frac{\text{Tr}(M) - 1}{2} = \frac{1}{3}$$

On retrouve $n = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}$ et $\sin(\theta) = -\frac{2\sqrt{2}}{3}$.

On aurait pu choisir aussi $n = -\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}$ et $\sin(\theta) = \frac{2\sqrt{2}}{3}$. Dans ce cas, on aurait $\theta = \arcsin(\frac{2\sqrt{2}}{3})$.

b) L'ensemble des vecteurs invariants est le plan $-x + y + z = 0$. Il s'agit donc d'une réflexion par rapport à ce plan.
On peut aussi remarquer que la matrice est symétrique. Etant également orthogonale, on a donc $M^TM = I_3$ et $M^T = M$, donc $M^2 = I_3$, donc M est la matrice d'une symétrie. M étant orthogonale, il s'agit d'une symétrie orthogonale, mais on ignore à ce stade s'il s'agit d'une réflexion (symétrie orthogonale par rapport à un plan) ou d'un demi-tour (symétrie orthogonale par rapport à une droite ou rotation d'angle π). La conclusion s'obtient par le calcul de $\det(M) = -1$, qui exclut le demi-tour.
c) L'ensemble des vecteurs invariants est $\{0\}$. Il s'agit donc d'une antirotation. Son axe se trouve en cherchant l'ensemble des vecteurs changés en leur opposé. On trouve la droite engendrée par

$n = \dfrac{1}{\sqrt{11}}\begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$. La restriction de l'antirotation au plan orthogonal à n est une rotation de ce plan dont on cherche l'angle θ. Soit $v = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ un vecteur unitaire orthogonal à n. Son image par M est $\dfrac{1}{3\sqrt{2}}\begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$. On a :

$$\cos(\theta) = \langle Mv, v\rangle = \frac{5}{6}$$

et on vérifiera que $v \wedge Mv$ est colinéaire est de même sens que n, donc $\theta \in [0, \pi]$.
L'isométrie est donc la composée de la réflexion par rapport au plan (orthogonal à n) d'équation $-x + 3y + z = 0$ et de la rotation d'axe n et d'angle $\arccos(\frac{5}{6})$.

On aurait pu aussi calculer $\cos(\theta)$ au moyen de $\dfrac{\mathrm{Tr(M)} + 1}{2} = \dfrac{5}{6}$ et utiliser $M - M^T = \dfrac{1}{3}\begin{pmatrix} 0 & -1 & 3 \\ 1 & 0 & 1 \\ -3 & -1 & 0 \end{pmatrix}$, conduisant à $2\sin(\theta)n = \dfrac{1}{3}\begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$, qui redonne $n = \dfrac{1}{\sqrt{11}}\begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$ et $\sin(\theta) = \dfrac{\sqrt{11}}{6}$.

d) L'ensemble des vecteurs invariants est la droite engendrée par $\dfrac{1}{\sqrt{14}}\begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}$. Il s'agit donc d'une rotation. Son angle θ est tel que $\cos(\theta) = \dfrac{\mathrm{Tr(M)} - 1}{2} = -1$, donc $\theta = \pi$. Il 'agit d'un demi-tour.
On peut aussi remarquer que la matrice est symétrique. Etant également orthogonale, on a $M^TM = I_3$ et $M^T = M$, donc $M^2 = I_3$, donc M est la matrice d'une symétrie. M étant orthogonale, il s'agit d'une symétrie orthogonale, soit une réflexion, soit un demi-tour. Comme $\det(M) = 1$, il s'agit d'un demi-tour.

Sol.7) On écrit que les colonnes sont orthonormées et que le déterminant vaut 1, ce qui donne :

$$\begin{cases} a^2 + b^2 + c^2 = 1 \\ ab + ac + bc = 0 \\ a^3 + b^3 + c^3 - 3abc = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} (a + b + c)^2 = 1 \\ ab + ac + bc = 0 \\ (a + b + c)^3 - 3(ab + ac + bc)(a + b + c) = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} a + b + c = 1 \\ ab + ac + bc = 0 \end{cases}$$

$\Leftrightarrow$ a, b, c sont les racines réelles d'un polynôme de la forme $(X - a)(X - b)(X - c) = X^3 - X^2 + k$, où $k = -abc$. Ce polynôme a pour dérivée $3X^2 - 2X$ qui s'annule en $x = 0$ et $x = \dfrac{2}{3}$. Le polynôme croît de $-\infty$ à k, puis décroît jusqu'à $-\dfrac{4}{27} + k$ puis croît jusqu'à $+\infty$. Il y a trois racines réelles si et seulement si $0 \le k \le \dfrac{4}{27}$.

Sol.8) a) On cherche le noyau de $I_n - A$:

$(I_n - A)X = 0 \Rightarrow AX = X$

$\Rightarrow \langle X, X\rangle = \langle X, AX\rangle = X^T AX = \langle A^T X, X\rangle = -\langle AX, X\rangle = -\langle X, X\rangle$

donc $\langle X,X\rangle = 0$ donc $X = 0$. Donc $I_n - A$ est inversible.

Montrons que M est orthogonale :

$M^T M = (I_n + A)^T ((I_n - A)^{-1})^T (I_n - A)^{-1} (I_n + A)$

$= (I_n + A^T) ((I_n - A)^T)^{-1} (I_n - A)^{-1} (I_n + A)$

$= (I_n - A) (I_n + A)^{-1} (I_n + A) (I_n - A)^{-1}$ car les matrices $(I_n - A)^{-1}$ et $I_n + A$ commutent

$= I_n$

Donc M est orthogonale.

Par l'absurde, si M possède la valeur propre –1, soit X un vecteur propre de M pour cette valeur propre (donc non nul) :

$MX = -X \Rightarrow (I_n - A)^{-1}(I_n + A)X = -X$

$\Rightarrow (I_n + A)X = -(I_n - A)X$

$\Rightarrow X + AX = -X + AX$

$\Rightarrow X = 0$

ce qui contredit l'hypothèse. Donc – 1 n'est pas valeur propre de M.

Enfin, $\det(M) = \dfrac{\det(I_n + A)}{\det(I_n - A)} = 1$ car $\det(I_n + A) = \det((I_n + A)^T) = \det(I_n - A)$.

b) Réciproquement, soit M une matrice orthogonale n'ayant pas la valeur propre –1. On cherche A antisymétrique telle que :

$(I_n - A)M = I_n + A$

$\Leftrightarrow \quad M - AM = I_n + A$

$\Leftrightarrow \quad A(I_n + M) = M - I_n$

$I_n + M$ est inversible puisque M ne possède pas de valeur propre – 1. On peut donc prendre :

$A = (M - I_n)(I_n + M)^{-1}$

On vérifie que A est antisymétrique.

$A^T = ((I_n + M)^{-1})^T (M - I_n)^T$

$= ((I_n + M)^T)^{-1} (M^T - I_n)$

$= (I_n + M^{-1})^{-1} (M^{-1} - I_n)$ car $M^T = M^{-1}$

$= (M + I_n)^{-1} M M^{-1}(I_n - M)$

$= (M + I_n)^{-1}(I_n - M) = -A$

Accessoirement, on a montré la propriété suivante :

M est orthogonale et n'admet pas la valeur propre – 1

$\Rightarrow \quad \exists$ A antisymétrique, $M = (I_n - A)^{-1}(I_n + A)$

$\Rightarrow \quad \det(M) = 1$ d'après a)

Dans le chapitre L3/HERMITN.PDF (ou plus bas en exercice), on montre que toute matrice orthogonale M est semblable à une matrice diagonale par blocs, avec des blocs 1×1 valant ± 1 et des blocs 2×2 de la forme $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. S'il n' y a pas de valeur propre –1, il n'y a que des blocs diagonaux 1×1 valant 1 ou des blocs 2×2 de déterminant 1, et retrouve le fait que $\det(M) = 1$.

c) $I_n - A = \begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix}$ a pour inverse $\dfrac{1}{1+t^2}\begin{pmatrix} 1 & -t \\ t & 1 \end{pmatrix}$ et le produit de cet inverse par $I_n + A = \begin{pmatrix} 1 & -t \\ t & 1 \end{pmatrix}$ donne :

$$\begin{pmatrix} \dfrac{1-t^2}{1+t^2} & -\dfrac{2t}{1+t^2} \\ \dfrac{2t}{1+t^2} & \dfrac{1-t^2}{1+t^2} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \text{ en posant } \tan(\frac{\theta}{2}) = t$$

Quand t varie de $-\infty$ à $+\infty$, $\frac{\theta}{2}$ varie de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$, et θ varie de $-\pi$ à π. On obtient toutes les rotations à l'exception du demi-tour, correspondant à $\theta = \pi$ ou $t = \infty$. Le demi-tour est la seule rotation du plan admettant la valeur propre -1.

d) X est invariant par M si et seulement si :

$(I_3 - A)^{-1}(I_3 + A)X = X \Leftrightarrow (I_3 + A)X = (I_3 - A)X \Leftrightarrow AX = 0 \Leftrightarrow \Omega \wedge X = 0 \Leftrightarrow$ X est colinéaire à Ω.

Si $\Omega = 0$, alors $A = 0$ et $M = I_3$

si $\Omega \neq 0$, l'axe de la rotation M est donc engendré par Ω. Cherchons l'expression de la rotation dans le plan orthogonal à Ω. Soit $e_3 = \frac{\Omega}{\| \Omega \|}$, e_1 unitaire orthogonal à e_3 et $e_2 = e_3 \wedge e_1$. Alors :

$Ae_1 = \Omega \wedge e_1 = \| \Omega \| \, e_3 \wedge e_1 = \| \Omega \| \, e_2$

$Ae_2 = \Omega \wedge e_2 = \| \Omega \| \, e_3 \wedge e_2 = -\| \Omega \| \, e_1$

Donc la matrice dans la base (e_1, e_2) de la restriction au plan Vect(e_1, e_2) de l'endomorphisme associé à A est :

$$\begin{pmatrix} 0 & -\| \Omega \| \\ \| \Omega \| & 0 \end{pmatrix}$$

qui est du type de la matrice antisymétrique du b), avec $t = \| \Omega \|$. Donc la matrice de la restriction de l'endomorphisme associé à M est :

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \dfrac{1-t^2}{1+t^2} & -\dfrac{2t}{1+t^2} \\ \dfrac{2t}{1+t^2} & \dfrac{1-t^2}{1+t^2} \end{pmatrix} = \frac{1}{1 + \| \Omega \|^2} \begin{pmatrix} 1 - \| \Omega \|^2 & -2\| \Omega \| \\ 2\| \Omega \| & 1 - \| \Omega \|^2 \end{pmatrix}$$

Donc l'angle θ de la rotation M vérifie $\cos(\theta) = \frac{1 - \| \Omega \|^2}{1 + \| \Omega \|^2}$, $\sin(\theta) = \frac{2\| \Omega \|}{1 + \| \Omega \|^2}$, ou encore $\tan(\frac{\theta}{2}) = \| \Omega \|$.

Comme dans le c), on n'obtient pas les demi-tours puisque $\tan(\frac{\theta}{2})$ n'est pas défini dans ce cas, à moins de faire tendre Ω vers l'infini.

e) Pour $\alpha = 1$, $\beta = a$, $\gamma = b$, $\delta = c$, on vérifiera que la matrice d'Euler est indentique à notre matrice $M = (I_3 - A)^{-1}(I_3 + A)$ avec $A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$.

α non nul ne donne rien de plus puisque la matrice d'Euler associée à $(\alpha, \beta, \gamma, \delta)$ est la même que celle associée à $(1, \frac{\beta}{\alpha}, \frac{\gamma}{\alpha}, \frac{\delta}{\alpha})$.

Pour $\alpha = 0$, la matrice d'Euler vaut $\frac{1}{\beta^2+\gamma^2+\delta^2} \begin{pmatrix} \beta^2-\gamma^2-\delta^2 & 2\beta\gamma & 2\beta\delta \\ 2\beta\gamma & -\beta^2+\gamma^2-\delta^2 & 2\gamma\delta \\ 2\beta\delta & 2\gamma\delta & -\beta^2-\gamma^2+\delta^2 \end{pmatrix}$. On vérifiera qu'il s'agit d'une matrice orthogonale, ayant pour vecteurs invariants les éléments de la droite engendrée

par $\begin{pmatrix}\beta\\ \gamma\\ \delta\end{pmatrix}$. Il s'agit donc d'une rotation. Elle a pour trace $-1 = 1 + 2\cos(\theta)$, donc $\theta = \pi$. Il s'agit donc d'un demi-tour.
La paramétrisation d'Euler permet d'obtenir toutes les rotations de $\mathbb{R}^3$.

Sol.9) a) On utilise la formule du double produit vectoriel $(a \wedge b) \wedge c = -<b, c>\, a + <a, c>\, b$, ainsi que la propriété $<f(a), b> = <a, f^*(b)>$, où f^* est l'adjoint de f.

$$\begin{aligned} Z \wedge x &= (u \wedge f(u) + v \wedge f(v) + w \wedge f(w)) \wedge x \\ &= -<f(u), x>u + <x, u>f(u) - <f(v), x>v + <x, v>f(v) - <f(w), x>w + <x, w>f(w) \\ &= -<f(u), x>u - <f(v), x>v - <f(w), x>w + <x, u>f(u) + <x, v>f(v) + <x, w>f(w) \\ &= \underbrace{-<u, f^*(x)>u - <v, f^*(x)>v - <w, f^*(x)>w}_{-f^*(x)} + \underbrace{f(<x, u>u + <x, v>v + <x, w>w)}_{f(x)} \end{aligned}$$

Ainsi, $Z \wedge x = f(x) - f^*(x)$. L'application $x \to Z \wedge x$ ne dépend pas de (u, v, w). Comme cette application permet de reconstituer Z, il en résulte que Z ne dépend pas de (u, v, w).
b) Pour une réflexion ou une projection orthogonale, $f = f^*$ donc $Z = 0$. Plus généralement, $Z = 0$ si f est autoadjoint. Pour une rotation d'angle θ et d'axe dirigé par le vecteur unitaire ω, $Z = 2\sin(\theta)\omega$; pour le voir, prendre une base orthonormée (u, v, w) avec $w = \omega$.
c) On utiliser la décomposition de f en la somme d'un endomorphisme symétrique et d'un endomorphisme antisymétrique :

$$\begin{aligned} f(x) &= \frac{1}{2}(f(x) + f^*(x)) + \frac{1}{2}(f(x) - f^*(x)) \\ &= g(x) + \frac{1}{2} Z \wedge x \end{aligned}$$

avec $g = \dfrac{f + f^*}{2}$ et $\Omega = \dfrac{Z}{2}$

On a donc $u \wedge f(u) + v \wedge f(v) + w \wedge f(w) = Z = 2\Omega$.
Pour une rotation d'angle θ et d'axe dirigé par le vecteur unitaire ω, on a :

$$f(x) = \cos(\theta)\, x + (1 - \cos(\theta))<x,\omega>\omega + \sin(\theta)\omega \wedge x$$

L'endomorphisme autoadjoint g associé à f est :

$$g(x) = \cos(\theta)\, x + (1 - \cos(\theta))<x,\omega>\omega$$

et $\quad \Omega = \sin(\theta)\omega$

On retrouve le résultat $Z = 2\sin(\theta)\omega$ du b).

Sol.10) Soit $U = \begin{pmatrix}1\\ \dots\\ 1\end{pmatrix}$ et V_j la j-ème colonne de A. Alors :

$$\left|\sum_{i,j} a_{ij}\right| = \left|<U, \sum_{j=1}^{n} V_j>\right| \leq \| U \|\, \| \sum_{j=1}^{n} V_j \| \qquad \text{d'après l'inégalité de Cauchy-Schwarz}$$

or $\| U \| = \sqrt{n}$ et $<\sum_{j=1}^{n} V_j, \sum_{j=1}^{n} V_j> = \sum_{j=1}^{n} \| V_j \|^2 = n$ donc $\| \sum_{j=1}^{n} V_j \| = \sqrt{n}$

Sol.11) a) On montre le résultat demandé par récurrence sur n. Si $n = 2$, $\exists\ \theta_1$, $u = \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix}$ et on a $R_1(\theta_1)e_1 = u$.

Supposons le résultat prouvé pour $n - 1$, montrons-le pour n. Posons $u = \begin{pmatrix} u_1 \\ ... \\ u_n \end{pmatrix}$.

Si $u_1 = 1$, alors $u = e_1$ et on prend tous les θ_k nuls de sorte que $R_{n-1}(\theta_{n-1})R_{n-2}(\theta_{n-2})...R_1(\theta_1) = I_n$.
Si $u_1 = -1$, alors $u = -e_1$ et on prend $\theta_1 = \pi$ et tous les autres θ_k nuls.

Sinon, $\exists\ \theta_1 \in\ ]0, \pi[$, $u_1 = \cos(\theta_1)$. $R_1(\theta_1)e_1 = \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \\ 0 \\ ... \\ 0 \end{pmatrix}$ a alors même première composante que u, et dans $\mathbf{R}^{n-1}$, $\begin{pmatrix} \sin(\theta_1) \\ 0 \\ ... \\ 0 \end{pmatrix}$ est non nul et a même norme que $\begin{pmatrix} u_2 \\ ... \\ u_n \end{pmatrix}$. D'après l'hypothèse de récurrence, il existe $S = S_{n-2}(\theta_{n-1})S_{n-3}(\theta_{n-2})...S_1(\theta_2)$ tel que $S\begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \end{pmatrix} = \dfrac{1}{\sin(\theta_1)}\begin{pmatrix} u_2 \\ ... \\ u_n \end{pmatrix}$ et donc $S\begin{pmatrix} \sin(\theta_1) \\ 0 \\ ... \\ 0 \end{pmatrix} = \begin{pmatrix} u_2 \\ ... \\ u_n \end{pmatrix}$.

Posons $R_{k+1}(\theta_{k+1})$ la matrice par blocs égale à $\begin{pmatrix} 1 & 0 \\ 0 & S_k(\theta_{k+1}) \end{pmatrix}$. Soit $T = R_{n-1}(\theta_{n-1})R_{n-2}(\theta_{n-2})...R_2(\theta_2)$, de sorte que T est égale à la matrice par blocs $\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}$. On a alors :

$$T\begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \\ 0 \\ ... \\ 0 \end{pmatrix} = u = R_{n-1}(\theta_{n-1})R_{n-2}(\theta_{n-2})...R_2(\theta_2)R_1(\theta_1)e_1$$

En dimension 3, il suffit de deux rotations, une d'axe Oz puis une d'axe Ox pour amener le vecteur unitaire directeur de Ox sur n'importe quel vecteur unitaire.

b) On que $G_n = SO_n(\mathbf{R})$ par récurrence sur n. Pour $n = 2$, c'est trivial puisque toute matrice de $SO_2(\mathbf{R})$ est exactement de la forme $R_1(\theta)$. Supposons la propriété prouvée au rang $n - 1$. Soit A une matrice de $SO_n(\mathbf{R})$. Notons u la première colonne de A. D'après le a), $\exists\ R \in G_n$ telle que la première colonne de R soit u. Donc $R^{-1}A$ a pour première colonne e_1. Comme $R^{-1}A$ est élément de SO_n, $R^{-1}A$ est nécessairement une matrice par blocs de la forme $\begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix}$, avec $A' \in SO_{n-1}(\mathbf{R})$. On applique l'hypothèse de récurrence sur A'. A' est produit de matrices $S_k(\theta) \in G_{n-1}$, donc $\begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix}$ est le produit des matrices $R_{k+1}(\theta)$ correspondantes telles qu'on les a définies dans le a), donc A est le produit de ces $R_{k+1}(\theta)$ par R.

On peut obtenir A avec au plus $\dfrac{n(n-1)}{2}$ matrices engendrant G_n.

En dimension 3, toute matrice de rotation est le produit d'au plus 3 matrices de rotation d'axe Ox ou Oz. Une fois amené le vecteur e_1 sur la première colonne de la matrice de rotation A à décomposer au moyen de deux rotations R_1 (d'axe Oz) puis R_2 (d'axe Ox), la décomposition cherchée est :

$$A = R_2R_1(R_1^{\mathsf{T}}R_2^{\mathsf{T}}A)$$

En effet, si u est la première colonne de A, alors $R_2R_1e_1 = u$ et $Ae_1 = u$ donc :

$$R_1^{\mathsf{T}}R_2^{\mathsf{T}}A\, e_1 = R_1^{\mathsf{T}}R_2^{\mathsf{T}}u = R_1^{-1}R_2^{-1}u = e_1$$

donc $R_1^{\mathsf{T}}R_2^{\mathsf{T}}A$ est une rotation d'axe Ox.

c) On applique le a) de façon à envoyer le vecteur $\begin{pmatrix}1\\0\\0\end{pmatrix}$ normal au plan $x = 0$ sur le vecteur $\dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$ normal au plan $x + y + z$. Soit θ_1 tel que $\cos(\theta_1) = \dfrac{1}{\sqrt{3}}$ et $\sin(\theta_1) = \dfrac{\sqrt{2}}{\sqrt{3}}$. On a :

$$R_1(\theta_1) = \begin{pmatrix}\frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} & 0\\ \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0\\ 0 & 0 & 1\end{pmatrix}$$

et $\quad R_1(\theta_1)\begin{pmatrix}1\\0\\0\end{pmatrix} = \dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\\sqrt{2}\\0\end{pmatrix}$

Il reste à envoyer le vecteur $\begin{pmatrix}1\\\sqrt{2}\\0\end{pmatrix}$ sur $\begin{pmatrix}1\\1\\1\end{pmatrix}$ et on l'obtient par une rotation d'axe Ox d'angle $\dfrac{\pi}{4}$:

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}\begin{pmatrix}1\\\sqrt{2}\\0\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix}$$

Donc la rotation finale cherchée est $\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} & 0\\ \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0\\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{\sqrt{2}}{2}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{2}\end{pmatrix}$

d) Prenons 2 comme longueur de côté du cube. On se propose d'amener le triangle de sommets $\begin{pmatrix}-1\\-1\\0\end{pmatrix}$, $\begin{pmatrix}-1\\0\\-1\end{pmatrix}$ et $\begin{pmatrix}0\\-1\\-1\end{pmatrix}$ à l'horizontale.

On effectue une rotation d'axe Oz d'angle $\frac{\pi}{4}$, de matrice $\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$, de façon à rendre le côté inférieur du triangle parallèle à Ox. Les sommets du triangle occupent maintenant les positions respectives $\begin{pmatrix} 0 \\ -\sqrt{2} \\ 0 \end{pmatrix}$, $\begin{pmatrix} -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ -1 \end{pmatrix}$ et $\begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ -1 \end{pmatrix}$.

Il convient ensuite d'effectuer une rotation d'axe Ox et d'angle $\arctan(\sqrt{2})$ de matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{pmatrix}$

de sorte que les sommets du triangle occupent les positions finales $\begin{pmatrix} 0 \\ -\frac{\sqrt{6}}{3} \\ -\frac{2\sqrt{3}}{3} \end{pmatrix}$, $\begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{3}}{3} \end{pmatrix}$, $\begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{3}}{3} \end{pmatrix}$. Ils se trouvent bien à la même hauteur.

Le produit des deux matrices vaut $\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{pmatrix}$.

e) Pour un tétraèdre de côté 1, les coordonnées des points sont :

$A = (-\frac{\sqrt{3}}{6}, -\frac{\sqrt{6}}{12}, -\frac{1}{2})$

$B = (-\frac{\sqrt{3}}{6}, -\frac{\sqrt{6}}{12}, \frac{1}{2})$

$C = (\frac{\sqrt{3}}{3}, -\frac{\sqrt{6}}{12}, 0)$

$D = (0, \frac{\sqrt{6}}{4}, 0)$

Le milieu de [AD] est M = $(-\frac{\sqrt{3}}{12}, \frac{\sqrt{6}}{12}, -\frac{1}{4})$ alors que le milieu de [BC] est N = $(\frac{\sqrt{3}}{12}, -\frac{\sqrt{6}}{12}, \frac{1}{4})$. Le vecteur $\boldsymbol{MN}$ vaut $(\frac{\sqrt{3}}{6}, -\frac{\sqrt{6}}{6}, \frac{1}{2})$ et est colinéaire au vecteur unitaire $a = (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}})$, axe de la rotation f d'angle θ demandée. Pour tout X de $\mathbb{R}^3$, on a :

$$f(X) = \cos(\theta)X + (1 - \cos(\theta))\langle X, a\rangle a + \sin(\theta)\, a \wedge X$$

ce qui permet de trouver la matrice de f dans la base canonique de $\mathbb{R}^3$:

$$M = \begin{pmatrix} \frac{1}{6} + \frac{5\cos(\theta)}{6} & -\frac{\sqrt{2}}{6} - \frac{\sqrt{2}\sin(\theta)}{2} + \frac{\sqrt{2}\cos(\theta)}{6} & \frac{\sqrt{3}}{6} - \frac{\sqrt{3}\sin(\theta)}{3} - \frac{\sqrt{3}\cos(\theta)}{6} \\ -\frac{\sqrt{2}}{6} + \frac{\sqrt{2}\sin(\theta)}{2} + \frac{\sqrt{2}\cos(\theta)}{6} & \frac{1}{3} + \frac{2\cos(\theta)}{3} & -\frac{\sqrt{6}}{6} + \frac{\sqrt{6}\cos(\theta)}{6} - \frac{\sqrt{6}\sin(\theta)}{6} \\ \frac{\sqrt{3}}{6} + \frac{\sqrt{3}\sin(\theta)}{3} - \frac{\sqrt{3}\cos(\theta)}{6} & -\frac{\sqrt{6}}{6} + \frac{\sqrt{6}\cos(\theta)}{6} + \frac{\sqrt{6}\sin(\theta)}{6} & \frac{1}{2} + \frac{\cos(\theta)}{2} \end{pmatrix}$$

Soit φ un angle élément de [0, π] tel que $\cos(\varphi) = \frac{1}{6} + \frac{5\cos(\theta)}{6}$.

Considérons la rotation $R_z = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Si (e_1, e_2, e_3) est la base canonique de $\mathbb{R}^3$, on a :

$$R_z(e_1) = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} + \frac{5\cos(\theta)}{6} \\ \sin(\varphi) \\ 0 \end{pmatrix}$$

qui a même première composante que Me_1. $\begin{pmatrix} \sin(\varphi) \\ 0 \end{pmatrix}$ a même norme euclidienne que les deux dernières composantes de Me_1 : $\begin{pmatrix} -\frac{\sqrt{2}}{6} + \frac{\sqrt{2}\sin(\theta)}{2} + \frac{\sqrt{2}\cos(\theta)}{6} \\ \frac{\sqrt{3}}{6} + \frac{\sqrt{3}\sin(\theta)}{3} - \frac{\sqrt{3}\cos(\theta)}{6} \end{pmatrix}$

Soit ψ tel que :

$$\cos(\psi)\sin(\varphi) = -\frac{\sqrt{2}}{6} + \frac{\sqrt{2}\sin(\theta)}{2} + \frac{\sqrt{2}\cos(\theta)}{6}$$

et $$\sin(\psi)\sin(\varphi) = \frac{\sqrt{3}}{6} + \frac{\sqrt{3}\sin(\theta)}{3} - \frac{\sqrt{3}\cos(\theta)}{6}$$

et soit $R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & -\sin(\psi) \\ 0 & \sin(\psi) & \cos(\psi) \end{pmatrix}$. On a $R_xR_ze_1 = Me_1$. Donc $R_z^{-1}R_x^{-1}M$ est une rotation d'axe e_1.

Les trois rotations d'axe e_1, e_3, e_1 dont le produit est M sont R_x, R_z, $(R_z^{-1}R_x^{-1}M)$.

Sol.12) Considérons le sous-espace vectoriel F orthogonal à Ker(I – A). Ker(I – A) est stable par A et, A étant orthogonale, l'orthogonal F de Ker(I – A) est aussi stable par A. Il en résulte que I – A est la matrice d'une application dont la restriction à F est bijective. Pour tout vecteur W élément de F, on a :

$$(\mathrm{I} - \mathrm{A})\,\frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}\,\mathrm{W} = \frac{1}{n+1}\,(\mathrm{I} - \mathrm{A}^{n+1})\mathrm{W}$$

or $\quad \lim\limits_{n\to\infty} \frac{1}{n+1}\,(\mathrm{I} - \mathrm{A}^{n+1})\mathrm{W} = 0 \qquad$ car $(\mathrm{A}^{n+1}\mathrm{W})$ est bornée

donc $\quad \lim\limits_{n\to\infty} (\mathrm{I} - \mathrm{A})\,\frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}\,\mathrm{W} = 0$

donc $\quad \lim\limits_{n\to\infty} \frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}\,\mathrm{W} = 0$

Si on décompose un vecteur X quelconque en la somme V + W d'un vecteur élément de Ker(I – A) et d'un élément de F, V est invariant par A donc aussi par $\frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}$, et l'on a :

$$\lim_{n\to\infty} \frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}\,\mathrm{X} = \mathrm{V}$$

Donc $\frac{\mathrm{I} + \mathrm{A} + ... + \mathrm{A}^n}{n+1}$ tend vers le projecteur orthogonal parallèlement à F.

Sol.13) Les matrices AA^{T} et $\mathrm{A}^{\mathrm{T}}\mathrm{A}$ sont symétriques donc diagonalisables avec une matrice de passage orthogonale. Il suffit donc de montrer que leur réduction donne la même matrice diagonale, i.e. elles ont mêmes valeurs propres avec même ordre de multiplicité.

❑ Soit λ une valeur propre non nulle de AA^{T} et V un vecteur propre de AA^{T} avec la valeur propre λ. Alors :

$\qquad \mathrm{AA}^{\mathrm{T}}\mathrm{V} = \lambda\mathrm{V}$

$\Rightarrow \qquad \mathrm{A}^{\mathrm{T}}\mathrm{AA}^{\mathrm{T}}\mathrm{V} = \lambda\mathrm{A}^{\mathrm{T}}\mathrm{V}$

$\Rightarrow \qquad \mathrm{A}^{\mathrm{T}}\mathrm{V}$ est vecteur propre de $\mathrm{A}^{\mathrm{T}}\mathrm{A}$ avec la valeur propre λ. $\mathrm{A}^{\mathrm{T}}\mathrm{V}$ est non nul sinon $\mathrm{AA}^{\mathrm{T}}\mathrm{V}$ le serait, donc $\lambda\mathrm{V}$ aussi, or $\lambda \neq 0$ et $\mathrm{V} \neq 0$.

En outre, si $\mathrm{V}_1, ..., \mathrm{V}_k$ sont k vecteurs propres linéairement indépendants de AA^{T} de même valeur propre λ, alors $(\mathrm{A}^{\mathrm{T}}\mathrm{V}_1, ..., \mathrm{A}^{\mathrm{T}}\mathrm{V}_k)$ est un système libre. En effet :

$\qquad \alpha_1\mathrm{A}^{\mathrm{T}}\mathrm{V}_1 + \alpha_2\mathrm{A}^{\mathrm{T}}\mathrm{V}_2 + ... + \alpha_k\mathrm{A}^{\mathrm{T}}\mathrm{V}_k = 0$

$\Rightarrow \qquad \alpha_1\mathrm{AA}^{\mathrm{T}}\mathrm{V}_1 + \alpha_2\mathrm{AA}^{\mathrm{T}}\mathrm{V}_2 + ... + \alpha_k\mathrm{AA}^{\mathrm{T}}\mathrm{V}_k = 0$

$\Rightarrow \qquad \lambda(\alpha_1\mathrm{V}_1 + \alpha_2\mathrm{V}_2 + ... + \alpha_k\mathrm{V}_k) = 0$

$\Rightarrow \qquad \alpha_1\mathrm{V}_1 + \alpha_2\mathrm{V}_2 + ... + \alpha_k\mathrm{V}_k = 0 \qquad$ car λ est non nul

$\Rightarrow \qquad \alpha_1 = ... = \alpha_k = 0 \qquad$ car $(\mathrm{V}_1, ..., \mathrm{V}_k)$ est libre

donc le sous-espace vectoriel propre associé à une valeur propre non nulle de $\mathrm{A}^{\mathrm{T}}\mathrm{A}$ a au moins la même dimension que celui de AA^{T}. Le rôle des deux matrices A et A^{T} étant symétriques, on a de même le fait que le sous-espace propre associé à une valeur propre non nulle de AA^{T} a au moins la même dimension que celui de $\mathrm{A}^{\mathrm{T}}\mathrm{A}$. Les deux sous-espaces propres ont donc même dimension.

❑ Le cas de $\lambda = 0$ se traite comme suit. Les deux sous-espaces propres associés à la valeur propre 0 ont même dimension car celle-ci est égale à n – la somme des dimensions des précédents sous-espaces propres.

On peut aussi écrire que :

$\qquad \mathrm{A}^{\mathrm{T}}\mathrm{AX} = 0$

$\Rightarrow \qquad \mathrm{X}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{AX} = 0 = (\mathrm{AX})^{\mathrm{T}}\mathrm{AX} = \| \mathrm{AX} \|^2$

$\Rightarrow \qquad \mathrm{AX} = 0$

$\Rightarrow \qquad \mathrm{A}^{\mathrm{T}}\mathrm{AX} = 0$

Donc $\mathrm{Ker}(\mathrm{A}^{\mathrm{T}}\mathrm{A}) = \mathrm{Ker}(\mathrm{A})$. Donc $\dim(\mathrm{Ker}(\mathrm{A}^{\mathrm{T}}\mathrm{A})) = n - \mathrm{rg}(\mathrm{A})$.

De même $\dim(\mathrm{Ker}(\mathrm{AA}^{\mathrm{T}})) = n - \mathrm{rg}(\mathrm{A}^{\mathrm{T}})$.

Mais rg(A) = rg(A^T). Donc Ker(A^TA) et Ker(AA^T) ont même dimension.

Sol.14) a) u étant orthogonal, on a $u^{-1} = u^*$, donc $u + u^{-1} = u + u^*$, endomorphisme symétrique puisque :

$$(u + u^*)^* = u^* + u^{**} = u^* + u = u + u^*$$

On peut aussi écrire que, pour tout x et tout y de E :

$$\langle x, (u + u^{-1})(y)\rangle = \langle x, u(y)\rangle + \langle x, u^{-1}(y)\rangle = \langle x, u(y)\rangle + \langle u(x), u(u^{-1}(y))\rangle$$
$$= \langle x, u(y)\rangle + \langle u(x), y\rangle \text{ expression symétrique en } x \text{ et } y$$

$u + u^{-1}$ est donc diagonalisable.
b) Soit x vecteur propre de $u + u^{-1}$, de valeur propre λ. On a alors $u(x) = \lambda x - u^{-1}(x)$. Soit F = Vect($x$, $u(x)$). Alors F est stable par u car $u(u(x)) = \lambda u(x) - x$.
Si x et $u(x)$ sont liés, F est une droite stable stable par u.
Si x et $u(x)$ sont indépendant, F est un plan stable par u.
c) Utiliser le fait que, u étant un endomorphisme orthogonal, si F est stable par u, $F^\perp$ est stable par u et itérer la question b) sur la restriction de u à $F^\perp$. On décompose ainsi E en une somme orthogonale de sous-espaces vectoriels stables par u, qui sont des droites ou des plans. Prendre une base orthonormée adaptée à cette décomposition. u étant orthogonale, si le sous-espace vectoriel stable est une droite, la restriction de u à cette droite est Id ou – Id (d'où les $\pm$ 1 sur la diagonale). Si le sous-espace vectoriel stable est un plan, la restriction de u à ce plan est une rotation ou une symétrie orthogonale, mais dans ce dernier cas, on aurait pu décomposer le plan en la somme orthogonale de deux droites stables.

d) $u + u^{-1}$ a pour matrice $\frac{1}{3}\begin{pmatrix} 0 & -3 & 0 & 3 \\ -3 & -4 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 3 & 0 & 1 & 4 \end{pmatrix}$ dont une base orthonormée de vecteurs propres est :

$$e_1 = \frac{1}{6}\begin{pmatrix} 3 \\ -1 \\ 1 \\ 5 \end{pmatrix} \text{ de valeur propre } 2$$

$$e_2 = \frac{1}{6}\begin{pmatrix} 3 \\ 5 \\ 1 \\ -1 \end{pmatrix} \text{ de valeur propre } -2$$

$$e_3 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 0 \\ -3 \\ 0 \end{pmatrix} \text{ et } e_4 = \frac{1}{3\sqrt{10}}\begin{pmatrix} 6 \\ -5 \\ 2 \\ -5 \end{pmatrix} \text{ de valeur propre } 0$$

On constate alors que $u(e_1) = e_1$, $u(e_2) = -e_2$, $u(e_3) = e_4$ et $u(e_4) = -e_3$. D'où la matrice de u dans la base (e_1, e_2, e_3, e_4) : $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. La restriction de u à Vect(e_3, e_4) est un quart de tour.

Sol.15) a) $f(x) = 0 \Rightarrow x$ est colinéaire à a, mais $f(a) = (1 + \alpha\| a \|^2)a = (1 + \alpha)a$. Donc :
si $\alpha = -1$, a appartient à Ker(f) et f n'est pas bijective,
si $\alpha \neq -1$, $f(a)$ est non nul donc le seul vecteur colinéaire à a qui a une image nulle par f est le vecteur nul, donc f est injective. Donc f est bijective puisqu'un endomorphisme injectif dans un

espace vectoriel de dimension finie est bijectif (grâce au théorème du rang). Dans ce cas, si $f(x) = y$, alors $x + \alpha<x \mid a>a = y$ et si on fait le produit scalaire par a, on obtient :

$$<x \mid a>(1 + \alpha) = <y \mid a>$$

donc $<x \mid a> = \dfrac{<y \mid a>}{1 + \alpha}$.

En reportant la valeur de $<x \mid a>$ dans l'égalité $x + \alpha<x \mid a>a = y$, on obtient :

$$x = y - \frac{\alpha}{1 + \alpha} <y \mid a>a = f^{-1}(y)$$

b) $<f(x) \mid y> = <x \mid y> + \alpha<x \mid a><y \mid a>$ symétrique en x et y

$= <x \mid f(y)>$

c) f est diagonalisable puisque symétrique. On a aussi :
λ est valeur propre s'il existe x non nul tel que $\lambda x = x + \alpha<x \mid a>a$. Donc

ou bien $\lambda = 1$ et le sous-espace propre E_1 est l'hyperplan orthogonal à a

ou bien $\lambda \neq 1$ et le sous-espace propre E_λ est nécessairement la droite engendrée par a, car x est colinéaire à a. Mais dans ce cas, on a $\lambda a = a + \alpha a$ donc $\lambda = 1 + \alpha$. On retrouve la valeur propre nulle si $\alpha = -1$.

d) La condition nécessaire et suffisante est $1 + \alpha = -1$, soit $\alpha = -2$. f est la réflexion par rapport à l'hyperplan orthogonal à a.

Sol.16) a) u étant symétrique est diagonalisable dans une base orthonormée. Dans une telle base, la matrice D de u est diagonale. Pour toute valeur propre λ de u de vecteur propre x, on a :

$$0 \leq <u(x), x> = \lambda \parallel x \parallel^2$$

donc $\lambda \geq 0$. Il suffit de prendre v de matrice diagonale $\sqrt{D}$ dans la base précédente, où $\sqrt{D}$ est la matrice ayant des $\sqrt{\lambda}$ sur la diagonale.

Si w est telle que $w^2 = u$, avec w symétrique positif, alors $wu = w^3 = uw$ donc w commute avec u et donc avec $u - \lambda\mathrm{Id}$ pour toute valeur propre λ de u. Donc tout sous-espace propre $\mathrm{Ker}(u - \lambda\mathrm{Id})$ est stable par w. La restriction de w à ce sous-espace vectoriel est encore symétrique positif, donc diagonalisable à valeurs propres positives ou nulles. Comme $w^2 = u$, les valeurs propres de cette restriction ne peuvent être que $\sqrt{\lambda}$, donc la restriction de w à ce sous-espace vectoriel est $\sqrt{\lambda}\,\mathrm{Id}$. w est donc unique sur tout sous-espace propre de u et donc unique sur E puisque la somme des sous-espaces propres de u est égal à E.

b) $f^* \circ f$ est un endomorphisme symétrique positif. En effet :

$$(f^* \circ f)^* = f^* \circ f^{**} = f^* \circ f$$

et $\forall x, <f^* \circ f(x), x> = <f(x), f(x)> \geq 0$

donc, d'après le a), $f^* \circ f$ est de la forme h^2 avec h symétrique positif. En outre, f étant inversible, il en est de même de h. Posons $g = f \circ h^{-1}$. Alors :

$$g^* \circ g = (h^{-1})^* \circ f^* \circ f \circ h^{-1} = (h^*)^{-1} \circ f^* \circ f \circ h^{-1} = h^{-1} \circ h^2 \circ h^{-1} = \mathrm{Id}$$

donc g est un automorphisme orthogonal.

Si $f = gh = g'h'$ avec g et g' orthogonaux et h et h' positifs, alors :

$$h^2 = h^* \circ g^* \circ g \circ h = f^* \circ f = h'^* \circ g'^* \circ g' \circ h' = h'^2$$

donc $h = h'$ par l'unicité du a), et donc $g = g'$, puisque h et h' sont inversibles.

c) $$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix} \times \frac{1}{5}\begin{pmatrix} 7 & -1 \\ -1 & 18 \end{pmatrix}$$

d) Montrons d'abord que $\mathrm{Ker}(f^* \circ f) = \mathrm{Ker}(f)$. L'inclusion $\mathrm{Ker}(f) \subset \mathrm{Ker}(f^* \circ f)$ est triviale. Réciproquement :

$$x \in \mathrm{Ker}(f^* \circ f)$$

$\Rightarrow \quad f^* \circ f(x) = 0$
$\Rightarrow \ <x, f^* \circ f(x)> = 0$
$\Rightarrow \quad <f(x), f(x)> = 0$
$\Rightarrow \quad \| f(x) \| = 0$
$\Rightarrow \quad f(x) = 0$
$\Rightarrow \quad x \in \mathrm{Ker}(f)$,

(On a déjà effectué ce raisonnement dans un exercice précédent). On a donc également $\mathrm{Ker}(f^* \circ f)^\perp = \mathrm{Ker}(f)^\perp$.

Montrons qu'on a également $\mathrm{Im}(f^* \circ f) = \mathrm{Im}(f^*) = \mathrm{Ker}(f)^\perp$. L'inclusion $\mathrm{Im}(f^* \circ f) \subset \mathrm{Im}(f^*)$ est évidente, et le théorème du rang donne :

$$\begin{aligned} \dim(\mathrm{Im}(f^* \circ f)) &= \dim(\mathrm{E}) - \dim(\mathrm{Ker}(f^* \circ f)) \\ &= \dim(\mathrm{E}) - \dim(\mathrm{Ker}(f)) \qquad \text{d'après l'égalité } \mathrm{Ker}(f^* \circ f) = \mathrm{Ker}(f) \\ &= \dim(\mathrm{Im}(f)) \end{aligned}$$

$\mathrm{Im}(f^* \circ f)$ et $\mathrm{Im}(f)$ ayant même dimension et l'un étant inclus dans l'autre, ils sont égaux.

Ensuite, $\mathrm{Im}(f^*) \subset \mathrm{Ker}(f)^\perp$ car :

$\quad y \in \mathrm{Im}(f^*)$
$\Rightarrow \quad \exists\, x \in \mathrm{E},\ y = f^*(x)$
$\Rightarrow \quad \forall\, u \in \mathrm{Ker}(f),\ <y, u> = <f^*(x), u> = <x, f(u)> = <x, 0> = 0$
$\Rightarrow \quad y \in \mathrm{Ker}(f)^\perp$

et $\dim(\mathrm{Im}(f^*)) = \mathrm{rg}(f^*) = \mathrm{rg}(f) = \dim(\mathrm{E}) - \dim(\mathrm{Ker}(f)) = \dim(\mathrm{Ker}(f)^\perp)$. $\mathrm{Im}(f^*)$ et $\mathrm{Ker}(f)^\perp$ ont même dimension et l'un est inclus dans l'autre, donc ils sont égaux.

Or $f^* \circ f$ est une application injective de $\mathrm{Ker}(f^* \circ f)^\perp$ sur $\mathrm{Im}(f^* \circ f)$, donc bijective car ces deux espaces ont même dimension. Puisque $\mathrm{Ker}(f^* \circ f)^\perp = \mathrm{Ker}(f)^\perp = \mathrm{Im}(f^* \circ f)$, $f^* \circ f$ est bijective de $\mathrm{Ker}(f)^\perp$ sur $\mathrm{Ker}(f)^\perp$ et la restriction de $f^* \circ f$ à ce sous-espace vectoriel est symétrique et positif. Il existe donc h endomorphisme symétrique positif bijectif de $\mathrm{Ker}(f)^\perp$ tel que $f^* \circ f = h^2$ sur $\mathrm{Ker}(f)^\perp$.

Considérons une base orthonormée $(e_1, ..., e_p)$ de $\mathrm{Ker}(f)^\perp$, et soit g de $\mathrm{Ker}(f)^\perp$ dans $\mathrm{Im}(f)$ définie par $g(e_i) = f(h^{-1}(e_i))$. g conserve le produit scalaire. En effet, pour tout (i, j), on a :

$$\begin{aligned} <g(e_i), g(e_j)> &= <f(h^{-1}(e_i)), f(h^{-1}(e_j))> = <h^{-1}(e_i), f^* \circ f(h^{-1}(e_j))> = <h^{-1}(e_i), h^2(h^{-1}(e_j))> \\ &= <h^{-1}(e_i), h(e_j)> = <h^*(h^{-1}(e_i), e_j> = <h(h^{-1}(e_i), e_j> = <e_i, e_j> = \delta_{ij} \end{aligned}$$

Comme $\dim(\mathrm{Ker}(f)^\perp) = \dim(\mathrm{Im}(f))$, g est une bijection entre ces deux espaces, qui envoie la base orthonormée $(e_1, ..., e_p)$ de $\mathrm{Ker}(f)^\perp$ sur une base orthonormée de $\mathrm{Im}(f)$.

On étend h à E on posant que, sur $\mathrm{Ker}(f)$, $h = 0$, et on étend g à E en définissant linéairement g de façon que g envoie une base orthonormée de $\mathrm{Ker}(f)$ sur une base orthonormée de $\mathrm{Im}(f)^\perp$. g sera alors une isométrie de E, et on aura $f = g \circ h$ pour tout x de E. Il n'y a pas unicité de la décomposition puisqu'on a une liberté sur l'extension de g.

e) $\mathrm{Ker}(f)$ est engendré par $e_2 = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}$, et $\mathrm{Ker}(f)^\perp$ est engendré par $e_1 = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. On vérifiera que :

$$(f^* \circ f)(e_1) = 10e_1$$

On prendra donc $h(e_1) = \sqrt{10}\ e_1$ et $h(e_2) = 0$. Dans la base canonique de $\mathbf{R}^2$, h a pour matrice $\dfrac{\sqrt{10}}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.

Puis $\mathrm{Im}(f)$ est engendré par $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $\mathrm{Im}(f)^\perp$ est engendré par $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

On définit ensuite $g(e_1) = f(h^{-1}(e_1)) = f(\frac{1}{\sqrt{10}}\, e_1) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $g(e_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (on pourrait prendre $g(e_2)$ de signe contraire). Dans la base canonique, la matrice de g est $\frac{1}{\sqrt{10}}\begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix}$, et l'on a :

$$\begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix} \times \frac{\sqrt{10}}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Sol.17) C'est équivalent à $\det(A^{\top}A) \leq (\frac{\mathrm{Tr}(A^{\top}A)}{n})^n$. $A^{\top}A$ est une matrice symétrique à valeurs propres positives ou nulles. Il suffit donc de montrer que $\det(S) \leq (\frac{\mathrm{Tr}(S)}{n})^n$ pour toute matrice symétrique S à valeurs propres positives ou nulles, ou bien, puisque toute matrice symétrique est diagonalisable, à $\det(D) \leq (\frac{\mathrm{Tr}(D)}{n})^n$ pour toute matrice diagonale D à coefficients positifs ou nuls. On tombe sur l'inégalité de convexité bien connue : $\sqrt[n]{\lambda_1 ... \lambda_n} \leq \frac{\lambda_1 + ... + \lambda_n}{n}$ (voir, dans le chapitre L1/DERIVEE.PDF, la partie traitant des fonctions convexes).

◆