Espaces euclidiens

Benoit Loisel

Version du 10 mars 2025

Table des matières

Introduction

Dans les trois premiers chapitre de ce cours, le corps de base sera $\mathbb{K} = \mathbb{R}$ et on ne considérera que des $\mathbb{R}$-espaces vectoriels, qu'on appellera simplement « espaces vectoriels » sans en rappeler le corps de base.

Dans ce cours, on va s'intéresser à des espaces vectoriels E munis d'une information supplémentaire de structure appelée « produit scalaire ». Cela signifie qu'on va définir une application sur l'espace vectoriel E qui a **deux vecteurs** $u, v \in E$ associe un **nombre**, avec de bonnes propriétés de compatibilité (associativité, commutativité, distributivité sur l'addition $+$ de vecteurs, compatibilité à la multiplication scalaire $\cdot$). C'est un peu comme si on définissait une « multiplication » entre les vecteurs, mais que cette multiplication renvoyait un nombre au lieu d'un vecteur. C'est ce qui motive la terminologie de « produit scalaire ». Dans le chapitre 3, nous verrons que dans le cas spécifique de la dimension 3, il est possible de définir d'autres opérations qui s'apparentent également à des produits entre les vecteurs. On parlera de « produit vectoriel » et de « produit mixte ».

Quel est l'intérêt de tout ça ? Les espaces euclidiens apparaissent naturellement dans de très nombreux contextes de mathématiques et de physique. À titre d'exemple, mais c'est une liste largement non exhaustive :

— En mécanique classique, c'est la géométrie naturelle dans laquelle on manipule les vecteurs, par exemple pour calculer le travail d'une force lors d'un déplacement.

— En géométrie, c'est le cadre classique historique, dit « Euclidien » inventé par Euclide. C'est dans ce cadre qu'on calcule les distances, les longueurs, les aires, les volumes. De plus, le cadre euclidien permet, comme on le verra au chapitre 3 de définir convenablement une notion d'angle et d'orientation du plan et de l'espace de dimension 3.

— En optique, l'étude des rayons lumineux passe par la compréhension de phénomène de réflexion, de réfraction. Cela fait intervenir l'angle entre le rayon et le plan (ou la surface) rencontrée par le rayon. Des techniques vectorielles d'espaces euclidien permettent de calcul l'angle d'incidence.

— En géométrie toujours, cela permet de définir une notion d'angle droit, d'orthogonalité. Cela va notamment permettre de calculer la distance entre un point et une droite, entre deux droites, entre un point et un plan, etc.

— En mécanique quantique, on va s'intéresser à des opérateurs sur les particules qu'on verra comme des endomorphismes d'un espace euclidien (ou hermitien, ou hilbertien, qui en sont des généralisations naturelles).

— En mécanique du solide, on aura besoin de comprendre l'évolution dans l'espace de notre solide. On aura donc besoin de comprendre les déplacements du solide dans l'espace : elles en préservent les longueurs et les angles si le solide est indéformable.

— En probabilités et statistiques, le caractère quadratique de la variance fait naturellement apparaître des espaces euclidiens, notamment quand on a plusieurs variables aléatoires non indépendantes et qu'on calcule des matrices de covariance par exemple.

— En astronomie, l'étude des coniques apparaît naturel pour comprendre le problème à deux corps, typiquement le mouvement d'une planète tournant autour de son étoile, ou d'un satellite tournant autour de sa planète.

— En construction de bâtiments, l'usage de surfaces réglées facilite l'ingénierie de la construction. C'est sur de telles techniques que sont construites les hautes cheminées d'évacuation de vapeur d'eau des centrales nucléaires par exemple.

— En analyse mathématiques, notamment en théorie des opérateurs, la généralisation d'espaces euclidiens à la dimension infinie (espaces de Hilbert, etc.) va permettre de mieux comprendre la structure d'espaces de fonction et donner des outils supplémentaires dans la résolution d'équations différentielles (EDO) et d'équations aux dérivées partielles (EDP).

— En analyse de Fourier, la dimension infinie est le cadre naturel pour comprendre les espaces de fonctions périodiques et effectuer des « transformations de Fourier ». Cela va permettre de transformer des fonctions périodiques en la donnée numérique d'une suite de valeurs numériques avec une perte d'information minime.

— En théorie du signal, et plus généralement dans l'étude des ondes (mécaniques, sonores, électromagnétiques, etc.) cela permet donc de comprendre et manipuler des ondes qui se mélangent, se superposent, s'annulent, entrent en résonance, etc.

— Plus généralement, en électromagnétique, de nombreuses fonctions issues de la physique ont un caractère périodique et pourront être étudiées plus facilement dans le langage de Fourier qui s'appuie sur la construction d'un produit scalaire sur l'espace vectoriel des fonctions périodiques.

— En analyse d'image et compression d'image, on peut aller chercher l'information principale par de l'analyse de Fourier.

— En relativité restreinte, l'espace de Minkowski (non euclidien) considéré est relié à la forme quadratique $q(x, y, z, t) = x^2 + y^2 + z^2 - t^2$.

La maîtrise des techniques d'espaces euclidien s'avère donc essentielle dans une démarche scientifique en mathématiques ou en physique.

Structure du cours Ce cours sera divisé en 5 chapitres. Dans un premier chapitre, nous présenterons la structure générale des espaces vectoriels réels muni d'un produit scalaire. Lorsque cela sera nécessaire, nous nous placeront en dimension finie, mais ça n'est pas systématiquement indispensable pour tous les résultats.

Du chapitre 2 au chapitre 4, nous nous restreindrons systématiquement au cas de la dimension finie, bien que la plupart des résultats puissent avoir un intérêt en dimension infinie, à condition d'être très prudent dans leur adaptation.

Dans le deuxième chapitre, nous nous intéresserons aux transformations vectorielles (endomorphismes) d'un espace euclidien.

Dans le troisième chapitre, nous verrons les spécificités de la dimension 2 ou 3. Cette étude nous amènera à définir une notion d'orientation (en toute dimension finie, pas seulement 2 ou 3). Nous disposerons d'outils supplémentaires spécifiques à la dimension 3 qui s'avèrent très utiles en mécanique du solide par exemple.

Dans le quatrième chapitre, nous nous intéresserons plus généralement à l'étude des formes bilinéaires et quadratiques. D'une part, elles enrichiront notre compréhension d'espaces munis d'une structure, pas nécessairement euclidienne et, d'autre part, elles seront le point de départ à l'étude des coniques du plan euclidien et quadriques de l'espace euclidien de dimension 3.

Enfin, si le temps le permet, dans le cinquième chapitre, nous reprendrons notre étude dans le cadre d'espaces vectoriels complexes. Les espaces vectoriels complexes de dimension finie seront alors appelés des espaces hermitiens. C'est, le plus souvent, le bon cadre pour l'analyse mathématique.

Chapitre I

Espaces vectoriels euclidiens

Le but de ce chapitre est de définir des notions de norme, distance et orthogonalité sur des espaces vectoriels. Ceux-ci peuvent être des systèmes de coordonnées $\mathbb{R}^n$, mais également des espaces vectoriels de matrices (e.g. $\mathcal{M}_n(\mathbb{R})$), des espaces vectoriels de fonctions (e.g. $\mathcal{C}^0(\mathbb{R})$, fonctions périodiques, etc.) ou encore des espaces vectoriels de polynômes (e.g. $\mathbb{R}[X]_{<d}$ polynômes de degré strictement inférieur à d).

I.1 Formes bilinéaires

I.1.a Définition

Dans cette partie, on ne suppose pas que E est de dimension finie.

Définition I.1.1. Soit E un espace vectoriel. Une *forme bilinéaire* sur E est une application

$$f : \begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (x,y) & \mapsto & f(x,y) \end{array}$$

vérifiant les deux propriétés :

1. (**linéarité à droite**) Pour tout $x \in E$, l'application $f(x,\cdot) : y \mapsto f(x,y)$ est linéaire, i.e.

$$\forall x \in E,\ \forall u,v \in E,\ \forall \lambda,\mu \in \mathbb{R},\ f(x,\lambda u+\mu v) = \lambda f(x,u) + \mu f(x,v)$$

2. (**linéarité à gauche**) Pour tout $y \in E$, l'application $f(\cdot,y) : x \mapsto f(x,y)$ est linéaire, i.e.

$$\forall y \in E,\ \forall u,v \in E,\ \forall \lambda,\mu \in \mathbb{R},\ f(\lambda u+\mu v,y) = \lambda f(u,y) + \mu f(v,y)$$

Les applications $f(x,\cdot)$ et $f(\cdot,y)$ sont des formes linéaires, c'est-à-dire des applications linéaires à valeurs dans le corps $\mathbb{R}$. L'application f est à valeurs dans le corps $\mathbb{R}$, d'où le terme « forme » et présente simultanément deux linéarités, d'où le terme « bilinéaire ».

Exercice 1. Montrer qu'une forme f est bilinéaire si, et seulement si,

$$\begin{aligned} &\forall x_1,x_2,y_1,y_2 \in E,\ \forall \lambda_1,\lambda_2,\mu_1,\mu_2, \\ &\qquad f(\lambda_1 x_1+\lambda_2 x_2, \mu_1 y_1+\mu_2 y_2) = \lambda_1\mu_1 f(x_1,y_1) + \lambda_1\mu_2 f(x_1,y_2) + \lambda_2\mu_1 f(x_2,y_1) + \lambda_2\mu_2 f(x_2,y_2) \end{aligned}$$

Notation I.1.2. On note $\mathcal{B}(E)$ (ou parfois $\mathrm{Bil}(E)$) l'ensemble des formes bilinéaires sur E.

Fait I.1.3. *L'ensemble des formes bilinéaires est un sous-espace vectoriel de l'espace des applications $E\times E \to \mathbb{R}$, c'est-à-dire que si $f,g \in \mathcal{B}(E)$, si $\lambda,\mu \in \mathbb{R}$, alors l'application*

$$h : \begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (x,y) & \mapsto & \lambda f(x,y) + \mu g(x,y) \end{array}$$

notée $h = \lambda f + \mu g$ est une application bilinéaire.

Exercice 2. Démontrer ce fait.

Définition I.1.4. On dit qu'une forme bilinéaire f est :

— *symétrique* si

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

— *anti-symétrique* si

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

Sinon, on dit seulement que c'est une forme bilinéaire.

Remarque I.1.5. L'ensemble des formes bilinéaires symétriques (resp. anti-symétriques) forme un sous-espace vectoriel de $\mathcal{B}(E)$.

Exemple I.1.6. 1. Soit $n \in \mathbb{N}^*$ et $E = \mathbb{R}^n$. Soit $1 \leqslant p \leqslant n$. Alors

$$\begin{array}{lccc} f: & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & \big((x_1,\dots,x_n),(y_1,\dots,y_n)\big) & \mapsto & \displaystyle\sum_{k=1}^{p} x_k y_k = x_1y_1 + \cdots + x_py_p \end{array}$$

est une forme bilinéaire symétrique.

2. L'application

$$\begin{array}{lccc} f: & \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ & \big((x,u),(y,v)\big) & \mapsto & xv - yu \end{array}$$

est une forme bilinéaire anti-symétrique.

3. Soit $E = \mathcal{C}^0([0,1])$ l'espace vectoriel des fonctions continues sur l'intervalle $[0,1]$. L'application

$$\begin{array}{lccc} f: & \mathcal{C}^0([0,1]) \times \mathcal{C}^0([0,1]) & \to & \mathbb{R} \\ & (u,v) & \mapsto & \displaystyle\int_0^1 u(t)v(t)dt \end{array}$$

est une forme bilinéaire symétrique.

4. Soit $E = \mathcal{M}_n(\mathbb{R})$ l'espace vectoriel des matrices carrées de taille n. Alors l'application

$$\begin{array}{lccc} f: & \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) & \to & \mathbb{R} \\ & (A,B) & \mapsto & \mathrm{Tr}({}^tAB) \end{array}$$

est une forme bilinéaire symétrique.

5. Soit E un espace vectoriel. L'application

$$\begin{array}{lccc} f: & E \times E & \to & E \\ & (x,y) & \mapsto & x+y \end{array}$$

n'est pas une forme bilinéaire car elle n'est pas à valeurs dans $\mathbb{R}$. De plus, si $x \neq 0$, l'application $f(x,\cdot)$ n'est pas linéaire car $f(x,0) = x \neq 0$. De même, l'application $f(\cdot,y)$ n'est pas linéaire.

6. Soit E un espace vectoriel et φ, ψ deux formes linéaires sur E. Alors l'application

$$\begin{array}{lccc} f: & E \times E & \to & \mathbb{R} \\ & (x,y) & \mapsto & \varphi(x)\psi(y) \end{array}$$

est bilinéaire. En général, elle n'est ni symétrique, ni anti-symétrique.

Exercice 3. Démontrer les affirmations des exemples précédents.

I.1.b Interprétation matricielle

On suppose dans cette sous-partie que E est de dimension finie n.

Définition I.1.7. Soit $f \in \mathcal{B}(E)$ une forme bilinéaire sur E.
Si $\mathcal{E} = (e_1,\dots,e_n)$ est une base de E, on appelle matrice de f dans la base $\mathcal{B}$ la matrice

$$\mathrm{Mat}(f,\mathcal{E}) := \begin{pmatrix} f(e_1,e_1) & \cdots & f(e_1,e_n) \\ \vdots & \ddots & \vdots \\ f(e_n,e_1) & \cdots & f(e_n,e_n) \end{pmatrix}$$

c'est-à-dire la matrice $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ dont le coefficient à la i-ème ligne et j-ème colonne est $a_{i,j} = f(e_i,e_j)$.

Remarque I.1.8. Soient $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n$ deux vecteurs de E et A la matrice de f dans la base $\mathcal{E}$.

D'une part, par bilinéarité de f on a alors

$$\begin{aligned}
f(x,y) &= f\left(\sum_{i=1}^n x_ie_i, \sum_{j=1}^n y_je_j\right) \\
&= \sum_{i=1}^n x_i f\left(e_i, \sum_{j=1}^n y_je_j\right) && \text{par linéarité à gauche} \\
&= \sum_{i=1}^n \sum_{j=1}^n x_iy_j f(e_i, e_j) && \text{par linéarité à droite}
\end{aligned}$$

D'autre part, notons $X = \mathrm{Mat}_{\mathcal{E}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \mathrm{Mat}_{\mathcal{E}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ les matrices colonnes associées aux vecteurs x et y dans la base $\mathcal{E}$. On observe alors que

$$\begin{aligned}
{}^tXAY &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\
&= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \sum_{j=1}^n a_{1,j}y_j \\ \vdots \\ \sum_{j=1}^n a_{n,j}y_j \end{pmatrix} \\
&= \sum_{i=1}^n x_ia_{i,j}y_j
\end{aligned}$$

Comme $a_{i,j} = f(e_i, e_j)$, on constate alors que

$$f(x,y) = {}^tXAY$$

ce qui donne un moyen matriciel de calculer les valeurs de la forme linéaire en écriture dans une base.

Inversement, donnons nous une matrice $A \in \mathcal{M}_n(\mathbb{R})$ et une base $\mathcal{E}$ d'un espace vectoriel E de dimension n. On peut alors définir une application $f : E \times E \to \mathbb{R}$ par

$$(x,y) = \left(\sum_{i=1}^n x_ie_i, \sum_{j=1}^n y_je_j\right) \mapsto \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Exercice 4. Montrer que f est une forme bilinéaire.

Ainsi, dès qu'on se fixe une base $\mathcal{E}$, on a établi une correspondance entre les matrices carrées et les formes bilinéaires. Ce qu'on peut résumer dans la proposition suivante :

Proposition I.1.9. *Soit E un espace vectoriel de dimension finie n. Soit $\mathcal{E} = (e_1, \ldots, e_n)$ une base de E. Alors l'application*

$$\begin{array}{rccc} \Phi : & \mathcal{B}(E) & \to & \mathcal{M}_n(\mathbb{R}) \\ & f & \mapsto & \mathrm{Mat}(f, \mathcal{E}) \end{array}$$

est une bijection, d'application réciproque

$$\begin{array}{rccc} \Phi^{-1} : & \mathcal{M}_n(\mathbb{R}) & \to & \mathcal{B}(E) \\ & A & \mapsto & \left(\begin{array}{ccc} E \times E & \to & \mathbb{R} \\ (x,y) & \mapsto & {}^t\mathrm{Mat}_{\mathcal{E}}(x) A \,\mathrm{Mat}_{\mathcal{E}}(y) \end{array}\right) \end{array}$$

Remarque I.1.10. On a mieux. L'application Φ est linéaire entre les espaces vectoriels $\mathcal{B}(E)$ et $\mathcal{M}_n(\mathbb{R})$. C'est donc un isomorphisme d'espaces vectoriels.

Ainsi, l'espace vectoriel des formes bilinéaires $\mathcal{B}(E)$ est de dimension n^2 lorsque E est de dimension n.

Corollaire I.1.11. *Lorsqu'elles sont écrites en coordonnées, on sait reconnaitre les formes bilinéaires par lecture graphique : elles sont de la forme*

$$f\big((x_1,\dots,x_n),(y_1,\dots,y_n)\big)=\sum_{1\leqslant i,j\leqslant n} a_{i,j}x_iy_j$$

En particulier, si l'un des termes d'une somme n'est pas le produit d'exactement un facteur en x_i et un en y_j (par exemple, il y en a trop ou pas assez), alors ça ne peut pas être une forme bilinéaire.

Proposition I.1.12. *Soit E un espace vectoriel de dimension finie. Soit $\mathcal{E}$ une base de E. Soit $f\in\mathcal{B}(E)$ une forme bilinéaire sur E. Alors f est symétrique si, et seulement si, sa matrice $A=\operatorname{Mat}(f,\mathcal{E})$ est symétrique.*

Remarque I.1.13. Cette propriété ne dépend pas de la base $\mathcal{E}$ choisie dans la proposition. C'est vrai dans n'importe quelle base $\mathcal{E}$ de E.

Démonstration. $\Rightarrow$ Supposons que f est symétrique. Alors pour tous x,y, on a $f(x,y)=f(y,x)$. En particulier, si $x=e_i$ et $y=e_j$, on a $a_{i,j}=f(e_i,e_j)=f(e_j,e_i)=a_{j,i}$. Ceci étant valable pour tout (i,j), on en déduit que $A={}^tA$, c'est-à-dire que A est symétrique.

$\Leftarrow$ Réciproquement, supposons que A est symétrique. Soient $x,y\in E$ deux vecteurs. Notons X,Y les matrices colonnes associées. Alors $f(x,y)={}^tXAY$ est un nombre qu'on peut voir comme une matrice de taille 1×1. Donc $f(x,y)={}^tf(x,y)={}^t({}^tXAY)={}^tY{}^tA{}^{tt}X={}^tYAX$ car ${}^tA=A$ et ${}^{tt}X=X$. Ainsi $f(x,y)=f(y,x)$. Ceci étant vrai pour tous $x,y\in E$, on en déduit que f est symétrique. □

Exercice 5. Montrer que f est anti-symétrique si et seulement si sa matrice $A=\operatorname{Mat}(f,\mathcal{E})$ est anti-symétrique, c'est-à-dire que ${}^tA=-A$.

Correction de l'exercice 5. $\Rightarrow$ Supposons que f est anti-symétrique. Alors pour tous x,y, on a $f(x,y)=-f(y,x)$. En particulier, si $x=e_i$ et $y=e_j$, on a $a_{i,j}=f(e_i,e_j)=f(e_j,e_i)=-a_{j,i}$. Ceci étant valable pour tout (i,j), on en déduit que ${}^tA=-A$, c'est-à-dire que A est anti-symétrique.

$\Leftarrow$ Réciproquement, supposons que A est anti-symétrique. Soient $x,y\in E$ deux vecteurs. Notons X,Y les matrices colonnes associées. Alors $f(x,y)={}^tXAY$ est un nombre qu'on peut voir comme une matrice de taille 1×1. Donc $f(x,y)={}^tf(x,y)={}^t({}^tXAY)={}^tY{}^tA{}^{tt}X=-{}^tYAX$ car ${}^tA=-A$ et ${}^{tt}X=X$. Ainsi $f(x,y)=-f(y,x)$. Ceci étant vrai pour tous $x,y\in E$, on en déduit que f est anti-symétrique. □

Corollaire I.1.14. *Soit E un espace vectoriel de dimension finie n. Alors*

1. *l'ensemble des formes bilinéaires symétriques est un espace vectoriel de dimension $\frac{n(n+1)}{2}$;*
2. *l'ensemble des formes bilinéaires anti-symétriques est un espace vectoriel de dimension $\frac{n(n-1)}{2}$.*

I.1.c Changement de base

Proposition I.1.15. *Soit E est un espace vectoriel de dimension finie n. Soient $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E et P la matrice de passage de $\mathcal{E}$ vers $\mathcal{E}'$.*

Si f est une forme bilinéaire sur E, alors

$$\operatorname{Mat}(f,\mathcal{E}')={}^tP\operatorname{Mat}(f,\mathcal{E})P$$

Remarque I.1.16. Si u est un endomorphisme d'un espace vectoriel de dimension finie n, on rappelle que $\operatorname{Mat}_{\mathcal{E}'}(u)=P^{-1}\operatorname{Mat}_{\mathcal{E}}(u)P$. On voit donc que l'opération de changement de base n'a pas le même effet pour les matrices des endomorphismes et pour les matrices des formes bilinéaires.

Dans l'écriture $A'=P^{-1}AP$, on dit que les matrices A et A' sont *semblables*.

Dans l'écriture $A'={}^tPAP$, on dit que les matrices A et A' sont *congruentes*.

Démonstration. Pour alléger les notations, notons $A=\operatorname{Mat}(f,\mathcal{E})$ et $A'=\operatorname{Mat}(f,\mathcal{E}')$.

Considérons la matrice de passage P de la base $\mathcal{E}$ à la base $\mathcal{E}'$ (on a écrit en colonne les éléments de $\mathcal{E}'$ dans la base $\mathcal{E}$).

Soit $x,y\in E$ deux vecteurs. On note X,Y (resp. X',Y') leurs matrices colonnes dans la base $\mathcal{E}$ (resp. $\mathcal{E}'$). Rappelons que par définition de la matrice de passage, on a alors $X=PX'$ et $Y=PY'$.

Dans la base $\mathcal{E}$, on a alors $f(x,y)={}^tXAY$.

Dans la base $\mathcal{E}$, on a alors $f(x,y)={}^tX'A'Y'$.

Mais on peut alors écrire

$${}^tX'A'Y'={}^tXAY={}^t(PX')A(PY')={}^tX'{}^tPAPY'$$

Ceci est valable pour n'importe quels $x, y \in E$, donc n'importe quelles matrices colonnes X', Y'. C'est en particulier vrai pour $X' = E_i$ et $Y' = E_j$ les matrices colonnes avec des 0 partout sauf à la ligne i (resp. j) où il y a un 1. Le (i, j)-ème coefficient de A' est alors

$$A'_{i,j} = {}^tE_i A' E_j = {}^tE_i {}^tPAPE_j = ({}^tPAP)_{i,j}$$

Ainsi $A' = {}^tPAP$. □

I.2 Produits scalaires

Dans $\mathbb{R}^3$, on connait déjà le produit scalaire $(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = x_1x_2 + y_1y_2 + z_1z_2$. Le but de cette section est de définir une opération similaire pour n'importe quel couple de vecteurs d'un espace vectoriel réel quelconque (dimension finie ou infinie), de sorte qu'on retrouve certaines propriétés de normes, d'angles et de distances très utiles.

I.2.a Produits scalaires et espaces euclidiens

Définition I.2.1. Soit E un espace vectoriel et $f \in \mathcal{B}(E)$ une forme bilinéaire sur E. On dit que

1. f est *positive* si $\forall x \in E,\ f(x, x) \geqslant 0$;
2. f est *négative* si $\forall x \in E,\ f(x, x) \leqslant 0$;
3. f est *définie positive* si $\forall x \in E \setminus \{0\},\ f(x, x) > 0$;
4. f est *définie négative* si $\forall x \in E \setminus \{0\},\ f(x, x) < 0$.

Exercice 6. Montrer que f est définie positive si, et seulement si, elle est positive et $f(x, x) = 0 \Leftrightarrow x = 0$.

Exercice 7. Dans l'exemple I.1.6, quelles sont les formes positives ? négatives ? définies positives ?

Définition I.2.2. Soit E un espace vectoriel.

On dit qu'une application $\varphi : E \times E \to \mathbb{R}$ est un *produit scalaire* si c'est une forme bilinéaire symétrique et définie positive.

On appelle *espace euclidien* la donnée d'un espace vectoriel de dimension finie munie E et d'un produit scalaire φ sur E. On dit alors que E est *muni* du produit scalaire φ.

Remarque I.2.3. Étant donné un espace vectoriel E, il est toujours possible de définir différents produits scalaires sur E. On en verra des exemples en exercice. Cependant, lorsqu'il n'y a pas d'ambiguïté sur le produit scalaire considéré, il sera parfois commode de noter $\langle x, y\rangle$ au lieu de $\varphi(x, y)$.

Parfois, les produits scalaires sont aussi notées $(x, y), (x|y), \langle x|y\rangle$

Remarque I.2.4. Si F est un sous-espace vectoriel d'un espace euclidien (E, φ), on peut définir une application

$$\begin{array}{rccc} \varphi_{|F\times F} : & F \times F & \to & \mathbb{R} \\ & (x, y) & \mapsto & \langle x, y\rangle \end{array}$$

par restriction de φ à $F \times F$. Alors $\varphi_{|F\times F}$ est bilinéaire, symétrique, définie positive donc est un produit scalaire sur F, ce qui fait de $(F, \varphi_{|F\times F})$ un espace euclidien.

I.2.b Exemples de produits scalaires

Exemple I.2.5.

1. Soit $E = \mathbb{R}^n$. L'application donnée par $\varphi\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = x_1y_1 + \cdots + x_ny_n$ est un produit scalaire sur E. On l'appelle le *produit scalaire canonique* sur $\mathbb{R}^n$.

   Si $\mathcal{E} = (e_1, \ldots, e_n)$ est la base canonique de $\mathbb{R}^n$, alors φ s'écrit matriciellement

   $$\varphi(x, y) = {}^tXY$$

   Ainsi, on a $\mathrm{Mat}(\varphi, \mathcal{E}) = I_n$. Autrement dit, le produit scalaire canonique de $\mathbb{R}^n$ est l'application bilinéaire dont la matrice dans la base canonique est l'identité.

2. Soit $E = \mathbb{R}^n$ et $1 \leqslant p < n$. Alors l'application donnée par $\varphi\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = x_1y_1 + \cdots + x_py_p$ est bilinéaire symétrique positive mais n'est pas définie positive. En effet, $\varphi(e_n, e_n) = 0$ où e_n est le n-ème vecteur de la base canonique, et pourtant $e_n \neq 0$.

3. Soit $[a, b]$ un intervalle fermé de $\mathbb{R}$ et $E = \mathcal{C}^0([a, b])$ l'espace vectoriel des fonctions continues sur $[a, b]$. Alors l'application donnée par

$$\varphi : (f, g) \in E \times E \mapsto \int_a^b f(x)g(x)dx$$

est un produit scalaire sur E. On a déjà vu qu'elle est bilinéaire symétrique et on vérifie aisément qu'elle est positive. Rappelons que si h est continue positive sur $[a, b]$, alors

$$\int_a^b h(t)dt = 0 \Leftrightarrow \forall t \in [a, b],\ h(t) = 0$$

Ainsi

$$\varphi(f, f) = 0 \Leftrightarrow \forall t \in [a, b],\ f(t)^2 = 0 \Leftrightarrow f = 0$$

Par conséquent, φ est définie positive, donc c'est un produit scalaire.

4. Soit $E = \mathcal{M}_n(\mathbb{R})$. Alors l'application définie sur $E \times E$ par

$$\varphi(A, B) = \mathrm{Tr}({}^tAB) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{k,\ell} b_{k,\ell},$$

où $A = (a_{i,j})$ et $B = (b_{i,j})$ est un produit scalaire. Pour le vérifier :
   (a) Montrer d'abord l'égalité entre les deux formules.
   (b) Montrer qu'elle est bilinéaire symétrique grâce à la formule $\mathrm{Tr}({}^tAB)$.
   (c) Montrer qu'elle est définie positive grâce à la formule $\sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{k,\ell} b_{k,\ell}$.

Exercice 8. Démontrer toutes les affirmations de ces exemples.

I.2.c Normes euclidiennes

Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

Notation I.2.6. Si $x \in E$, on sait que $\langle x, x \rangle \geqslant 0$. On peut donc poser

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

Proposition I.2.7. *Pour tout $x, y \in E$ et $\lambda \in \mathbb{R}$, on a :*
1. *$\|x\| \geqslant 0$ et $\|x\| = 0 \Leftrightarrow x = 0$ **(séparation)** ;*
2. *$\|\lambda x\| = |\lambda| \|x\|$ **(positive homogénéité)** ;*
3. *$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$ **(identité de polarisation 1)**.*

Démonstration. Pour la séparation, on a directement

$$\|x\| = 0 \Leftrightarrow \langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

Pour la positive homogénéité, on a

$$\begin{aligned} \|\lambda x\| &= \sqrt{\langle \lambda x, \lambda x \rangle} && \text{par définition,} \\ &= \sqrt{\lambda^2 \langle x, x \rangle} && \text{par bilinéarité,} \\ &= |\lambda| \sqrt{\langle x, x \rangle} && \text{car } \sqrt{\lambda^2} = |\lambda| \\ &= |\lambda| \|x\| \end{aligned}$$

Pour la formule de polarisation, on développe par bilinéarité :

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle && \text{par définition,} \\ &= \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle && \text{par bilinéarité,} \\ &= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle && \text{par symétrie,} \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \end{aligned}$$

□

Remarque I.2.8. Pour la retenir, la formule de polarisation généralise l'identité remarquable $(a+b)^2 = a^2+2ab+b^2$ avec des produits scalaires partout là où c'est possible.

Corollaire I.2.9 (Identités)**.** *Soient $x, y \in E$. On a :*

1. $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ ***(identité de la médiane)**;*
2. $\|x+y\|^2 - \|x-y\|^2 = 4\langle x, y\rangle$ ***(identité de polarisation 2)**.*

Démonstration. Cela découle directement des formules

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \qquad \text{et } \|x-y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2.$$

□

Théorème I.2.10 (Inégalité de Cauchy-Schwarz)**.** *On a*

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

avec égalité si, et seulement si, les vecteurs x et y sont colinéaires.

Démonstration. Soient $x, y \in E$ deux vecteurs quelconques.

Si $x = y = 0$, la formule $|\langle x, y\rangle| \leqslant \|x\|\|y\| = 0$ est évidente.

Par symétrie, on peut donc supposer que $y \neq 0$. Considérons alors l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(t) = \|x + ty\|^2$.

D'après la Proposition I.2.7, on a

$$f(t) = \|x\| + 2\langle x, ty\rangle + \|ty\|^2 = \|y\|^2 t^2 + 2\langle x, y\rangle t + \|x\|^2.$$

On constate que $f(t)$ est un polynôme en t. Comme on a supposé $y \neq 0$, on sait que $\|y\|^2 = \langle y, y\rangle \neq 0$ car le produit scalaire est défini positif. Donc f est un polynôme de degré 2 qui est toujours positif ou nul. Son discriminant est donc négatif, autrement dit

$$(2\langle x, y\rangle)^2 - 4\|x\|^2\|y\|^2 \leqslant 0$$

ce qui nous donne bien

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|.$$

Le cas d'égalité signifie que le discriminant est nul. Le polynôme P admet donc une racine (double) en un nombre réel λ. Cela signifie que $P(\lambda) = \|x + \lambda y\|^2 = 0$. On en déduit que $x + \lambda y = 0$ et donc x et y sont colinéaires.

Réciproquement, supposons que x et y sont colinéaires. Par symétrie, on peut supposer sans perte de généralité qu'on peut écrire $y = \lambda x$. On a alors

$$|\langle x, y\rangle| = |\langle x, \lambda x\rangle| = |\lambda|\|x\|^2 = \|x\| \cdot |\lambda|\|x\| = \|x\|\|y\|.$$

Ce qui terminer de démontrer que le cas d'égalité se produit si, et seulement si, x et y sont colinéaires. □

Remarque I.2.11. L'inégalité de Cauchy-Schwarz est utile en géométrie pour généraliser la notion d'angle. Lorsque x et y sont deux vecteurs **non nuls**, la quantité

$$C(x, y) = \frac{\langle x, y\rangle}{\|x\|\|y\|} \in [-1, 1]$$

est un nombre compris entre -1 et 1 qu'on peut interpréter comme le cosinus d'un angle. On appelle *angle non orienté* entre x et y le nombre réel

$$\widehat{(x, y)} = \arccos\big(C(x, y)\big) \in [0, \pi],$$

de sorte que

$$\langle x, y\rangle = \|x\|\|y\| \cos\big(\widehat{(x, y)}\big).$$

On ne peut évidemment pas définir d'angle avec un vecteur nul.

Corollaire I.2.12 (Inégalité triangulaire)**.** *Pour tout $x, y \in E$, on a*

1. $\|x+y\| \leqslant \|x\| + \|y\|$ ***(inégalité triangulaire directe)**;*
2. $\big|\|x\| - \|y\|\big| \leqslant \|x-y\|$ ***(inégalité triangulaire renversée)**.*

Démonstration. Soient $x, y \in E$. Par l'inégalité de CAUCHY-SCHWARZ, on a immédiatement

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

Par croissante de l'application $\sqrt{\cdot}$, on en déduit

$$\|x+y\| \leqslant \|x\| + \|y\|,$$

d'où la première inégalité triangulaire.

Pour l'inégalité renversée, on applique l'inégalité triangulaire à $x' = x - y$ et $y' = y$. On a alors

$$\|x' + y'\| = \|x\| \leqslant \|x'\| + \|y'\| = \|x - y\| + \|y\|$$

En soustrayant $\|y\|$, on a donc

$$\|x\| - \|y\| \leqslant \|x - y\|.$$

Ceci étant valable pour tout $x, y \in E$, on peut donc également l'appliquer à $x' = y$ et $y' = x$. Ainsi

$$\|y\| - \|x\| \leqslant \|x - y\|.$$

Finalement

$$\big|\|x\| - \|y\|\big| = \max(\|x\| - \|y\|, \|y\| - \|x\|) \leqslant \|x - y\|.$$

□

Pour résumé, dans un espace euclidien, si $x, y \in E$ et $\lambda \in \mathbb{R}$, on a les propriétés :

1. $\|x\| \geqslant 0$ (**positivité**) et $\|x\| = 0 \Leftrightarrow x = 0$ (**séparation**) ;
2. $\|\lambda x\| = |\lambda|\|x\|$ (**positive homogénéité**) ;
3. $\|x+y\| \leqslant \|x\| + \|y\|$ (**inégalité triangulaire directe**) ;
4. $\big|\|x\| - \|y\|\big| \leqslant \|x - y\|$ (**inégalité triangulaire renversée**) ;
5. $\big|\langle x, y\rangle\big| \leqslant \|x\|\|y\|$ (**inégalité de Cauchy-Schwarz**) ;
6. $\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$ (**identité de polarisation 1**) ;
7. $\|x+y\|^2 - \|x-y\|^2 = 4\langle x, y\rangle$ (**identité de polarisation 2**) ;
8. $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (**identité de la médiane** ou **du parallélogramme**).

Les identités de polarisation et de la médiane se retiennent facilement en pensant aux identités remarquables.

Exercice 9. Montrer qu'on a également

$$\forall x, y \in E,\ \|x\|^2 - \|y\|^2 = \langle x + y, x - y\rangle,$$

ce qui est l'analogue d'une autre identité remarquable.

Remarque I.2.13. Si une application $N : E \to \mathbb{R}$ vérifie les quatre propriétés de positivité, séparation, positive homogénéité et inégalité triangulaire, on dit que N est une *norme* sur E.

En particulier, l'application $\|\cdot\|$ qu'on a étudié est une norme sur E. On dit qu'elle est issue du produit scalaire $\langle\cdot,\cdot\rangle$.

En général, il existe des normes qui ne sont pas issues de produit scalaires, par exemple la norme $N_1(x_1, \ldots, x_n) = |x_1| + \cdots + |x_n|$ vérifie la positivité, la séparation, la positive homogénéité et l'inégalité triangulaire mais n'est pas issue d'un produit scalaire. Pour le démontrer, on peut par exemple utiliser l'identité de polarisation 1 et constater que l'application $b(x, y) = \frac{1}{2}\left(N_1(x+y)^2 - N_1(x)^2 - N_1(y)^2\right)$ n'est pas une forme bilinéaire dès que $n \geqslant 2$.

I.3 Orthogonalité

On a vu dans la remarque I.2.11 qu'on pouvait définir une notion d'angle entre les vecteurs. Un par particulier est celui de l'angle droit, cela correspond au cas d'annulation du produit scalaire.

Dans la suite de cette section, soit E un espace vectoriel muni d'un produit scalaire $\langle\cdot,\cdot\rangle$.

I.3.a Vecteurs orthogonaux, espaces orthogonaux

Définition I.3.1 (vecteurs orthogonaux)**.** Soient $x, y \in E$ deux vecteurs. On di que x est *orthogonal* à x lorsque $\langle x, y\rangle = 0$. On note alors $x \perp y$.

Fait I.3.2. *Si x est orthogonal à y, alors y est aussi orthogonal à x par symétrie. On peut dire plus simplement que x et y sont orthogonaux.*

Définition I.3.3 (ensembles orthogonaux)**.** Deux sous-ensembles non vides $A, B \subseteq E$ sont dit *orthogonaux* lorsque

$$\forall a \in A,\ \forall b \in B,\ \langle a, b\rangle = 0.$$

Si A est une partie non vide de E, on appelle *orthogonal de A* et on note $A^\perp$ l'ensemble

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

Exercice 10. Montrer que si $A \subseteq B$, alors $B^\perp \subseteq A^\perp$.

Proposition I.3.4. *L'ensemble $A^\perp$ est un sous-espace vectoriel de E.*

Démonstration. Il suffit de montrer que $A^\perp$ et stable par somme et multiplication par un scalaire.

Soient $x, y \in A^\perp$ et $\lambda \in \mathbb{R}$. On a

$$\forall a \in A,\ \langle x, a\rangle = 0 \text{ et } \langle y, a\rangle = 0$$

donc par linéarité à gauche du produit scalaire, on a

$$\forall a \in A,\ \langle x + \lambda y, a\rangle = \langle x, a\rangle + \lambda\langle y, a\rangle = 0 + \lambda 0 = 0.$$

□

Fait I.3.5. *On a $\{0\}^\perp = E$ et $E^\perp = 0$.*

Exercice 11. Si $x \in E$, alors $\langle x, 0\rangle = 0$ donc $x \in \{0\}^\perp$. Ainsi $E \subseteq \{0\}^\perp$, d'où l'égalité.

Si $x \in E^\perp$, comme $x \in E$, on a $\langle x, x\rangle = 0$. Donc $x = 0$. Ainsi $E^\perp \subseteq \{0\}$, d'où l'égalité.

Proposition I.3.6. *Soit F un sous-espace vectoriel de E et $(f_1, \dots, f_p)$ une famille génératrice de F. Alors $F^\perp = \{f_1, \dots, f_p\}^\perp$.*

Démonstration. On procède par double inclusion pour démontrer l'égalité. Soit $x \in E$.

Supposons que $x \in F^\perp$. Soit i tel que $1 \leqslant i \leqslant p$. On a alors $\langle x, f_i\rangle = 0$ car $f_i \in F$ et $x \in F^\perp$. Ceci étant vrai pour tout i, on a donc $x \in \{f_1, \dots, f_p\}^\perp$. Ceci démontre que $F^\perp \subseteq \{f_1, \dots, f_p\}^\perp$.

Réciproquement, supposons que $x \in \{f_1, \dots, f_p\}^\perp$. Soit $f \in F$. Comme la famille $(f_1, \dots, f_p)$ est génératrice de F, il existe des réels $\lambda_1, \dots \lambda_p$ (pas nécessairement uniques) tels que $f = \sum_{i=1}^p \lambda_i f_i$. Alors par linéarité à droite du produit scalaire, on a

$$\langle x, f\rangle = \langle x, \sum_{i=1}^p \lambda_i f_i\rangle = \sum_{i=1}^p \lambda_i \underbrace{\langle x, f_i\rangle}_{=0} = 0.$$

Ceci étant vrai pour tout $f \in F$, on en déduit que $x \in F^\perp$. Donc $\{f_1, \dots, f_p\}^\perp \subseteq F^\perp$. □

Ainsi, pour calculer explicitement l'orthogonal d'un sous-espace vectoriel F de E, il sera plus commode de calculer d'abord une base $(f_1, \dots, f_p)$ de F puis de résoudre le système d'équations $\langle x, f_i\rangle = 0$ pour $1 \leqslant i \leqslant p$.

Exemple I.3.7. Supposons que $E = \mathbb{R}^2$ est muni du produit scalaire canonique.

1. Soit $D = \mathrm{Vect}(\begin{pmatrix}1\\0\end{pmatrix})$ la droite des abscisses. Alors

$$\begin{aligned} D^\perp &= \left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}^\perp \\ &= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ \langle\begin{pmatrix}x\\y\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\rangle = 0\right\} \\ &= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ 1x + 0y = 0\right\} \\ &= \mathrm{Vect}(\begin{pmatrix}0\\1\end{pmatrix}) \end{aligned}$$

est la droite des ordonnées.

2. Soit $D = \text{Vect}(\begin{pmatrix}1\\1\end{pmatrix})$. Alors

$$\begin{aligned}
D^\perp &= \left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}^\perp \\
&= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ \langle \begin{pmatrix}x\\y\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \rangle = 0\right\} \\
&= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ 1x + 1y = 0\right\} \\
&= \text{Vect}(\begin{pmatrix}-1\\1\end{pmatrix})
\end{aligned}$$

3. Plus généralement, si $D = \text{Vect}(\begin{pmatrix}a\\b\end{pmatrix})$ est la droite du plan dirigée par le vecteur de coordonnées $(a, b) \neq (0, 0)$. Alors

$$\begin{aligned}
D^\perp &= \left\{\begin{pmatrix}a\\b\end{pmatrix}\right\}^\perp \\
&= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ \langle \begin{pmatrix}x\\y\end{pmatrix}, \begin{pmatrix}a\\b\end{pmatrix} \rangle = 0\right\} \\
&= \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2,\ ax + by = 0\right\} \\
&= \text{Vect}(\begin{pmatrix}-b\\a\end{pmatrix})
\end{aligned}$$

Exemple I.3.8. Supposons que $E = \mathbb{R}^3$ est muni du produit scalaire canonique.

1. Si $D = \text{Vect}(\begin{pmatrix}a\\b\\c\end{pmatrix})$ est la droite de l'espace euclidien de dimension 3 dirigée par le vecteur de coordonnées $(a, b, c) \neq (0, 0, 0)$. Alors

$$\begin{aligned}
D^\perp &= \left\{\begin{pmatrix}a\\b\\c\end{pmatrix}\right\}^\perp \\
&= \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ \langle \begin{pmatrix}x\\y\\z\end{pmatrix}, \begin{pmatrix}a\\b\\c\end{pmatrix} \rangle = 0\right\} \\
&= \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ ax + by + cz = 0\right\}
\end{aligned}$$

est le plan d'équation $ax + by + cz = 0$. Le vecteur (a, b, c) est orthogonal à ce plan. On dit que c'est un vecteur *normal* au plan.

2. Si $P = \text{Vect}(\begin{pmatrix}a_1\\b_1\\c_1\end{pmatrix}, \begin{pmatrix}a_2\\b_2\\c_2\end{pmatrix})$ est le plan de l'espace euclidien de dimension 3 engendré par les vecteurs non colinéaires (a_1, b_1, c_1) et (a_2, b_2, c_3). Alors

$$\begin{aligned}
P^\perp &= \left\{\begin{pmatrix}a_1\\b_1\\c_1\end{pmatrix}, \begin{pmatrix}a_2\\b_2\\c_2\end{pmatrix}\right\}^\perp \\
&= \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ \langle \begin{pmatrix}x\\y\\z\end{pmatrix}, \begin{pmatrix}a_1\\b_1\\c_1\end{pmatrix} \rangle = 0 \text{ et } \langle \begin{pmatrix}x\\y\\z\end{pmatrix}, \begin{pmatrix}a_2\\b_2\\c_2\end{pmatrix} \rangle = 0\right\} \\
&= \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ \left\{\begin{array}{l} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{array}\right.\right\}
\end{aligned}$$

est la droite donnée par le système d'équations $\left\{\begin{array}{l} a_1x + b_1y + c_1z = 0 \\ a_2x + b_2y + c_2z = 0 \end{array}\right.$

On peut également s'intéresser à l'orthogonalité dans des espaces de matrices ou de fonctions.

Exercice 12. Soit $E = \mathcal{C}^0([0,\pi])$ l'espace vectoriel des fonctions continues sur l'intervalle fermé $[0,\pi]$. On munit E du produit scalaire $\langle u, v\rangle = \int_0^\pi u(t)v(t)dt$.

Montrer que $\cos \perp \sin$ pour ce produit scalaire.

Exercice 13. On se place dans $E = \mathcal{M}_n(\mathbb{R})$ muni du produit scalaire $\varphi(A,B) = \mathrm{Tr}({}^tAB)$. On note $\mathcal{D}_n(\mathbb{R})$ (resp. $\mathcal{S}_n(\mathbb{R})$, $\mathcal{AS}_n(\mathbb{R})$) le sous-ensemble de $\mathcal{M}_n(\mathbb{R})$ formé des matrices diagonales (resp. symétriques, resp. antisymétriques). Déterminer $\mathcal{D}_n(\mathbb{R})^\perp$ (resp. $\mathcal{S}_n(\mathbb{R})^\perp$, $\mathcal{AS}_n(\mathbb{R})^\perp$)

Exercice 14. On considère l'ensemble E des fonctions de classe $\mathcal{C}^1$ sur $[-1,1]$ d'intégrale nulle sur $[-1,1]$, c'est-à-dire que $f \in E$ si $f : [-1,1] \to \mathbb{R}$ est dérivable, de dérivée continue et telle que $\int_{-1}^1 f(t)dt = 0$.

1. Montrer que E est un espace vectoriel.
2. Montrer que $\varphi : E \times E \to \mathbb{R}$ donné par $\varphi(u,v) = \int_{-1}^1 u'(t)v'(t)dt$ est un produit scalaire sur E.
3. Montrer que la fonction $u(t) = t$ est un élément de E, puis que $u^\perp$ est l'ensemble des fonctions $f \in E$ telles que $f(1) = f(-1)$.

I.3.b Familles et bases orthogonales et orthonormées

Supposons qu'on veuille calculer un certain nombre de produits scalaire en ayant écrit certains vecteurs dans une base donnée. Par exemple, pour le produit scalaire canonique de $\mathbb{R}^n$ et des vecteurs écrits dans la base canonique, on a

$$\langle x_1e_1 + \cdots + x_ne_n, y_1e_1 + \cdots + y_ne_n\rangle = x_1y_1 + \cdots + x_ny_n$$

ce qui nous fait $\Theta(n)$ multiplications et additions à effectuer.

En revanche, pour un produit scalaire quelconque et une base quelconque, on aurait

$$\langle x_1e_1 + \cdots + x_ne_n, y_1e_1 + \cdots + y_ne_n\rangle = \sum_{i=1}^n x_iy_j\langle e_i, e_j\rangle$$

ce qui ferait $\Theta(n^2)$ multiplications et additions à effectuer.

Il semble que si la première situation nous est plus favorable que la seconde, c'est parce qu'on a souvent l'égalité $\langle e_i, e_j\rangle = 0$. On va donc s'intéresser à des bases ayant cette propriété.

Définition I.3.9. Un vecteur $x \in E$ est dit *unitaire* si $\|x\| = 1$.

Soit $(f_1, \ldots, f_p)$ une famille de vecteurs. On dit qu'elle est

1. *orthogonale* si $\forall i \neq j,\ \langle f_i, f_j\rangle = 0$;
2. *orthonormée* si $\forall i, j \langle f_i, f_j\rangle = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$

Une *base orthogonale* (resp. *base orthonormée*) est une famille orthogonale (resp. orthonormée) qui est une base de E.

Théorème I.3.10 (Théorème de Pythagore)**.** *Soit $\mathcal{F} = (f_1, \ldots, f_p)$ une famille orthogonale de vecteurs. Alors*

1. $\|\sum_{i=1}^p f_i\|^2 = \sum_{i=1}^p \|f_i\|^2$ *;*
2. *si, de plus, les vecteurs de la famille sont **tous non nuls**, alors la famille est libre.*

En particulier, une famille orthonormée est libre.

Démonstration. Pour l'égalité, on procède par récurrence sur p en montrant

$$\mathcal{P}(p) : \forall (f_1, \ldots, f_p) \text{ libre },\ \|\sum_{i=1}^p f_i\|^2 = \sum_{i=1}^p \|f_i\|^2.$$

L'initialisation pour $\mathcal{P}(p = 1)$ est immédiate.

Hérédité : Supposons $\mathcal{P}(p)$ et montrons $\mathcal{P}(p+1)$. Soit $(f_1, \ldots, f_{p+1})$ une famille orthogonale. On observe que $(f_1, \ldots, f_p)$ est également une famille orthogonale. On a donc

$$\|\sum_{i=1}^p f_i\|^2 = \sum_{i=1}^p \|f_i\|^2.$$

Posons $x = \sum_{i=1}^{p} f_i$ et $y = f_{p+1}$. Par l'identité de polarisation, on sait que

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2.$$

Mais

$$\langle x, y\rangle = \langle \sum_{i=1}^{p} f_i, f_{p+1}\rangle = 0$$

par orthogonalité de la famille. Ainsi

$$\|x+y\|^2 = \|\sum_{i=1}^{p} f_i + y\|^2 = \sum_{i=1}^{p} \|f_i\|^2 + \|y\|^2 = \sum_{i=1}^{p+1} \|f_i\|^2.$$

D'où l'hérédité.

Pour la liberté, donnons-nous une combinaison linéaire nulle

$$\sum_{i=1}^{p} \lambda_i f_i = 0$$

Alors pour tout $1 \leqslant j \leqslant p$, on a

$$0 = \langle \sum_{i=1}^{p} \lambda_i f_i, f_j\rangle = \sum_{i=1}^{p} \lambda_i \underbrace{\langle f_i, f_j\rangle}_{=0 \text{ sauf si } i=j} = \lambda_j \langle f_j, f_j\rangle$$

Comme $f_j \neq 0$, on a $\langle f_j, f_j\rangle \neq 0$ donc $\lambda_j = 0$. Ceci étant vrai pour tout $1 \leqslant j \leqslant p$, on en déduit que la famille est libre. □

Proposition I.3.11. *Soit $(E, \langle\cdot,\cdot\rangle$ un espace euclidien et $\mathcal{E} = (e_1, \dots, e_n)$ une base.*

1. *La base $\mathcal{E}$ est orthonormée si, et seulement si, $\mathrm{Mat}(\langle\cdot,\cdot\rangle, \mathcal{E}) = I_n$.*
2. *Supposons que $\mathcal{E}$ est orthonormée. Soit $x \in E$ qu'on écrit $x = \sum_{i=1}^{n} x_i e_i$ dans la base. Alors $x_i = \langle x_i, e_i\rangle$. Ainsi, trouver les coordonnées d'un vecteur dans une base orthonormée se lit par des produits scalaires.*
3. *Supposons que $\mathcal{E}$ est orthonormée. Soient $x, y \in E$ qu'on écrit $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$ dans la base $\mathcal{E}$. Alors*
$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \qquad \textit{et} \qquad \langle x, x\rangle = \sum_{i=1}^{n} x_i^2$$
Ainsi, écrire des vecteurs dans une base orthonormée nous donne des formules comme si on travaillait avec le produit scalaire canonique.

Démonstration. Pour 1, la base $\mathcal{E}$ est orthonormée si, et seulement si, le coefficient $a_{i,j}$ de la matrice $\mathrm{Mat}(\langle\cdot,\cdot\rangle, \mathcal{E})$ est

$$a_{i,j} = \langle e_i, e_j\rangle = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

On reconnait alors les coefficients de la matrice I_n.

Pour 2, pour $1 \leqslant j \leqslant n$, on calcule

$$\langle x, e_j\rangle = \langle \sum_{i=1}^{n} x_i e_i, e_j\rangle = \sum_{i=1}^{n} x_i \underbrace{\langle e_i, e_j\rangle}_{=0 \text{ sauf si } i=j} = x_j \langle e_j, e_j\rangle = x_j.$$

Pour 3, on calcule

$$\langle x, y\rangle = \langle \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \langle e_i, e_j\rangle = \sum_{1 \leqslant i=j \leqslant n} x_i y_i \langle e_i, e_i\rangle = \sum_{i=1}^{n} x_i y_i$$

D'où la formule souhaitée. □

I.3.c Procédé d'orthogonalisation de Gram-Schmidt

On dispose d'un algorithme qui permet de construire des bases orthonormées.

Théorème I.3.12 (Procédé d'orthonormalisation de Gram-Schmidt)**.** *Soit E un espace vectoriel muni d'un produit scalaire et $(f_1, \dots, f_p)$ une famille libre de E.*

On définit récursivement simultanément deux familles $(g'_1, \dots, g'_p)$ et $(g_1, \dots, g_p)$ en posant :

$$\begin{cases} g'_1 = f_1 \\ g_1 = \frac{g'_1}{\|g'_1\|} \end{cases}$$

et pour $1 \leqslant k \leqslant p-1$:

$$\begin{cases} g'_{k+1} = f_{k+1} - \sum_{i=1}^{k} \langle f_{k+1}, g_i \rangle g_i \\ g_{k+1} = \frac{g'_{k+1}}{\|g'_{k+1}\|} \end{cases}$$

Alors :

1. *la famille $(g'_1, \dots, g'_p)$ est orthogonale ;*
2. *la famille $(g_1, \dots, g_p)$ est orthonormée ;*
3. *pour tout $1 \leqslant k \leqslant p$, on a*

$$\mathrm{Vect}(f_1, \dots, f_k) = \mathrm{Vect}(g'_1, \dots, g'_k) = \mathrm{Vect}(g_1, \dots, g_k).$$

Démonstration. On procède par récurrence sur p. C'est évident si $p = 1$.

Hérédité : Soit $(f_1, \dots, f_{p+1})$ une famille libre de E. On sait par hypothèse de récurrence qu'on peut alors construire comme indiqué des familles $(g'_1, \dots, g'_p)$ et $(g_1, \dots, g_p)$ avec les propriétés voulues. On pose alors $g'_{p+1} = f_{p+1} - \sum_{i=1}^{p} \langle f_{p+1}, g_i \rangle g_i$.

Alors pour $1 \leqslant i \leqslant p$, on sait déjà que $g'_i \in \mathrm{Vect}(f_1, \dots, f_p) \subseteq \mathrm{Vect}(f_1, \dots, f_{p+1})$. De plus, g'_{p+1} est une combinaison linéaires de vecteurs dans $\mathrm{Vect}(f_1, \dots, f_{p+1})$ donc $g'_{p+1} \in \mathrm{Vect}(f_1, \dots, f_{p+1})$. Ainsi $\mathrm{Vect}(g'_1, \dots, g'_{p+1}) \subseteq \mathrm{Vect}(f_1, \dots, f_{p+1})$

Réciproquement, on observe également que pour $1 \leqslant i \leqslant p$, on sait déjà que $f_i \in \mathrm{Vect}(g'_1, \dots, g'_p) \subseteq \mathrm{Vect}(g'_1, \dots, g'_{p+1})$. Comme $f_{p+1} = g'_{p+1} + \sum_{i=1}^{p} \frac{\langle f_{p+1}, g_i \rangle \|}{\|g'_i\|} g'_i$, on a également écrit f_{p+1} comme combinaison linéaire en les g'_i. Donc $f_{p+1} \in \mathrm{Vect}(g'_1, \dots, g'_{p+1})$. Ainsi $\mathrm{Vect}(g'_1, \dots, g'_{p+1}) \supseteq \mathrm{Vect}(f_1, \dots, f_{p+1})$, d'où l'égalité par double inclusion.

Montrons que la famille $(g'_1, \dots, g'_{p+1})$ est orthogonale. On a tout fait pour. En effet, pour $1 \leqslant j \leqslant p$, on a

$$\langle g'_{p+1}, g'_j \rangle = \langle f_{p+1}, g'_j \rangle - \sum_{i=1}^{p} \langle f_{p+1}, g_i \rangle \langle g_i, g'_j \rangle$$

Or $\langle g_i, g'_j \rangle = \|g'_j\| \langle g_i, g_j \rangle = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$. Donc

$$\langle g'_{p+1}, g'_j \rangle = \|g'_j\| \langle f_{p+1}, g_j \rangle - \|g'_j\| \langle f_{p+1}, g_j \| = 0$$

D'où l'orthogonalité de la famille $(g'_1, \dots, g'_{p+1})$.

Il reste à voir s'il est possible de définir g_{p+1}. Pour cela, on doit s'assurer que $g'_{p+1} \neq 0$. Si par l'absurde $g'_{p+1} = 0$, alors on aurait

$$f_{p+1} = \sum_{i=1}^{p} \langle f_{p+1}, g_i \rangle g_i \in \mathrm{Vect}(f_1, \dots, f_p)$$

ce qui contredit le fait que la famille $(f_1, \dots, f_{p+1})$ est libre. Donc $g'_{p+1} \neq 0$. Donc g_{p+1} est bien défini.

Or, comme la famille $(g'_1, \dots, g'_{p+1})$ est orthogonale, la famille $(g_1, \dots, g_{p+1})$ l'est aussi. Elle est, de plus, orthonormée car pour $1 \leqslant i \leqslant p+1$, on a

$$\|g_i\|^2 = \langle g_i, g_i \rangle = \langle \frac{g'_i}{\|g'_i\|}, \frac{g'_i}{\|g'_i\|} \rangle = \frac{\langle g'_i, g'_i \rangle}{\|g'_i\|^2} = 1.$$

□

Pour mieux comprendre le procédé d'orthonormalisation de Gram-Schmidt, traitons un exemple simple.

Exemple I.3.13. Soit $E = \mathbb{R}^3$ et $e_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $e_2 = \begin{pmatrix}1\\2\\0\end{pmatrix}$, $e_3 = \begin{pmatrix}1\\0\\0\end{pmatrix}$. On observe que (e_1, e_2, e_3) est une base de $\mathbb{R}$. En effet, la matrice de passage de la base canonique vers cette famille est $P = \begin{pmatrix}1 & 1 & 1\\0 & 2 & 0\\1 & 0 & \end{pmatrix}$ qui est inversible car $\det(P) = -2 \neq 0$.

Appliquons alors le procédé de Gram-Schmidt.

Premiers vecteurs $g_1' = e_1$ et $g_1 = \frac{g_1'}{\|g_1'\|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}$.

Deuxièmes vecteurs $g_2' = e_2 - \langle e_2, g_1\rangle g_1 = \begin{pmatrix}1\\2\\0\end{pmatrix} - \frac{1+0+0}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1-\frac{1}{2}\\2-0\\0-\frac{1}{2}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\4\\-1\end{pmatrix}$

$g_2 = \frac{g_2'}{\|g_2'\|} = \frac{2}{\sqrt{1^2+4^2+(-1)^2}}\frac{1}{2}\begin{pmatrix}1\\4\\-1\end{pmatrix} = \frac{1}{3\sqrt{2}}\begin{pmatrix}1\\4\\-1\end{pmatrix}$

Troisièmes vecteurs $g_3' = e_3 - \langle e_3, g_1\rangle g_1 - \langle e_3, g_2\rangle g_2 = \begin{pmatrix}1\\0\\0\end{pmatrix} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{3\sqrt{2}}\frac{1}{3\sqrt{2}}\begin{pmatrix}1\\4\\-1\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix} - \frac{1}{18}\begin{pmatrix}1\\4\\-1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}\frac{4}{9}\\-\frac{2}{9}\\-\frac{4}{9}\end{pmatrix}$

$g_3 = \frac{g_3'}{\|g_3'\|} = \frac{9}{\sqrt{4^2+2^2+(-4)^2}}\begin{pmatrix}\frac{4}{9}\\-\frac{2}{9}\\-\frac{4}{9}\end{pmatrix} = \frac{9}{6}\begin{pmatrix}\frac{4}{9}\\-\frac{2}{9}\\-\frac{4}{9}\end{pmatrix} = \begin{pmatrix}\frac{2}{3}\\-\frac{1}{3}\\-\frac{2}{3}\end{pmatrix}$

Ainsi $g_1 = \begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}$, $g_2 = \begin{pmatrix}\frac{1}{3\sqrt{2}}\\\frac{2\sqrt{2}}{3}\\-\frac{1}{3\sqrt{2}}\end{pmatrix}$, $g_3 = \begin{pmatrix}\frac{2}{3}\\-\frac{1}{3}\\-\frac{2}{3}\end{pmatrix}$ est une base orthonormée de $\mathbb{R}^3$.

Remarque I.3.14. On remarque en particulier que si E est un espace euclidien, alors il existe des bases dans E donc des bases orthonormées.

Il n'y a pas du tout unicité des bases orthonormé, ça n'a aucun sens de dire « la base orthonormée » de E car on ne sait pas comment la choisir en général. Il faudra veiller à préciser de quelle base orthonormée on parle à chaque fois.

Théorème I.3.15 (base orthonormée incomplète)**.** *Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension n. Soit $(e_1, \ldots, e_p)$ une famille orthonormée de E. Alors il existe $(e_{p+1}, \ldots, e_n)$ tels que $(e_1, \ldots, e_n)$ est une base orthonormée de E.*

Démonstration. D'après le théorème de Pythagore, la famille $(e_1, \ldots, e_p)$ est libre. On peut donc la compléter en une base de E, c'est-à-dire $(e_1, \ldots, e_p, f_{p+1}, \ldots, f_n)$. En appliquant le procédé de Gram-Schmidt à cette famille, on trouve une famille orthonormée $(g_1 = e_1, \ldots, g_p = e_p, g_{+1}, \ldots, g_n)$ qui engendre $\mathrm{Vect}(e_1, \ldots, f_n) = E$. Donc $(g_1, \ldots, g_n)$ est une famille génératrice de E, et libre car orthonormée. C'est donc une base orthonormée de E qui complète $(e_1, \ldots, e_p)$. □

Théorème I.3.16. *Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension n. Soit F un sous-espace vectoriel de E de dimension p. Alors $F^\perp$ est un supplémentaire de F, i.e. $E = F \oplus F^\perp$.*

En particulier $\dim(F^\perp) = n - p$.

Démonstration. Si $F = 0$, on l'a déjà vu.

Sinon, soit $(e_1, \ldots, e_p)$ une base orthonormée de F qu'on complète en $(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ base orthonormée de E. Posons $G = \mathrm{Vect}(e_{p+1}, \ldots, e_n)$. Montrons que $G = F^\perp$.

Si $x \in G$. Écrivons $x = \sum_{i=p+1}^n \lambda_i e_i$. Alors pour tout $1 \leqslant j \leqslant p$, on a

$$\langle x, e_j\rangle = \sum_{i=p+1}^n \lambda_i \langle e_i, e_j\rangle = 0$$

Donc $x \in F^\perp$. Ainsi $G \subseteq F^\perp$.

Réciproquement, soit $y \in F^\perp$ qu'on écrit $y = \sum_{i=1}^n \lambda_i e_i$. Soit $1 \leqslant j \leqslant p$. Alors

$$0 = \langle y, e_j \rangle = \sum_{i=1}^n \lambda_i \langle e_i, e_j \rangle = \lambda_j.$$

Donc $y = \sum_{i=p+1}^n \lambda i e_i \in G$. Ainsi $F^\perp \subseteq G$ et donc en fait $F^\perp = \mathrm{Vect}(e_{p+1}, \dots, e_n)$.

De plus, par définition, on a $F + G = E$. Il reste à montrer que $F \cap F^\perp = 0$. Soit $x \in F \cap F^\perp$. Alors $\langle x, x \rangle = 0$ par définition de $F^\perp$, donc $x = 0$ par séparation. □

Notation I.3.17. On dit alors que F et $F^\perp$ sont *en somme directe orthogonale* et on note $F \overset{\perp}{\oplus} F^\perp = E$.

Plus généralement, soit $F_1, \dots, F_d$ des sous-espaces vectoriels de E qui sont deux à deux orthogonaux et tels que $E = F_1 + \cdots + F_d$. Donnons-nous une base orthonormée $\mathcal{E}_i$ de chaque F_i. Alors la concaténation $\mathcal{E}$ des bases $\mathcal{E}_i$ est une base orthonormée de E car elle est une famille orthonormée et génératrice. En particulier, on observe par récurrence que les F_i sont en somme directe. On note alors $E = F_1 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} F_d$ cette somme directe orthogonale.

Ici, l'orthogonalité pour un produit scalaire joue un rôle essentielle. Cette propriété de somme directe, qui se lit sur l'orthogonalité deux à deux, serait complètement fausse sans l'hypothèse d'orthogonalité : trois droites distinctes de $\mathbb{R}^2$ s'intersectent trivialement et pourtant, elles ne sont pas en somme directe.

Exercice 15. Soit E un espace euclidien et F, G deux sous-espaces vectoriels de E. Montrer que s'équivalent :

(i) $F^\perp \perp G^\perp$;

(ii) $F^\perp \subseteq G$;

(iii) $G^\perp \subseteq F$.

I.3.d Projections orthogonales

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien

Remarque I.3.18. En algèbre linéaire, pour définir un projecteur, il nous fallait deux sous-espaces en somme directe $E = F \oplus G$, ce qui permettait alors de définir la projection orthogonale $p_{F\|G} : E \to E$ sur F parallèlement à G de sorte que $\mathrm{im}(p_{F\|G}) = F$ et $\ker p_{F\|G} G$.

Définition I.3.19. Soit F un sous-espace vectoriel de E. On appelle *projection orthogonal sur F* l'endomorphisme $p_F : E \to E$ qui est la projection orthogonal sur F parallèlement à $F^\perp$, c'est-à-dire que si $x \in E$ s'écrit $x = y + z$ avec $y \in F$ et $z \in F^\perp$, alors $p_F(x) = y$.

Remarque I.3.20. Ceci est bien défini quel que soit le sous-espace vectoriel F car on a vu qu'alors $E = F \overset{\perp}{\oplus} F^\perp$. De plus, $F^\perp$ est uniquement déterminé par F, donc la définition de p_F ne dépend que de F.

Proposition I.3.21. *Soit E un espace euclidien et F un sous-espace vectoriel de E. Alors,*

1. *si $(e_1, \dots, e_k)$ est une base orthonormée de F, alors*

$$\forall x \in E,\ p_F(x) = \sum_{i=1}^k \langle x, e_i \rangle e_i, \tag{I.1}$$

2. *on a $\forall x \in E,\ x = p_F(x) + p_{F^\perp}(x)$, autrement dit $\mathrm{id}_E = p_F + p_{F^\perp}$,*
3. *si $(e_1, \dots, e_\ell)$ est une base orthonormée de $F^\perp$, alors*

$$\forall x \in E,\ p_F(x) = x - \sum_{j=1}^\ell \langle x, e_j \rangle e_j. \tag{I.2}$$

Quand on veut calculer des projections, en fonction de la dimension k de F et ℓ de $F^\perp$, il peut être plus commode d'utiliser la formule (I.1) ou (I.2).

Démonstration. On complète la famille orthonormée $(e_1, \dots, e_k)$ en une base orthonormée $(e_1, \dots, e_n)$ de E et on écrit $x = \sum_{i=1}^n x_i e_i$ dans cette base. On sait alors que $x_i = \langle x, e_i \rangle$. Ainsi

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

De plus, on sait également que $(e_{k+1}, \dots, e_n)$ est une base orthonormée de $F^\perp$. Donc

$$p_{F^\perp}(x) = \sum_{i=k+1}^{n} \langle x, e_i \rangle e_i$$

D'après ce qui précède. Ainsi

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i = \sum_{i=1}^{k} \langle x, e_i \rangle e_i + \sum_{i=k+1}^{n} \langle x, e_i \rangle e_i = p_F(x) + p_{F^\perp}(x)$$

Si $(e_1, \dots, e_\ell)$ est une base orthonormée de $F^\perp$, on a donc en appliquant (I.1) à $p_{F^\perp}$ que

$$p_F(x) = x - p_{F^\perp}(x) = x - \sum_{j=1}^{\ell} \langle x, e_j \rangle e_j.$$

□

Remarque I.3.22 (Interprétation de l'algorithme de Gram-Schmidt). Grâce à cette formule, on comprend désormais mieux ce qui a motivé l'algorithme de Gram-Schmidt : On avait en fait posé $F_k = \mathrm{Vect}(f_1, \dots, f_k)$ muni d'une base orthonormée $(g_1, \dots, g_k)$ et

$$g'_{k+1} = f_{k+1} - \sum_{i=1}^{k} \langle f_{k+1}, g_i \rangle g_i = f_{k+1} - p_{F_k}(f_{k+1}) = p_{F_k^\perp}(f_{k+1})$$

Autrement dit, l'élément qu'on va ajouter pour agrandir la famille orthogonale des (g'_i) avec un g'_{k+1} est le projeté orthogonal de f_{k+1} sur l'orthogonal $F_k^\perp$ de F_k, ce qui nous assure bien l'orthogonalité de g'_{k+1} avec les (g'_i).

I.3.e Distance euclidienne

On a vu que $\|\cdot\|$ était une norme sur E. Lorsqu'on dispose d'une norme sur un espace E, on peut lui associer une distance d sur E. C'est une application $d : E \times E \to \mathbb{R}$ qui vérifie :

positivité $\forall x, y \in E,\ d(x, y) \geqslant 0$;

séparation $\forall x, y \in E,\ d(x, y) = 0 \Leftrightarrow x = y$;

inégalité triangulaire $\forall x, y, z \in E,\ d(x, z) \leqslant d(x, y) + d(y, z)$.

Fait I.3.23. *On définit* $\begin{array}{rccc} d : & E \times E & \to & \mathbb{R} \\ & (x, y) & \mapsto & \|y - x\| \end{array}$. *C'est une distance sur* E.

Exercice 16. Le vérifier.

Cela nous permet de définir des distances entre les vecteurs d'un espace euclidien, qu'on pourra identifier à des points en géométrie. Plus généralement, on voudrait définir la distance d'un point à une partie. On pose

Définition I.3.24. Soit $v \in E$ et $A \subseteq E$ une partie **non vide** de E. On appelle *distance* de v à A, la quantité

$$d(v, A) = \inf\{d(v, x),\ x \in A\}$$

Soient $A, B \subseteq E$ deux parties de E. On appelle *distance* de A à B la quantité

$$d(A, B) = \inf\{d(x, y),\ x \in A,\ y \in B\}$$

En général, il est délicat de calculer des distances et il faut être méticuleux. Il y a un cas particulier dans lequel on sait faire aisément.

Proposition I.3.25. *Soit F un sous-espace vectoriel de E et $v \in E$ un vecteur. Alors*

$$d(v, F) = \|v - p_F(v)\| = \|p_{F^\perp}(v)\|$$

Démonstration. Posons $y = p_F(v)$. Soit $x \in F$ et posons $z = x - y \in F$ de sorte que $x = y + z$. On a alors $v - y = v - p_F(v) = p_{F^\perp}(v) \in F^\perp$. Donc $v - y$ et z sont orthogonaux. Ainsi, par le théorème de Pythagore, on a

$$\|v - x\|^2 = \|v - y - z\|^2 = \|v - y\|^2 + \|z\|^2 \geqslant \|v - y|^2$$

On observe donc que

$$\forall x \in F,\ \|v - x\| \geqslant \|v - y\|$$

avec égalité lorsque $x = y$. L'infimum des $d(v, x) = \|v - x\|$ pour $x \in F$ est donc atteint en $x = y = p_F(v)$. Donc $d(v, F) = \|v - p_F(v)\|$. Enfin, la seconde égalité découle de $v - p_F(v) = p_{F^\perp}(v)$. □

Le point qui réalise la distance est en fait unique, ce qu'on laisse en exercice.

Exercice 17. On suppose que F est un sous-espace vectoriel de E et $v \in E$. Montrer que $x \in F$ et $\|v, x\| = d(v, F)$ si, et seulement si $x = p_F(v)$.

Chapitre II

Endomorphismes remarquables d'un espace euclidien

Dans tout ce chapitre, on se fixe $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien de dimension $n > 0$.

II.1 Adjoint d'un endomorphisme

La bilinéarité des produits scalaires permet d'interpréter différemment les applications linéaires. Pour commencer, intéressons-nous aux formes linéaires.

Théorème II.1.1 (Dualité entre produit scalaire et formes linéaires)**.** *Soit $f : E \to \mathbb{R}$ une forme linéaire. Alors il existe un unique $z \in E$ tel que*

$$\forall x \in E,\ f(x) = \langle x, z \rangle.$$

Remarque II.1.2. Attention ! La dimension finie de E est essentielle dans ce théorème. Il existera un théorème analogue, appelé *théorème de représentation de Riesz*, dans des espaces de dimension infinie munis d'un produit scalaire, mais cela nécessitera une hypothèse de continuité sur la forme linéaire f, ce qui est automatique en dimension finie.

Démonstration. Pour l'unicité, supposons que z, z' répondent à la condition. Alors

$$\forall x \in E,\ 0 = f(x) - f(x) = \langle x, z \rangle - \langle x, z' \rangle = \langle x, z - z' \rangle$$

On a donc par définition de l'orthogonal que $(z - z') \in E^{\perp} = 0$. Ainsi $z = z'$.

Pour l'existence, si $f \equiv 0$ est la forme linéaire nulle, c'est immédiat en posant $z = 0$. Sinon, on sait alors que $\ker f = F$ est un sous-espace vectoriel de E de dimension $n-1$. Comme $E = F \oplus F^{\perp}$, il existe alors $z_0 \in F^{\perp}$ tel que $f(z_0) \neq 0$ car sinon f serait nulle sur E. Posons alors $z = \frac{f(z_0)}{\|z_0\|^2} z_0 \neq 0$ car $z_0 \neq 0$ et $f(z_0) \neq 0$. Soit $x \in E$ qu'on écrit $x = y + \alpha z$ avec $y \in F$ et $\alpha \in \mathbb{R}$, ce qui est possible car $z \neq 0$ et $F^{\perp}$ est de dimension 1. On a alors, d'une part,

$$\langle x, z \rangle = \langle y + \alpha z, z \rangle = \underbrace{\langle y, z \rangle}_{=0} + \alpha \langle z, z \rangle = \alpha \frac{f(z_0)^2}{\|z_0\|^4} \langle z_0, z_0 \rangle = \alpha \frac{f(z_0)^2}{\|z_0\|^2}$$

D'autre part,

$$f(x) = f(y + \alpha z) = \underbrace{f(y)}_{=0} + \alpha f(z) = \alpha \frac{f(z_0)}{\|z_0\|^2} f(z_0)$$

Ainsi, on a bien

$$\forall x \in E,\ \langle x, z \rangle = f(z).$$

□

Remarque II.1.3. On peut donc plus simplement comprendre les formes linéaires d'un espace euclidien comme des produits scalaires contre des vecteurs.

C'est un point de vue qui sera à la base de toute une partie de la résolution de systèmes différentiels au moyen d'intégrales. On définira des formes linéaires sur des espaces de fonction et on cherchera la « fonctionnelle » qui correspond à une forme linéaire donnée pour un produit scalaire défini par une intégrale.

Intéressons-nous maintenant aux endomorphismes de E.

Définition et théorème II.1.4 (Adjoint d'un endomorphisme)**.** Soit $u \in \mathrm{End}(E)$. Alors il existe un unique endomorphisme $u^* \in \mathrm{End}(E)$ tel que

$$\forall x, y \in E,\ \langle u(x), y\rangle = \langle x, u^*(y)\rangle$$

On appelle u^* l'*endomorphisme adjoint* de u dans E.

Démonstration. Pour l'unicité, supposons qu'il existe deux endomorphismes v, w vérifiant la propriétés. Soit $y \in E$, alors

$$\forall x \in E, 0 = \langle u(x), y\rangle - \langle u(x), y\rangle = \langle x, v(y)\rangle - \langle x, w(y)\rangle = \langle x, v(y) - w(y)\rangle$$

Donc $v(y) - w(y) \in E^{\perp} = 0$. Ceci étant vrai pour tout $y \in E$, on a donc

$$\forall y \in E,\ v(y) = w(y).$$

D'où l'unicité de u^* si toutefois il en existe un.

Pour l'existence, soit $y \in E$. Alors l'application $f_y : x \in E \mapsto \langle u(x), y\rangle \in \mathbb{R}$ est une forme linéaire sur E. Par le théorème de dualité II.1.1, il existe un unique $z_y \in E$ tel que

$$\forall x \in E,\ f_y(x) = \langle x, z_y\rangle$$

On pose alors $u^*(y) := z_y$. Comme le z_y est uniquement déterminé, cela permet de définir une fonction $u^* : E \to E$.

Il reste à montrer que u^* est linéaire. Soient $y, y' \in E$ et $\lambda, \mu \in E$. Alors

$$\begin{aligned}\forall x \in E,\ \langle x, u^*(\lambda y + \mu y)\rangle = \langle u(x), \lambda y + \mu y'\rangle &= \lambda\langle u(x), y\rangle + \mu\langle u(x), y'\rangle \\ &= \lambda\langle x, u^*(y)\rangle + \mu\langle x, u^*(y')\rangle = \langle x, \lambda u^*(y) + \mu u^*(y)\rangle\end{aligned}$$

Ainsi

$$u^*(\lambda y + \mu y') - \lambda u^*(y) - \mu u^*(y') \in E^{\perp} = 0$$

Ce qui démontre que u^* est linéaire. □

On a alors les propriétés suivantes sur l'adjoint.

Proposition II.1.5. *Soient* $u, v \in \mathrm{End}(E)$ *et* $\lambda, \mu \in \mathbb{R}$. *Alors*

1. $(\lambda u + \mu v)^* = \lambda u^* + \mu v^*$ *;*
2. $(u \circ v)^* = v^* \circ u^*$ *;*
3. $(u^*)^* = u$ *;*
4. $\mathrm{id}_E^* = \mathrm{id}_E$ *;*
5. *si* $u \in \mathrm{GL}(E)$*, alors* $u^* \in \mathrm{GL}(E)$ *et* $(u^{-1})^* = (u^*)^{-1}$.

Démonstration. On va simplement utiliser l'unicité de l'adjoint et la bilinéarité du produit scalaire

$$\begin{aligned}\forall x, y \in E,\ \langle x, (\lambda u + \mu v)^*(y)\rangle = \langle(\lambda u + \mu v)(x), y\rangle &= \lambda\langle u(x), y\rangle + \mu\langle v(x), y\rangle \\ &= \lambda\langle x, u^*(y)\rangle + \mu\langle x, v^*(y)\rangle = \langle x, \lambda u^*(y) + \mu v^*(y)\rangle\end{aligned}$$

Ainsi, par unicité de l'adjoint, on a $(\lambda u + \mu v)^* = \lambda u^* + \mu v^*$.

$$\forall x, y \in E,\ \langle x, (u \circ v)^*(y)\rangle = \langle u \circ v(x), y\rangle = \langle v(x), u^*(y)\rangle = \langle x, v^* \circ u^*(y)\rangle$$

Donc $(u \circ v)^* = v^* \circ u^*$ par unicité de l'adjoint.

$$\forall x, y \in E,\ \langle x, (u^*) * (y)\rangle = \langle u^*(x), y\rangle = \langle y, u^*(x)\rangle = \langle u(y), x\rangle = \langle x, u(y)\rangle$$

D'où $(u^*)^* = u$.

$$\forall x, y \in E,\ \langle x, \mathrm{id}_E^*(y)\rangle = \langle \mathrm{id}_E(x), y\rangle = \langle x, y\rangle = \langle x, \mathrm{id}_E(y)\rangle$$

D'où $\mathrm{id}_E^* = \mathrm{id}_E$.

Pour la dernière assertion, si u est inversible, alors

$$\mathrm{id}_E = \mathrm{id}_E^* = (u \circ u^{-1})^* = (u^{-1})^* \circ u^*$$

Donc $(u^{-1})^*$ est l'inverse de u^*, ce qui dit en particulier que u^* est inversible. □

Comme on est en dimension finie, on peut réinterpréter ces théorèmes abstraits dans le langage matriciel, ce qui pourra nous simplifier certains calculs.

Proposition II.1.6. *Soit $\mathcal{E}$ une base **orthonormée** de E. Soit $A = \mathrm{Mat}_{\mathcal{E}}(u)$ la matrice de l'endomorphisme u exprimée dans la base orthonormée $\mathcal{E}$. Alors*

$$\mathrm{Mat}_{\mathcal{E}}(u^*) = {}^tA.$$

Remarque II.1.7. Attention : la proposition est fausse si la base $\mathcal{E}$ n'est pas orthonormée.

Démonstration. Notons $A = (a_{i,j}) = \mathrm{Mat}_{\mathcal{E}}(u)$ et $B = \mathrm{Mat}_{\mathcal{E}}(u^*) = (b_{i,j})$. Soient $1 \leqslant i, j \leqslant n$. Alors

$$\langle u(e_i), e_j\rangle = \langle \sum_{k=1}^n a_{k,i} e_k, e_j\rangle = \sum_{k=1}^n a_{k,i}\langle e_k, e_j\rangle = a_{j,i}$$

et

$$\langle e_i, u^*(e_j)\rangle = \langle e_i, \sum_{k=1}^n b_{k,j} e_k\rangle = \sum_{k=1}^n b_{k,j}\langle e_i, e_k\rangle = b_{i,j}$$

Comme $\langle u(e_i), e_j\rangle = \langle e_i, u^*(e_j)\rangle$, on a donc $a_{j,i} = b_{i,j}$ et ceci, pour tout i, j. Cela signifie que $B = {}^tA$. □

Exercice 18. Redémontrer la Proposition II.1.5 en utilisant des matrices.

On a un lien entre l'adjoint et l'orthogonalité.

Proposition II.1.8. *Soit $u \in \mathrm{End}(E)$ un endomorphisme et F un sous-espace vectoriel de E stable par u (c'est-à-dire $u(F) \subseteq F$). Alors $F^\perp$ est stable par u^* (i.e. $u^*(F^\perp) \subseteq F^\perp$).*

Démonstration. Soit $y \in F^\perp$. On veut montrer que $u^*(y) \in F^\perp$.

$$\forall z \in F,\ \langle z, u^*(y)\rangle = \langle u(z), y\rangle = 0$$

car $u(z) \in F$ et $y \in F^\perp$. Cela démontre que $u^*(y)$ est orthogonal à tout $z \in F$ donc $u^*(y) \in F^\perp$. □

II.2 Endomorphismes orthogonaux

On a vu précédemment une notion de distance et norme sur un espace euclidien. On va s'intéresser aux endomorphismes qui préservent la norme euclidienne.

II.2.a Définition des endomorphismes orthogonaux

Définition II.2.1. On dit qu'un endomorphisme $u \in \mathrm{End}(E)$ est *orthogonal* s'il préserve la norme, c'est-à-dire que

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

Notation II.2.2. On note $\mathrm{O}(E)$ l'ensemble des endomorphismes orthogonaux de E.

On peut caractériser les endomorphismes orthogonaux comme suit :

Proposition II.2.3. *S'équivalent*

(i) u est orthogonal, i.e. $\forall x \in E,\ \|u(x)\| = \|x\|$;

(ii) u préserve le produit scalaire, i.e. $\forall x, y \in E,\ \langle u(x), u(y)\rangle = \langle x, y\rangle$;

(iii) $u \in \mathrm{GL}(E)$ et $u^* = u^{-1}$;

Démonstration. **(i)** ⇒ **(ii)** Supposons u orthogonal. Alors, d'un côté

$$\forall x, y \in E,\ \|u(x+y)\|^2 = \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$$

D'un autre côté,

$$\forall x, y \in E,\ \|u(x+y)\|^2 = \|u(x)\|^2 + \|u(y)\|^2 + 2\langle u(x), u(y)\rangle = \|x\|^2 + \|y\|^2 + 2\langle u(x), u(y)\rangle$$

On a donc

$$\forall x, y \in E,\ 2\langle x, y\rangle = 2\langle u(x), u(y)\rangle$$

Ainsi u préserve le produit scalaire.

(ii) $\Rightarrow$ (iii) C'est immédiat par adjonction car

$$\forall x, y \in E, \langle x, y\rangle = \langle u(x), u(y)\rangle = \langle x, u^* \circ u(y)\rangle$$

Donc $u^* \circ u = \mathrm{id}_E$ donc $u^* = u^{-1}$.

(iii) $\Rightarrow$ (i) Si $u^* = u^{-1}$,alors

$$\forall x \in E,\ \|u(x)\|^2 = \langle u(x), u(x)\rangle = \langle x, u^* \circ u(x)\rangle = \langle x, \mathrm{id}_E(x)\rangle = \langle x, x\rangle = \|x\|^2.$$

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

Exercice 19. Montrer l'équivalence des conditions précédentes avec

(iv) u^* est orthogonal, i.e. $\forall x \in E,\ \|u^*(x)\| = \|x\|$.

II.2.b Matrices orthogonales

Petit interlude sur les matrices : l'égalité $u^* = u^{-1}$ invite à s'intéresser aux matrices inversibles Ω telles que ${}^t\Omega = \Omega^{-1}$, ce qui revient à :

Définition II.2.4. Une matrice Ω est dite *orthogonale* si

$${}^t\Omega\Omega = \Omega{}^t\Omega = I_n$$

Notation II.2.5. On note $\mathrm{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales.

C'est un sous-groupe (non commutatif si $n \geqslant 3$) de $\mathrm{GL}_n(\mathbb{R})$. Cela signifie qu'on a les propriétés suivantes :

Proposition II.2.6. *1.* $I_n \in \mathrm{O}_n(\mathbb{R})$ *;*

2. $\forall A, B \in \mathrm{O}_n(\mathbb{R}),\ AB \in \mathrm{O}_n(\mathbb{R})$ *;*

3. $\forall A \in \mathrm{O}_n(\mathbb{R}),\ A^{-1} \in \mathrm{O}_n(\mathbb{R})$.

Exercice 20. Le démontrer.

Remarque II.2.7. Si $\Omega \in \mathrm{O}_n(\mathbb{R})$ on sait déjà que $\Omega \in \mathrm{GL}_n(\mathbb{R})$ car $\Omega^{-1} = {}^t\Omega$ par définition. Mais

$$1 = \det(I_n) = \det(\Omega{}^t\Omega) = \det(\Omega)\det({}^t\Omega) = \det(\Omega)^2$$

Donc $\det(\Omega) \in \{-1, 1\}$. Il y a donc deux types de matrices orthogonales, celles de déterminant 1 et celles de déterminant -1. On note alors

$$\mathrm{SO}_n(\mathbb{R}) = \{\Omega \in \mathrm{O}_n(\mathbb{R}),\ \det(\Omega) = 1\} = \mathrm{O}_n(\mathbb{R}) \cap \mathrm{SL}_n(\mathbb{R})$$

et

$$\mathrm{O}_n^-(\mathbb{R}) = \{\Omega \in \mathrm{O}_n(\mathbb{R}),\ \det(\Omega) = -1\}$$

ce qui nous donne une partition

$$\mathrm{O}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}) \sqcup \mathrm{O}_n^-(\mathbb{R})$$

Exercice 21. Montrer que $\mathrm{SO}_n(\mathbb{R})$ est un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Montrer que $\mathrm{O}_n^-(\mathbb{R})$ n'est pas un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

II.2.c Lien avec les bases orthonormées

On va maintenant voir le lien avec les bases orthonormées.

Lemme II.2.8. *Soient $\mathcal{E}$ et $\mathcal{E}'$ deux bases orthonormées de E et P la matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$. Alors $P \in \mathrm{O}_n(\mathbb{R})$.*

Démonstration. Notons $\mathcal{E} = (e_1, \ldots, e_n)$ et $\mathcal{E}' = (e'_1, \ldots, e'_n)$. Notons également Q la matrice de passage de la base $\mathcal{E}'$ à la base $\mathcal{E}$, de sorte que $Q = P^{-1}$.

Comme $\mathcal{E}$ est orthonormée, on sait qu'on peut écrire

$$\forall 1 \leqslant j \leqslant n,\ e'_j = \sum_{i=1}^{n} \langle e'_j, e_i\rangle e_i$$

Le (i, j)-ème coefficient de P est donc $P_{i,j} = \langle e'_j, e_i$.

De même, comme $\mathcal{E}'$ est orthonormée, on sait qu'on peut écrire

$$\forall 1 \leqslant i \leqslant n,\ e_i = \sum_{j=1}^{n} \langle e_i, e'_j \rangle e_j$$

Le (j,i)-ème coefficient de Q est donc $Q_{j,i} = \langle e_i, e'_j \rangle = P_{i,j}$.

On constate ainsi que $Q = {}^tP = P^{-1}$. Donc P et Q sont des matrices orthogonales. □

Puisque les endomorphismes orthogonaux préservent le produit scalaire, les conditions de produit scalaire valant 0 ou 1 pour les bases orthonormées vont être conservées. Les endomorphismes orthogonaux envoient donc les bases orthonormées sur des bases orthonormées. Plus précisément :

Théorème II.2.9. *Soit $u \in \mathrm{End}(E)$ un endomorphisme. S'équivalent :*

(i) u est orthogonal ;

(ii) il existe une base orthonormée $\mathcal{E}$ telle que $\mathrm{Mat}_{\mathcal{E}}(u) \in \mathrm{O}_n(\mathbb{R})$;

(iii) pour toute base orthonormée $\mathcal{E}$, on a $\mathrm{Mat}_{\mathcal{E}}(u) \in \mathrm{O}_n(\mathbb{R})$;

(iv) il existe une base orthonormée $\mathcal{E} = (e_1, \dots, e_n)$ telle que $\big(u(e_1), \dots, u(e_n)\big)$ est une base orthonormée ;

(v) pour toute base orthonormée $\mathcal{E} = (e_1, \dots, e_n)$ la famille $\big(u(e_1), \dots, u(e_n)\big)$ est une base orthonormée.

Démonstration. On va démontrer (i) ⇒ (v) ⇒ (iv) ⇒ (ii) ⇒ (iii) ⇒ (i).

(i) ⇒ (v) : Soit $\mathcal{E} = (e_1, \dots, e_n)$ une base orthonormée. Alors

$$forall 1 \leqslant i, j \leqslant n,\ \langle u(e_i), u(e_j) \rangle = \langle e_i, e_j \rangle = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Donc la famille $\big(u(e_1), \dots, u(e_n)\big)$ est bien une base orthonormée.

(v) ⇒ (iv) est immédiat.

(iv) ⇒ (ii) : La matrice $\mathrm{Mat}_{\mathcal{E}}(u)$ est la matrice dont les colonnes sont les vecteurs $u(e_i)$ écrits dans la base des (e_i). C'est donc la matrice de passage de $(e_1, \dots, e_n)$ vers $\big(u(e_1), \dots, u(e_n)\big)$. D'après le lemme II.2.8, c'est donc une matrice orthogonale.

(ii) ⇒ (iii) : Soit $\mathcal{E}$ une base orthonormée dans laquelle la matrice de u est orthogonale. Soit $\mathcal{E}'$ une autre base orthonormée de E et P la matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$. Alors

$$\mathrm{Mat}_{\mathcal{E}'}(u) = P^{-1} \mathrm{Mat}_{\mathcal{E}}(u) P$$

est le produit de trois matrices dans $\mathrm{O}_n(\mathbb{R})$. C'est donc une matrice dans $\mathrm{O}_n(\mathbb{R})$ qui est stable par multiplication (groupe).

(iii) ⇒ (i) : Soit $\mathcal{E}$ une base orthonormée de E, soit $P \in \mathrm{O}_n(\mathbb{R})$ la matrice de u dans cette base. Soit $x \in E$ et X sa matrice colonne dans la base $\mathcal{E}$. Alors

$$\langle u(x), u(x) \rangle = {}^t(PX)(PX) = {}^tX{}^tPPX = {}^tXX = \langle x, x \rangle$$

Ceci étant vrai pour tout $x \in E$, on en déduit que u est orthogonal. □

Exercice 22. Démontrer les autres implications possibles entre les assertions.

Remarque II.2.10. Si $\mathcal{E}$ est une base orthonormée de $\mathcal{E}$ (n'importe laquelle), d'après le théorème II.2.9, on peut donc définir une fonction

$$\begin{array}{rccc} \Phi : & \mathrm{O}(E) & \to & \mathrm{O}_n(\mathbb{R}) \\ & u & \mapsto & \mathrm{Mat}_{\mathcal{E}}(u) \end{array}$$

C'est en fait une bijection, et même un isomorphisme de groupes.

Cela permet alors de définir naturellement

$$\mathrm{SO}(E) = \Phi^{-1}\big(\mathrm{SO}_n(\mathbb{R})\big) \qquad \text{et} \qquad \mathrm{O}^-(E) = \Phi^{-1}\big(\mathrm{O}_n^-(\mathbb{R})\big)$$

II.2.d Sous-espaces et endomorphismes orthogonaux

Proposition II.2.11. *Soit $u \in \mathrm{O}(E)$. Alors*

1. *Si F est un sous-espace vectoriel de E stable par u, alors $F^\perp$ est stable par u (i.e. $u(F^\perp) \subseteq F^\perp$).*
2. *si $\lambda \in \mathbb{R}$ est une valeur propre de u, alors $\lambda \in \{-1, 1\}$.*
3. *On a $E = \ker(u - \mathrm{id}_E) \overset{\perp}{\oplus} \mathrm{im}(u - \mathrm{id}_E)$.*

4. On a $E = \ker(u + \mathrm{id}_E) \overset{\perp}{\oplus} \mathrm{im}(u + \mathrm{id}_E)$.

Démonstration. Supposons que F est u-stable, i.e. $u(F) \subseteq F$. Comme u est un automorphisme de E, on a donc $u(F) = F$ car u préserve la dimension des sous-espaces. Soit $x \in F^\perp$. Montrons que $u(x) \in F^\perp$. Soit $y \in F = u(F)$. Alors il existe $z \in F$ tel que $y = u(z)$. Donc

$$\langle u(x), y\rangle = \langle u(x), u(z)\rangle = \langle x, z\rangle = 0$$

Ceci étant vrai pour tout $y \in F$, on en déduit que $u(x) \in F^\perp$.

Soit λ une valeur propre de u et x un vecteur propre associé à λ. Alors

$$\|x\| = \|u(x)\| = \|\lambda x\| = |\lambda|\|x\|$$

Comme $x \neq 0$, en divisant par $\|x\| \neq 0$, on en déduit que $|\lambda| = 1$, donc $\lambda = \{-1, 1\}$.

Montrons simultanément les deux dernières assertions. Soit $\varepsilon \in \{-1, 1\}$. Soit $F = \ker(u - \varepsilon\,\mathrm{id}_E)$ et $G = \mathrm{im}(u - \varepsilon\,\mathrm{id}_E)$. On va montrer que $G = F^\perp$. Soit $x \in G$ qu'on écrit donc $x = (u - \varepsilon\,\mathrm{id}_E)(y) = u(y) - \varepsilon y$. Soit $z \in F$. Alors $(u - \varepsilon\,\mathrm{id}_E)(z) = u(z) - \varepsilon z$ donc $u(z) = \varepsilon z$. Ainsi

$$\langle x, z\rangle = \langle u(y) - \varepsilon y, z\rangle = \langle u(y), z\rangle - \varepsilon\langle y, z\rangle = \langle u(y), \varepsilon u(z)\rangle - \varepsilon\langle y, z\rangle = \varepsilon\langle y, z\rangle - \varepsilon\langle y, z\rangle = 0$$

Ceci étant vrai pour tout $z \in F$, on a donc $x \in F^\perp$. Donc $G \subseteq F^\perp$.

Mais, d'après le théorème du rang, on a

$$\dim(E) = \dim\ker(u - \varepsilon\,\mathrm{id}_E) + \mathrm{rg}(u - \varepsilon\,\mathrm{id}_E) = \dim F + \dim(G)$$

Ainsi G et $F^\perp$ sont de même dimension $\dim E - \dim F$ et inclus l'un dans l'autre. Donc $G = F^\perp$. □

II.3 Endomorphismes auto-adjoints

Une autre famille d'endomorphismes remarquable est celle des endomorphismes égaux à leurs adjoints.

Définition II.3.1. On dit qu'un endomorphisme $u \in \mathrm{End}(E)$ est *auto-adjoint* si $u^* = u$.

II.3.a Interprétation matricielle

Comme précédemment, on a une interprétation matricielle favorable en base orthonormée.

Théorème II.3.2. *Soit* $u \in \mathrm{End}(E)$*. S'équivalent :*

(i) u est auto-adjoint ;

(ii) il existe une base orthonormée $\mathcal{E}$ telle que $\mathrm{Mat}_\mathcal{E}(u)$ est symétrique ;

(iii) pour toute base orthonormée $\mathcal{E}$, la matrice $\mathrm{Mat}_\mathcal{E}(u)$ est symétrique.

Remarque II.3.3. Puisque les endomorphismes auto-adjoints sont fortement liés aux matrices symétriques, on les appelle aussi parfois *endomorphismes symétriques*.

Démonstration. **(i) ⇒ (iii) :** Soit $\mathcal{E}$ une base orthonormée. D'après la Proposition II.1.6, on sait que

$$\mathrm{Mat}_\mathcal{E}(u) = \mathrm{Mat}_\mathcal{E}(u^*) = {}^t\mathrm{Mat}_\mathcal{E}(u)$$

Donc $\mathrm{Mat}_\mathcal{E}(u)$ est symétrique.

(iii) ⇒ (ii) est immédiat.

(ii) ⇒ (i) : Il existe une base orthonormée $\mathcal{E}$ telle que $\mathrm{Mat}_\mathcal{E}(u) = {}^t\mathrm{Mat}_\mathcal{E}(u)$. Dans cette base orthonormée, par la Propostion II.1.6, on a donc $\mathrm{Mat}_\mathcal{E}(u) = \mathrm{Mat}_\mathcal{E}(u^*)$. Puisque $\mathcal{E}$ est une base, cela implique que $u = u^*$ est donc auto-adjoint. □

Exemple II.3.4. Attention : là encore, il est très important de ne considérer que des bases orthonormées. Prenons par exemple $E = \mathbb{R}^2$ muni du produit scalaire canonique et $e_1 = \begin{pmatrix}1\\0\end{pmatrix}$ et $e_2 = \begin{pmatrix}1\\1\end{pmatrix}$ qui forme une base (non orthogonale) de $\mathbb{R}^2$. Considérons l'endomorphisme u défini par $u(ae_1 + be_2) = ae_1 - be_2$. Alors dans la base $\mathcal{E} = (e_1, e_2)$ on a $\mathrm{Mat}_\mathcal{E}(u) = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$ qui est diagonale donc symétrique. Pourtant on a $u^*(e_1) = 3e_1 + 2e_2$ et $u^*(e_2) = e_1 + e_2$. On voit donc que $u^* \neq u$ donc u n'est pas auto-adjoint.

Dans la base canonique, on a en fait $\mathrm{Mat}_{\mathcal{E}_{\mathrm{can}}}(u) = \begin{pmatrix}1 & -2\\0 & -1\end{pmatrix}$ qui n'est pas symétrique.

Notation II.3.5. On note alors $\mathrm{Sym}(E)$ l'ensemble des endomorphismes auto-adjoints de E.

On note également $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques réelles de taille $n \times n$.

Exercice 23. En utlisant les propriétés de l'adjoint (proposition II.1.5), montrer que $\mathrm{Sym}(E)$ est un sous-espace vectoriel de $\mathrm{End}(E)$.

Remarque II.3.6. Étant donné une base orthonormée $\mathcal{E}$ de E, on peut alors définir une bijection (qui dépend du choix de $\mathcal{E}$) par

$$\begin{array}{ccc} \mathrm{Sym}(E) & \to & \mathcal{S}_n(\mathbb{R}) \\ u & \mapsto & \mathrm{Mat}_{\mathcal{E}}(u) \end{array}.$$

C'est en fait une application linéaire.

Exercice 24. Montrer que la bijection de la remarque précédente est linéaire et en déduire la dimension de $\mathrm{Sym}(E)$.

Proposition II.3.7. *Soit $u \in \mathrm{Sym}(E)$. Si F est stable par u (i.e. $u(F) \subseteq F$) alors $F^{\perp}$ est stable par u (i.e. $u(F^{\perp}) \subseteq F^{\perp}$).*

Démonstration. D'après la Propostion II.1.8, on sait que $F^{\perp}$ est stable par u^*. Comme $u^* = u$, on a le résultat. $\square$

Lemme II.3.8. *Soit $u \in \mathrm{Sym}(E)$. Soit F un sous-espace u-stable. Alors les endomorphismes induits $u_F \in \mathrm{End}(F)$ et $u_{F^{\perp}} \in \mathrm{End}(F^{\perp})$ sont autoadjoints.*

Démonstration. Soit $\mathcal{E}'$ une base orthonormée de F et $\mathcal{E}''$ une base orthonormée de $F^{\perp}$. Alors $\mathcal{E} = \mathcal{E}' \cup \mathcal{E}''$ est une base orthonormée de E. Dans cette base, on sait que la matrice de u est symétrique et elle s'écrit, par blocs

$$\mathrm{Mat}_{\mathcal{E}}(u) = \left(\begin{array}{c|c} \mathrm{Mat}_{\mathcal{E}'}(u_F) & 0 \\ \hline 0 & \mathrm{Mat}_{\mathcal{E}''}(u_{F^{\perp}}) \end{array}\right)$$

En transposant

$${}^t\mathrm{Mat}_{\mathcal{E}}(u) = \left(\begin{array}{c|c} {}^t\mathrm{Mat}_{\mathcal{E}'}(u_F) & 0 \\ \hline 0 & {}^t\mathrm{Mat}_{\mathcal{E}''}(u_{F^{\perp}}) \end{array}\right)$$

Donc

$${}^t\mathrm{Mat}_{\mathcal{E}'}(u_F) = \mathrm{Mat}_{\mathcal{E}'}(u_F) \qquad \text{et} \qquad {}^t\mathrm{Mat}_{\mathcal{E}''}(u_{F^{\perp}}) = \mathrm{Mat}_{\mathcal{E}''}(u_{F^{\perp}})$$

Les matrices de u_F et $u_{F^{\perp}}$ sont symétriques dont les endomorphismes u_F et $u_{F^{\perp}}$ sont auto-adjoints. $\square$

Exercice 25. Soit F un sous-espace de E et G un supplémentaire de F dans E. Soit p la projection sur F parallèlement à G. Montrer que p est la projection orthogonale sur F (i.e. $G = F^{\perp}$) si, et seulement si, p est auto-adjoint.

II.3.b Théorème spectral

On va maintenant avoir besoin du lemme d'algèbre linéaire générale suivant, valable pour les $\mathbb{R}$-espaces vectoriels.

Lemme II.3.9. *Soit $u \in \mathrm{End}(E)$ (pas nécessairement symétrique). Alors u admet un sous-espace stable de dimension 1 ou 2.*

Démonstration. Si $u = 0$, tout sous-espace vectoriel est stable donc on peut prendre n'importe lequel.

Sinon, soit P le polynôme minimal de u, c'est-à-dire le polynôme unitaire P de plus petit degré tel que $P(u) = 0$ (P est uniquement déterminé par u). Soit $\lambda \in \mathbb{C}$ une racine de P.

Supposons $\lambda \in \mathbb{R}$. Alors P se factorise en $P = (X - \lambda)Q$ avec $Q \in \mathbb{R}[X]$. Il existe $y_0 \in E$ tel que $x_0 := Q(u)(y_0) \neq 0$ car sinon, cela contredirait la minimalité de P. Alors $0 = P(u)(y_0) = (u - \lambda\,\mathrm{id}_E)(Q(u)(y_0)) = u(x_0) - \lambda x_0$. Ainsi, $x_0 \neq 0$ est un vecteur propre de u est $F = \mathrm{Vect}(x_0)$ est un sous-espace u-stable de dimension 1.

Supposons $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Alors $P(\overline{\lambda}) = \overline{P(\lambda)} = 0$ car $P \in \mathbb{R}[X]$. Donc $R = (X - \lambda)(X - \overline{\lambda}) = X^2 - bX - c \in \mathbb{R}[X]$, avec $b = 2\Re(\lambda)$ et $c = -|\lambda|^2$, divise P et P se factorise en $P = RQ$ pour un certain polynôme unitaire $Q \in \mathbb{R}[X]$. Il existe $y_0 \in E$ tel que $x_0 = Q(u)(y_0) \neq 0$ car sinon cela contredirait la minimalité de P. Donc $0 = P(u)(x_0) = R(u)(Q(u)(y_0)) = u^2(x_0) - bu(x_0) - cx_0$. Ainsi $u^2(x_0) \in \mathrm{Vect}(u(x_0), x_0)$. Donc $F = \mathrm{Vect}(u(x_0), x_0)$ est u-stable et de dimension 1 ou 2 convient. $\square$

Lemme II.3.10. *Soit W un espace euclidien de dimension 2 et $u \in \mathrm{Sym}(W)$ un endomorphisme symétrique de W. Alors u admet une valeur propre réelle.*

De plus, u est diagonalisable.

Démonstration. Soit $\mathcal{E}$ une base orthonormée de W et $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ la matrice (symétrique) de u dans la base $\mathcal{E}$. Le polynôme caractéristique de A est

$$\chi_A = (X-a)(X-b) - c^2 = X^2 - (a+b)X + ab - c^2$$

Son discriminant est

$$\Delta = (a+b)^2 + 4c^2 - 4ab = a^2 + 2ab + b^2 - 4ab + 4c^2 = (a-b)^2 + 4c^2 \geqslant 0$$

Donc χ_A admet au moins une racine réelle, elle est double si, et seulement si, $a = b$ et $c = 0$.

Donc u admet soit deux valeurs propres réelles distinctes et donc est diagonalisable, soit u est diagonale dans la base $\mathcal{E}$ et admet une valeur propre double. □

On a vu que u est diagonalisable mais pour l'instant on ne sait rien de la base. En procédant désormais par récurrence, on va montrer

Théorème II.3.11 (Théorème spectral)**.** *Soit $u \in \mathrm{Sym}(E)$. Alors il existe une base **orthonormée** $\mathcal{E}$ telle que* $\mathrm{Mat}\,\mathcal{E}(u)$ *est diagonale.*

Remarque II.3.12. On peut retenir la phrase « un endomorphisme symétrique est diagonalisable en base orthonormée ». Il y a deux informations dans ce théorème très court à énoncer :

— D'une part, les endomorphismes auto-adjoints sont diagonalisables.

— D'autre part, on peut trouver une base de diagonalisation orthonormée.

En général, un endomorphisme diagonalisable ne l'est pas toujours dans une base orthonormé comme l'a montré l'exemple II.3.4.

Démonstration. On procède par récurrence sur $n = \dim(E)$. Si $n = 1$, il suffit de prendre un vecteur de norme 1.

Hérédité : On suppose $n \geqslant 2$. Soit $u \in \mathrm{Sym}(E)$. D'après le lemme II.3.9, u admet une valeur propre réelle donc un vecteur propre x qu'on peut choisir de norme 1. Alors $F = \mathrm{Vect}(x)$ est un sous-espace u-stable. D'après la propostion II.3.7, on sait que $F^\perp$ est un supplémentaire u-stable de F. D'après le lemme II.3.8, l'endomorphisme induit sur $F^\perp$ est également symétrique. Par hypothèse de récurrence, il existe une base orthonormée $\mathcal{E}'$ de $F^\perp$ de vecteurs propres de $u_{F^\perp}$. Donc $\mathcal{E} = \mathcal{E}' \cup \{x\}$ est une base orthonormée formée de vecteurs propres de u. □

Notation II.3.13. Soit $u \in \mathrm{End}(E)$. Rappelons qu'on appelle spectre de u, noté $\mathrm{Sp}_{\mathbb{C}}(u)$ l'ensemble des valeurs propres complexes de u.

Le théorème spectral nous dit, en particulier, que si $u \in \mathrm{Sym}(E)$, alors $\mathrm{Sp}_{\mathbb{C}}(u) \subset \mathbb{R}$.
Dans $\mathbb{R}$, la notion de signe prend sens, contrairement à $\mathbb{C}$.

Définition II.3.14. On dit qu'un endomorphisme auto-adjoint $u \in \mathrm{Sym}(E)$ est :

— *positif* si $\mathrm{Sp}(u) \subset \mathbb{R}_+ = [0, +\infty[$;

— *défini positif* si $\mathrm{Sp}(u) \subset \mathbb{R}_+^* =]0, +\infty[$.

Notation II.3.15. On note $\mathrm{Sym}^+(\mathbb{R})$ l'ensemble des endomorphismes symétriques positifs et $\mathrm{Sym}^{++}(\mathbb{R})$ l'ensemble des endomorphismes symétriques définis positifs.

Remarque II.3.16. Les ensembles d'endomorphismes $\mathrm{Sym}^+(\mathbb{R})$ et $\mathrm{Sym}^{++}(\mathbb{R})$ **ne sont pas** des espaces vectoriels : si u est positif, les valeurs propres de $-u$ sont négatives dont $-u$ n'est pas positif !

Corollaire II.3.17 (Théorème spectral matriciel)**.** *Soit $S \in \mathcal{S}_n(\mathbb{R})$ une matrice symétrique réelle. Alors il existe une matrice $\Omega \in \mathrm{O}_n(\mathbb{R})$ telle que $D = \Omega^{-1} S \Omega$ est diagonale.*

Remarque II.3.18. Dans cette écriture, comme $\Omega \in \mathrm{O}_n(\mathbb{R})$, on a $D = \Omega^{-1} S \Omega = {}^t\Omega S \Omega$. Donc la matrice S est à la fois semblable et congruence à une matrice diagonale, via une matrice orthogonale.

Si, on veut l'écrire dans l'autre sens, $S = \Omega D \Omega^{-1} = \Omega D {}^t\Omega$.

Démonstration. Soit $S \in \mathcal{S}_n(\mathbb{R})$. Soit $\mathcal{E}$ la base canonique de $\mathbb{R}^n$ et $u \in \mathrm{End}(\mathbb{R}^n)$ l'endomorphisme donc la matrice dans la base canonique est S. D'après le théorème spectral, on sait qu'il existe une base orthonormée $\mathcal{E}'$ de vecteurs propres de u. Soit Ω la matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$. La matrice D de u dans la base $\mathcal{E}'$ est donc diagonale, c'est-à-dire que

$$D = \mathrm{Mat}_{\mathcal{E}'}(u) = \Omega^{-1} S \Omega$$

Mais comme Ω est la matrice de passage d'une base orthonormée vers une autre base orthonormée, on sait d'après le lemme II.2.8 que $\Omega \in \mathrm{O}_n(\mathbb{R})$. □

Remarque II.3.19. On dispose alors d'une méthode pour diagonaliser une matrice symétrique S :

étape 1 : Déterminer les valeurs propres $(\lambda_1, \dots, \lambda_d)$ de S (par exemple en factorisant son polynôme caractéristique).

étape 2 : Déterminer les espaces propres $E_{\lambda_i}(S)$.

étape 3 : Pour chaque λ_i, déterminer une base orthonormée $\mathcal{E}_i$ de $E_{\lambda_i}(S)$ en appliquant le procédé de Gram-Schmidt. C'est une base de vecteurs propres associés à la valeur propre λ_i.

étape 4 : La base $\mathcal{E}$, réunion des $\mathcal{E}_i$ est orthonormée. On écrit alors Ω la matrice de passage de la base canonique vers $\mathcal{E}$, dont les colonnes sont les vecteurs de $\mathcal{E}$ écrits dans la base canonique. On a alors $\Omega \in \mathrm{O}_n(\mathbb{R})$ et on écrit D la matrice diagonale des valeurs propres avec multiplicité.

étape 5 : On calcule enfin Ω^{-1}. Contrairement à la diagonalisation usuelle, comme $\Omega^{-1} = {}^t\Omega$, on a juste une transposition à faire pour conclure.

Le point délicat réside dans l'étape 1. Dans la plupart des exercices, on proposera des matrices pour lesquelles il est possible de déterminer les racines du polynôme caractéristique. Cependant, pour une matrice quelconque, on ne peut en général pas déterminer explicitement ses valeurs propres. Il existe des techniques d'approximation des valeurs propres qui pourront être étudiées dans la suite de vos études.

Exemple II.3.20. Considérons $S = \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. On constate que $S \in \mathcal{S}_3(\mathbb{R})$ est symétrique. On calcule son polynôme caractéristique :

$$\chi_S(x) = \begin{vmatrix} 2-x & & \\ 2 & 0 & \\ 2 & -1-x & 0 \\ 0 & 0 & 1-x \end{vmatrix} = [(2-x)(-1-x)-4](1-x) = [x^2-x-6](1-x) = (x-3)(x+2)(1-x)$$

Donc $\chi_S(x)$ admet trois valeurs propres distinctes $\{-2, 1, 3\}$. On a

$$E_{-2}(S) = \ker(S + 2I_3) = \ker \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \mathrm{Vect}(v_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix})$$

$$E_1(S) = \ker(S - I_3) = \ker \begin{pmatrix} 1 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathrm{Vect}(v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix})$$

$$E_1(S) = \ker(S - 3I_3) = \ker \begin{pmatrix} -1 & 2 & 0 \\ 2 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \mathrm{Vect}(v_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix})$$

Ici, on n'a pas encore orthonormalisé les bases de vecteurs propres. Faisons-le.

$$w_1 = \frac{v_1}{\|v_1\|} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \\ 0 \end{pmatrix} \qquad w_2 = \frac{v_2}{\|v_2\|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad w_3 = \frac{v_3}{\|v_3\|} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}$$

On a alors $\Omega = \begin{pmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}$. Ainsi $D = {}^t\Omega S \Omega$, est diagonale, soit

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ 0 & 1 & 0 \end{pmatrix}$$

Exemple II.3.21. Soit $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$. On observe que la matrice A est symétrique et

$$\chi_A(x) = \begin{vmatrix} 1-x & i \\ i & -1-x \end{vmatrix} = (x-1)(x+1) - i^2 = x^2 - 1 + 1 = x^2$$

Donc 0 est valeur propre double de A. Si A est diagonalisable, alors il existe $P \in \mathrm{GL}_2(\mathbb{C})$ telle que $P^{-1}AP = 0$. Donc $A = P0P^{-1} = 0$, ce qui est absurde.

Quel est le problème ?

Ici, on a supposé A diagonalisable et on a abouti a une contradiction. Donc A n'est pas diagonalisable. On voit que le théorème spectral ne s'applique pas aux matrices complexes ! C'est un théorème qui est spécifique aux matrices à coefficients **réels**.

II.3.c Décomposition polaire

À l'aide des espaces euclidiens, on peut obtenir de nombreuses décompositions matricielles qui s'avèrent utiles dans différents problèmes d'analyse en mathématique.

Notation II.3.22. Comme pour les endomorphismes symétriques, on note également $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) l'ensemble des matrices symétriques dont le spectre est positif (resp. strictement positif)

La décomposition polaire va généraliser naturellement la forme trigonométrique des nombres complexes inversibles $z = re^{i\theta}$ où $r \in \mathbb{R}_+^* \simeq \mathcal{S}_1^{++}(\mathbb{R})$ et $e^{i\theta} \in \mathbb{U} \simeq \mathcal{SO}_2(\mathbb{R})$ s'identifie à une rotation de $\mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel de dimension 2.

Proposition II.3.23 (Unicité de la racine carrée)**.** *Toute matrice $M \in \mathcal{S}_n^+(\mathbb{R})$ admet une racine carrée $N \in \mathcal{S}_n^+(\mathbb{R})$, c'est-à-dire $N^2 = M$.*

Si, de plus, $M \in \mathcal{S}_n^{++}(\mathbb{R})$, alors $N \in \mathcal{S}_n^{++}(\mathbb{R})$ est unique.

Démonstration. On écrit $M = P^{-1}DP$ avec $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ diagonale constituée des valeurs propres réelles de M, donc (resp. strictement) positives.

Existence : Posons $E = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. Alors $N = P^{-1} \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})P$ convient.

Unicité : Soit $R \in \mathcal{S}_n^{++}(\mathbb{K})$ telle que $R^2 = M$ une racine carrée de M. Alors R est diagonalisable et commute à M. Soit Q le polynôme interpolateur de Lagrange donné par $Q(\lambda_i) = \sqrt{\lambda_i}$. Alors

$$Q(M) = Q(P^{-1}DP) = P^{-1}Q(D)P = P^{-1}EP = N.$$

Comme R et $Q(R^2) = Q(M) = N$ commutent, elles sont codiagonalisables. Soit $U = R^{-1}N$. Alors $U \in \mathcal{S}_n^{++}(\mathbb{R})$ car ${}^t(R^{-1}N) = {}^tN{}^t(R^{-1}) = NR^{-1} = R^{-1}N$ et les valeurs propres de U sont des produits d'une valeur propre de N et d'une valeur propre de R^{-1} donc strictement positives. De plus $U^2 = R^{-2}N^2 = M^{-1}M = I_n$, donc les valeurs propres de U sont dans $\{\pm 1\}$. Ainsi $U = I_n$ donc $R = N$. □

Théorème II.3.24 (Décomposition polaire)**.** *Pour toute matrice $M \in \mathrm{GL}_n(\mathbb{R})$, il existe un unique couple $O \in \mathcal{O}_n(\mathbb{K})$ et $S \in \mathcal{S}_n^{++}(\mathbb{R})$ tel que $M = OS$.*

Démonstration. **Unicité :** Si $M = OS$, alors ${}^tMM = {}^tS{}^tOOS = {}^tSS = S^2$. Ainsi nécessairement $S = \sqrt{{}^tMM}$ et donc $O = MS^{-1}$.

Existence : Il reste à montrer que $O = MS^{-1}$ est orthogonale.

$$O{}^tO = (MS^{-1}){}^t(MS^{-1}) = MS^{-1}{}^t(S^{-1}){}^tM = MS^{-2}{}^tM = M\left({}^tMM\right)^{-1}{}^tM = I_n$$

□

II.4 Symétries et réflexions orthogonales

Pour finir cette étude des endomorphismes d'un espace euclidien E, on va s'intéresser aux symétries.

II.4.a Symétries vectorielles

Définition II.4.1. Soit F un sous-espace vectoriel de E et G un supplémentaire de F (i.e. $E = F \oplus G$). On appelle *symétrie par rapport à F parallèlement à G* l'application $s : E \to E$ telle que

$$\forall x \in F,\ \forall y \in G,\ s(x+y) = x - y$$

Fait II.4.2. *Si s est la symétrie par rapport à F parallèlement à G, alors*

$$F = \ker(s - \mathrm{id}_E) = \mathrm{im}(s + \mathrm{id}_E)$$

et

$$G = \ker(s + \mathrm{id}_E) = \mathrm{im}(s - \mathrm{id}_E).$$

Démonstration. Soit $y \in F$ et $z \in G$. Alors, d'une part, $(s - \mathrm{id}_E)(y+z) = y - z - y - z = -2z$ Donc

$$\ker(s - \mathrm{id}_E) = \{y + z \mid y \in F,\ z \in G,\ -2z = 0\} = F$$

$$\mathrm{im}(s - \mathrm{id}_E) = \{-2z \mid y \in F,\ z \in G\} = G$$

D'autre part, $(s + \mathrm{id}_E)(y+z) = y - z + y + z = 2y$ Donc

$$\ker(s + \mathrm{id}_E) = \{y + z \mid y \in F,\ z \in G,\ 2y = 0\} = G$$

$$\mathrm{im}(s + \mathrm{id}_E) = \{2y \mid y \in F,\ z \in G\} = F$$

□

Proposition II.4.3. *Soit $s \in \mathrm{End}(E)$. S'équivalent :*

(i) *s est une symétrie ;*

(ii) *$s \circ s = \mathrm{id}_E$;*

(iii) *s est diagonalisable et $\mathrm{Sp}_{\mathbb{C}}(s) \subseteq \{-1, 1\}$.*

Démonstration. **(i) ⇒ (ii) :** Supposons que s est la symétrie par rapport à F parallèlement à G. Soit $y \in F$ et $z \in G$, alors $s \circ s(y+z) = s(y-z) = y+z$. Comme $F + G = E$, on a $s \circ s = \mathrm{id}_E$.

(ii) ⇒ (iii) : Si $s \circ s = \mathrm{id}_E$, alors le polynôme $X^2 - 1$ annule s. Ce polynôme est scindé à racines simples, donc s est diagonalisable et ses valeurs propres sont contenues dans les racines de $X^2 - 1 = (X+1)(X-1)$.

(iii) ⇒ (i) : Si s est diagonalisable et ses valeurs propres sont dans $\{-1, 1\}$. Posons $F = \ker(s - \mathrm{id}_E)$ et $G = \ker(s + \mathrm{id}_E)$. Alors, par le lemme des noyaux, on a $E = F \oplus G$. De plus, si $y \in F$ et $z \in G$, on a $s(y) = y$ et $s(z) = -z$ donc $s(y+z) = y - z$. Ainsi s est la symétrie par rapport à F parallèlement à G. □

Exercice 26. Soient F un sous-espace vectoriel de E et G un supplémentaire de F dans E Soit p la projection sur F parallèlement à G et s la symétrie par rapport à F parallèlement à G. Montrer que $s = 2p - \mathrm{id}_E$.

II.4.b Symétries orthogonales

Proposition II.4.4. *Soit F un sous-espace vectoriel de E et G un supplémentaire de F dans E. Soit s la symétrie par rapport à F parallèlement à G. S'équivalent :*

(i) *$s \in \mathrm{O}(E)$ est orthogonale ;*

(ii) *$s \in \mathrm{Sym}(E)$ est auto-adjoint ;*

(iii) *$G = F^\perp$.*

Définition II.4.5. Lorsque les conditions de la proposition II.4.4 sont vérifiées, on dit que s est *la symétrie orthogonale* par rapport à F et on la note s_F.

De plus, si F est un hyperplan de E (i.e. $\dim F = \dim E - 1$), on appelle s_F la *réflexion orthogonale* par rapport à F.

Si F est de codimension 2 dans E (i.e. $\dim F = \dim E - 2$), on appelle s_F le *retournement orthogonal* par rapport à F.

Démonstration. **(i) ⇒ (iii) :** Si $s \in \mathrm{O}(E)$, montrons que F et G sont orthogonaux. Soit $y \in F$ et $z \in G$. Alors

$$\langle y, z \rangle = \langle s(y), s(z) \rangle = \langle y, -z \rangle = -\langle y, z \rangle.$$

Donc $\langle y, z \rangle = 0$. Ainsi $G \subseteq F^\perp$ et lui est égal par dimension.

(iii) ⇒ (ii) : Supposons $G = F^\perp$. Soit p la projection orthogonale sur F. On a vu que id_E et $p \in \mathrm{Sym}(E)$. Comme $\mathrm{Sym}(E)$ est un espace vectoriel, on en déduit que $s = 2p - \mathrm{id}_E \in \mathrm{Sym}(E)$.

(ii) ⇒ (i) : Si $s \in \mathrm{Sym}(E)$ alors $s^* = s$. Mais $s \circ s = \mathrm{id}_E$ car s est une symétrie. Donc $s^* \circ s = \mathrm{id}_E$. □

II.4.c Réflexions orthogonales

Le cas d'une réflexion par rapport à un hyperplan est souvent plus commode à manipuler. Physiquement, c'est l'image que nous renvoie un miroir par exemple.

Soit $H \subsetneq E$ un hyperplan. Soit $(e_1, \ldots, e_{n-1})$ une base orthonormée de H qu'on complète en une base orthonormée $(e_1, \ldots, e_n)$ de E. On a donc $H^\perp = \mathrm{Vect}(e_n)$.

Fait II.4.6. *Si s_H est la réflexion orthogonale par rapport à H, alors on a*

$$\mathrm{Mat}_{\mathcal{E}}(s_H) = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$$

En particulier, $\det(s_H) = -1$.

Proposition II.4.7. *On suppose que $\dim(E) \geqslant 2$. Soit H un hyperplan de E et $v \in H^\perp \setminus \{0\}$ (de sorte que $H^\perp = \mathrm{Vect}(v)$. Alors , pour tout $x \in E$, on a*

$$s_H(x) = x - 2\frac{\langle x, v \rangle}{\langle v, v \rangle} v.$$

Démonstration. Soit $x \in E$ qu'on écrit $x = h + \lambda v$ avec $h \in H$ et $\lambda \in \mathbb{R}$. Alors

$$s_H(x) = h - \lambda v = x - 2\lambda v$$

Or $\langle x, v\rangle = \langle h + \lambda v, v\rangle = \underbrace{\langle h, v\rangle}_{=0} + \lambda\langle v, v\rangle$. Comme $v \neq 0$, on a $\langle v, v\rangle \neq 0$ donc $\lambda = \frac{\langle x,v\rangle}{\langle v,v\rangle}$. Ce qui démontre la formule souhaitée. □

Proposition II.4.8. *Soient $x, y \in E$. S'équivalent :*

(i) il existe une réflexion orthogonale s_H telle que $s_H(x) = y$;

(ii) $\|x\| = \|y\|$.

Démonstration. Le sens direct est immédiat car une réflexion orthogonale est une isométrie.

Réciproquement, supposons que $\|x\| = \|y\|$.

Si $x = y$**,** alors il suffit de prendre H contenant $x = y$ car alors $s_H(x) = x = y$.

Si $x \neq y$**,** posons $H = \big(\operatorname{Vect}(x-y)\big)^\perp$ qui est un hyperplan de E car de dimension $\dim(E) - 1$ car $x - y \neq 0$. Considérons la réflexion orthogonale s_H par rapport à H et remarquons que $v = x - y \in H^\perp$ par construction.

On a $\langle x + y, x - y\rangle = \|x\|^2 - \|y\|^2 = 0$ donc $x + y \in (H^\perp)^\perp = H$. Ainsi

$$s_H(x + y) = x + y \text{ car } x + y \in H,$$

$$s_H(x - y) = -(x - y) \text{ car } x - y \in H^\perp.$$

Par somme des deux égalités, on en déduit :

$$2s_H(x) = s_H(x + y) + s_H(x - y) = (x + y) - (x - y) = 2y.$$

On a donc bien $s_H(x) = y$. □

La réflexion de la proposition précédente est uniquement déterminée lorsque $x \neq y$, c'est :

Lemme II.4.9. *Soient $x, y \in E$ tels que $x \neq y$. Si H est un hyperplan tel que la réflexion orthogonale par rapport à H vérifie $s_H(x) = y$, alors $H = \big(\operatorname{Vect}(x - y)\big)^\perp$.*

Démonstration. On a $s_H(x) = y$ et $s_H(y) = s_H \circ s_H(x) = x$. Donc $s_H(x - y) = y - x$. Ainsi $x - y \in \ker s_H + \mathrm{id}_E = H^\perp$. Comme $H^\perp$ est de dimension 1, on en déduit $H^\perp = \operatorname{Vect}(x-y)$. D'où le résultat en passant l'égalité précédente à l'orthogonal. □

Théorème II.4.10. *Soit E un espace euclidien. Soit $u \in \mathrm{O}(E)$ et $d = \dim(E) - \dim\ker(u - \mathrm{id}_E) = \operatorname{rg}(u - \mathrm{id}_E)$. Alors, il existe d hyperplans de E $H_1, \ldots, H_d$ tels que $u = s_{H_d} \circ \cdots \circ s_{H_1}$, où, par convention, id_E correspond à la composition de $d = 0$ réflexions.*

On peut retenir de ce théorème par la phrase : « Toute isométrie est un *produit de* réflexions. » bien que l'énoncé du théorème soit un peu plus précis.

Démonstration. On montre par récurrence sur d la propriété :

$$\begin{aligned}\mathcal{P}(d) : &\text{ Pour toute isométrie } u \in \mathcal{O}(E) \text{ telle que } d = \dim(E) - \dim\ker(u - \mathrm{id}_E),\\ &\text{il existe des hyperplans } H_1, \ldots, H_d \text{ tels que } u = s_{H_d} \circ \cdots \circ s_{H_1}.\end{aligned}$$

Si $d = 0$, cela signifie que $u = \mathrm{id}_E$. La propriété $\mathcal{P}(0)$ est donc vraie.

Pour alléger les notations, lorsque $w \in \operatorname{End}(E)$, notons $E_1(w) = \ker w - \mathrm{id}_E$.

Hérédité : Soit $d \geqslant 1$. Supposons $\mathcal{P}(d-1)$ vraie. Soit E un espace euclidien et $u \in \mathcal{O}(E)$ telle que $d = \dim(E) - \dim E_1(u)$. Comme $d \geqslant 1$, il existe $x \notin E_1(u)$ de sorte que $u(x) \neq x$. Comme $u \in \mathrm{O}(E)$ est une isométrie, on a $\|u(x)\| = \|x\|$. L'hyperplan $H_d = \big(\operatorname{Vect}(u(x) - x)\big)^\perp$ vérifie donc $S_{H_d}(x) = u(x)$. Posons $v = s_{H_d} \circ u$.

Tout d'abord, on remarque que $E_1(u) \cap \operatorname{Vect}(x) = 0$ car $x \notin E_1(u)$, donc les espaces $\operatorname{Vect}(x)$ et $E_1(u)$ sont en somme directe. On va montrer par double inclusion que

$$E_1(v) = \operatorname{Vect}(x) \oplus E_1(u).$$

Montrons $\subseteq$. Remarquons que $x \notin H_d$ car $s_{H_d}(x) = u(x) \neq x$. Ainsi $E = \mathrm{Vect}(x) \oplus H_d$. Soit $y \in E_1(v)$ qu'on écrit alors $y = \lambda x + z$ avec $\lambda \in \mathbb{R}$ et $z \in H_d$. Alors

$$\begin{aligned}
\lambda u(x) + u(z) &= u(y) && \text{par linéarité} \\
&= s_{H_d}(v(y)) && \text{par définition de } v \text{ et } s_{H_d}^2 = \mathrm{id}_E \\
&= s_{H_d}(y) && \text{car } v(y) = y \text{ puisque } y \in E_1(v) \\
&= \lambda s_{H_d}(x) + s_{H_d}(z) && \text{par linéarité de } v \\
&= \lambda u(x) + s_{H_d}(z) && \text{car } u(x) = s_{H_d}(x)
\end{aligned}$$

Donc

$$u(z) = s_{H_d}(z) = z$$

car $z \in H_d$. Ainsi $z \in E_1(u)$ donc $y = \lambda x + z \in \mathrm{Vect}(x) + E_1(u)$.

Réciproquement, montrons $\supseteq$. D'une part, on a bien $x \in E_1(v)$ car $v(x) = s_{H_d} \circ u(x) = s_{H_d}^2(x) = x$. D'autre part, soit $y \in E_1(u)$. Montrons que $y \in H_d$. Alors

$$\begin{aligned}
\langle x - u(x), y\rangle &= \langle x, y\rangle - \langle u(x), y\rangle && \text{par linéarité à gauche} \\
&= \langle x, y\rangle - \langle u(x), u(y)\rangle && \text{car } y \in E_1(u) \\
&= \langle x, y\rangle - \langle x, y\rangle && \text{car } u \in \mathrm{O}(E) \text{ préserve le produit scalaire} \\
&= 0
\end{aligned}$$

Ainsi $y \in \mathrm{Vect}\left(x - u(x)\right)^{\perp} = H_d$. Donc

$$v(y) = s_{H_d} \circ u(y) = s_{H_d}(y) = y.$$

Ainsi $E_1(u) \subseteq E_1(v)$ et $x \in E_1(v)$ donc $E_1(u) + \mathrm{Vect}(x) \subseteq E_1(v)$.

D'où l'égalité $E_1(v) = E_1(u) \oplus \mathrm{Vect}(x)$. Mais alors, $\dim E - \dim E_1(v) = \dim E - \dim E_1(u) - 1 = d - 1$. Par hypothèse de récurrence appliquée à v, il existe des hyperplans de E tels que $v = s_{H_{d-1}} \circ \cdots \circ s_{H_1}$. Donc $u = s_{H_d} \circ v = s_{H_d} \circ s_{H_{d-1}} \circ \cdots \circ s_{H_1}$. Ce qui termine l'hérédité. □

En un certain sens, ce théorème est optimal, car u ne peut pas être le produit de moins de réflexions que d :

Exercice 27. Soit $u \in \mathrm{O}(E)$ et $d = \dim E - \dim \ker(u - \mathrm{id}_E)$. On suppose que $d > 0$. Montrer que si $e < d$, $e \in \mathbb{N}$, et si $H_1, \ldots, H_e$ sont des hyperplans de E, alors

$$u \neq s_{H_e} \circ \cdots \circ s_{H_1}.$$

Indication : on pourra estimer la dimension de $\ker(s_{H_k} \circ \cdots \circ s_{H_1})$ *pour* $0 \leqslant k \leqslant e$.

Chapitre III

Isométries en dimension 2 et 3 et orientation

Dans le chapitre précédent, on s'est intéressé à certains endomorphismes remarquables d'un espace euclidien du point de vue de l'algèbre linéaire. Désormais, on va tenter d'adopter un point de vue plus géométrique. On va principalement se concentrer sur les cas d'un espace euclidien de dimension 2 ou 3.

Rappelons que si $u \in \mathrm{O}(E)$ et si $\mathcal{E}$ est une base orthonormée de E, on a $\det\big(\mathrm{Mat}_{\mathcal{E}}(u)\big) = \pm 1$, et cette valeur ne dépend pas du choix de u.

Notation III.0.1. On note alors $\det(u) = \det\big(\mathrm{Mat}_{\mathcal{E}}(u)\big)$.

Définition III.0.2. On appelle *rotation vectorielle* un élément de $\mathrm{SO}(E) = \{u \in \mathrm{O}(E),\ \det(u) = 1\}$.

On appelle *antirotation vectorielle* un élément de $\mathrm{O}^-(E) = \{u \in \mathrm{O}(E),\ \det(u) = -1\}$.

III.1 Isométries vectorielles de dimension 2

Soit $\mathcal{E} = (e_1, e_2)$ une base orthonormée de E. On a alors une bijection (c'est même un isomorphisme) donnée par :

$$\begin{array}{ccc} \mathrm{O}(E) & \to & \mathrm{O}_2(\mathbb{R}) \\ u & \mapsto & \mathrm{Mat}_{\mathcal{E}}(u) \end{array}$$

qui induit également des bijections $\mathrm{SO}(E) \to \mathrm{SO}_2(\mathbb{R})$ et $\mathrm{O}^-(E) \to \mathrm{O}_2^-(\mathbb{R})$. On peut donc directement étudier les matrices carrées de $\mathrm{O}_2(\mathbb{R})$.

III.1.a Étude des rotations vectorielles du plan

Soit $\Omega = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R})$. On sait que $\Omega^{-1} = {}^t\Omega$. Donc

$$\frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Or $\det(\Omega) = ad - bd = 1$ nous donne donc le système

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

Ainsi, il existe $\theta \in \mathbb{R}$ (unique à un élément de $2\pi\mathbb{Z}$ près tel que $a = \cos\theta$ et $c = \sin\theta$. Ainsi

$$\Omega = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Notation III.1.1. Pour $\theta \in \mathbb{R}$, on note $\Omega(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ et on l'appelle *rotation d'angle θ.*

Proposition III.1.2. *Soit $\alpha, \beta \in \mathbb{R}$. Alors*

1. $\Omega(\alpha + \beta) = \Omega(\alpha) \cdot \Omega(\beta)$;

2. $\Omega(\alpha)^{-1} = \Omega(-\alpha)$.

Démonstration.

$$\Omega(\alpha)\cdot\Omega(\beta) = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}$$
$$= \begin{pmatrix}\cos\alpha\cos\beta-\sin\alpha\sin\beta & -\cos\alpha\sin\beta-\cos\beta\sin\alpha\\ \sin\alpha\cos\beta+\cos\alpha\sin\beta & -\sin\alpha\sin\beta+\cos\alpha\cos\beta\end{pmatrix} = \begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix} = \Omega(\alpha+\beta)$$

Comme $\Omega(0) = I_2$, on a en particulier que

$$I_2 = \Omega(0) = \Omega(\alpha-\alpha) = \Omega(\alpha)\Omega(-\alpha).$$

Donc $\Omega(-\alpha)$ est l'inverse de $\Omega(\alpha)$. □

Théorème III.1.3. *Le groupe* $\mathrm{SO}_2(\mathbb{R})$ *est commutatif.*

Démonstration. Si $\Omega(\alpha), \Omega(\beta)$ sont deux éléments quelconques de $\mathrm{SO}_2(\mathbb{R})$, alors

$$\Omega(\alpha)\Omega(\beta) = \Omega(\alpha+\beta) = \Omega(\beta+\alpha) = \Omega(\beta)\Omega(\alpha).$$

□

III.1.b Étude des antirotations vectorielles du plan

Soit $\Lambda = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{O}_2^-(\mathbb{R})$. On sait que $\Lambda^{-1} = {}^t\Lambda$. Donc

$$\frac{1}{ad-bc}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix} = \begin{pmatrix}a & c\\ b & d\end{pmatrix}$$

Or $\det(\Omega) = ad - bd = -1$ nous donne donc le système

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

Ainsi, il existe $\varphi \in \mathbb{R}$ (unique à un élément de $2\pi\mathbb{Z}$ près tel que $a = \cos\varphi$ et $c = \sin\varphi$. Ainsi

$$\Lambda = \begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}$$

Notation III.1.4. Pour $\varphi \in \mathbb{R}$, on note $\Lambda(\varphi) = \begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}$.

Ainsi

$$\mathrm{O}_2^-(\mathbb{R}) = \{\Lambda(\varphi),\ \varphi\in\mathbb{R}\}$$

Proposition III.1.5. *Soit $\mathcal{E} = (e_1, e_2)$ une base orthonormée d'un espace euclidien de dimension 2 et $u \in \mathrm{O}_2^-(E)$. Si $\mathrm{Mat}_{\mathcal{E}}(u) = \Lambda(\varphi)$, avec $\varphi\in\mathbb{R}$, alors u est la réflexion orthogonale par rapport à la droite $\mathrm{Vect}(x_\varphi)$ où $x_\varphi = \cos\frac{\varphi}{2}e_1 + \sin\frac{\varphi}{2}e_2$.*

Démonstration. On pose également $y_\varphi = \sin\frac{\varphi}{2}e_1 - \cos\frac{\varphi}{2}e_2$. Alors

$$\Lambda(\varphi)\,\mathrm{Mat}(x_\varphi) = \begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}\begin{pmatrix}\cos\frac{\varphi}{2}\\ \sin\frac{\varphi}{2}\end{pmatrix} = \begin{pmatrix}\cos\varphi\cos\frac{\varphi}{2}+\sin\varphi\sin\frac{\varphi}{2}\\ \sin\varphi\cos\frac{\varphi}{2}-\cos\varphi\sin\frac{\varphi}{2}\end{pmatrix} = \begin{pmatrix}\cos(\varphi-\frac{\varphi}{2})\\ \sin(\varphi-\frac{\varphi}{2})\end{pmatrix} = \begin{pmatrix}\cos\frac{\varphi}{2}\\ \sin\frac{\varphi}{2}\end{pmatrix}$$

Ainsi $u(x_\varphi) = x_\varphi$.

On montre par un calcul similaire que $u(y_\varphi) = -y_\varphi$.

Donc $\mathrm{Mat}_{(x_\varphi,y_\varphi)} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, autrement dit, u est la réflexion orthogonale par rapport à $\mathrm{Vect}(x_\varphi)$. □

Exercice 28. Soient $\theta, \varphi, \varphi_1, \varphi_2 \in \mathbb{R}$. Montrer les égalités :

1. $\Lambda(\varphi)\Omega(\theta) = \Lambda(\varphi-\theta)$;
2. $\Omega(\theta)\Lambda(\varphi) = \Lambda(\varphi+\theta)$;
3. $\Lambda(\varphi_1)\Lambda(\varphi_2) = \Omega(\varphi_1-\varphi_2)$;
4. $\Lambda(\varphi)^{-1}\Omega(\theta)\Lambda(\varphi) = \Omega(-\theta)$.

Illustrer par des figures les situations correspondant à ces quatre égalités.

III.2 Orientation

Exemple III.2.1. Soit E un espace euclidien de dimension 2. Soit $\mathcal{E} = (e_1, e_2)$ une base orthonormée de E. Soit $\varphi \in \mathbb{R}$. Posons $f_1 = \cos(\varphi)e_1 + \sin(\varphi)e_2$ et $f_2 = -\sin(\varphi)e_1 + \cos(\varphi)e_2$ de sorte que $\mathcal{F} = (f_1, f_2)$ est une base orthonormée de E. Posons également $g_1 = \cos(\varphi)e_1 + \sin(\varphi)e_2 = f_1$ et $g_2 = \sin(\varphi)e_1 - \cos(\varphi)e_2 = -f_2$ de sorte que $\mathcal{G} = (g_1, g_2)$ est aussi une base orthonormée.

Soit $\theta \in \mathbb{R}$ et $u \in \mathrm{SO}(E)$ l'isométrie telle que $\mathrm{Mat}_{\mathcal{E}}(u) = \Omega(\theta)$. Demandons-nous ce que devient la matrice de u dans les bases $\mathcal{F}$ et $\mathcal{G}$.

Soit P la matrice de passage de $\mathcal{E}$ à $\mathcal{F}$. On a

$$P = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = \Omega(\varphi).$$

Ainsi, on a

$$\mathrm{Mat}_{\mathcal{F}}(u) = P^{-1}\,\mathrm{Mat}_{\mathcal{E}}(u)P = \Omega(-\varphi)\Omega(\theta)\Omega(\varphi) = \Omega(\theta)$$

Soit Q la matrice de passage de $\mathcal{E}$ à $\mathcal{G}$. On a

$$Q = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \Lambda(\varphi).$$

Ainsi, on a

$$\mathrm{Mat}_{\mathcal{G}}(u) = Q^{-1}\,\mathrm{Mat}_{\mathcal{E}}(u)Q = \Lambda(\varphi)^{-1}\Omega(\theta)\Lambda(\varphi) = \Omega(-\theta)$$

C'est embêtant parce qu'on aimerait dire que u est la rotation d'angle θ dans les bases $\mathcal{E}$ et $\mathcal{F}$ mais qu'elle est la rotation d'angle $-\theta$ dans la base $\mathcal{G}$.

Quel est le problème ici ?

On observe que $\det(P) = 1$ mais $\det(Q) = -1$, il y a une symétrie entre $\mathcal{F}$ et $\mathcal{G}$ qui a changé f_2 en $g_2 = -f_2$, c'est comme si on avait regardé ce que fait u « de l'autre côté du miroir ».

Il semble donc que l'écriture de la matrice d'une rotation $u \in \mathcal{O}(E)$ dépende, au signe près, du choix de la base. On observe, qu'en fait, seul le signe de θ dépend du choix de cette base car le paramètre φ fait que les bases orthonormées de E ont toutes été décrites.

Pour parler de l'angle d'une rotation dans un espace de dimension 2, on a donc besoin de décider qu'une base orthonormée arbitraire est « mieux » que les autres, en ce sens qu'on se l'est donnée. Cela nous dit alors que la « moitié » des bases permettront d'écrire les rotations de la même manière et l'autre « moitié » en renverseront le signe de l'angle. On est alors amené à poser la définition suivante :

Définition III.2.2 (Orientation)**.** Soit E un espace euclidien de dimension $n \geqslant 1$ (pas nécessairement $n = 2$). On dit que deux bases orthonormées $\mathcal{E}$ et $\mathcal{F}$ définissent la même *orientation* si la matrice de passage P de $\mathcal{E}$ à $\mathcal{F}$ vérifie $P \in \mathrm{SO}_n(\mathbb{R})$. Sinon, on a alors $P \in \mathrm{O}_n^-(\mathbb{R})$ et on dit alors que $\mathcal{E}$ et $\mathcal{F}$ définissent des orientations opposées.

Exercice 29. Sur l'ensemble des bases orthonormée $\mathcal{B}$ d'un espace euclidien de dimension finie $n \geqslant 1$, on considère la relation binaire $\mathscr{R}$ donnée par :

$$\mathcal{E}\mathscr{R}\mathcal{F} \text{ si } \mathcal{E} \text{ et } \mathcal{F} \text{ définissent la même orientation.}$$

1. Montrer que $\mathscr{R}$ est une relation d'équivalence sur $\mathcal{B}$.
2. Montrer qu'il y a exactement deux classes d'équivalents pour cette relation, i.e. $\mathrm{Card}\,(\mathcal{B}/\mathscr{R}) = 2$.

Définition III.2.3 (Bases orthonormées directes et indirectes)**.** Si $\mathcal{E}$ est une base orthonormée de E, on dit alors que $\big(E, \langle\cdot,\cdot\rangle, \mathcal{E}\big)$ est un *espace euclidien orienté*.

Dans un espace euclidien orienté $\big(E, \langle\cdot,\cdot\rangle, \mathcal{E}\big)$, si $\mathcal{F}$ est une base orthonormée, on dit que c'est :

— une *base orthonormée directe* (BOND) si $\mathcal{F}$ définit la même orientation que $\mathcal{E}$;

— une *base orthonormée indirecte* (BONI) si $\mathcal{F}$ définit une orientation opposée à $\mathcal{E}$.

Remarque III.2.4. En choisissant $\mathcal{E}$ une base orthonormée, on a choisi un représentant d'une classe d'équivalence $\mathrm{cl}_{\mathscr{R}}(\mathcal{E})$. L'orientation de E choisie ne dépend ainsi que de la classe d'équivalence de $\mathcal{E}$.

On n'oubliera pas que dans une famille/base de vecteurs, l'ordre dans lequel on les écrit est important. Échanger deux vecteurs consécutifs d'une base orthonormée directe en fait une base orthonormée indirecte par exemple.

Définition III.2.5. Soit $(E, \langle\cdot,\cdot\rangle, \mathcal{E})$ un espace euclidien orienté de dimension 2.

Si $u \in \mathrm{SO}(E)$ et $\theta \in \mathbb{R}$, on dit que u est la *rotation d'angle* θ lorsque dans n'importe quelle base orthonormée directe $\mathcal{F}$, on a

$$\mathrm{Mat}_{\mathcal{F}}(u) = \Omega(\theta) \in \mathrm{SO}_2(\mathbb{R})$$

Remarque III.2.6. On a vu que cette définition ne dépend donc pas de la base orthonormée directe $\mathcal{F}$ choisie mais ça dépend de l'orientation de E.

Si on renverse l'orientation de E, on observe que u devient la rotation d'angle $-\theta \mod 2\pi$.

On fera donc attention à ne jamais parler d'angle d'une rotation sans avoir au préalable défini une orientation sur E.

III.3 Notion d'angle en dimension 2

Soit E un espace euclidien de dimension 2 et $x, y \in E$ deux vecteurs **non nuls**. On souhaiterait définir l'angle entre x et y.

Tout d'abord, avec le théorème de Cauchy-Schwarz, on sait que

$$C(x,y) = \frac{\langle x, y\rangle}{\|x\|\|y\|} \in [-1, 1]$$

Par conséquent :

Définition III.3.1 (Angle non orienté)**.** Si E est un espace euclidien (non orienté) de dimension 2 et si $x, y \in E$ sont deux vecteurs **non nuls**, alorsl existe un unique $\theta(x,y) = \arccos C(x,y) \in [0, \pi]$ de telle sorte que

$$\langle x, y\rangle = \|x\|\|y\| \cos\big(\theta(x,y)\big)$$

On appelle $\theta(x,y)$ *l'angle non orienté* entre x et y.

Remarque III.3.2. Pour tout $x, y \in E$ non nuls, on a

$$\theta(x,y) = \theta(y,x)$$

par symétrie du produit scalaire.

Autrement dit, l'angle non orienté ne dépend pas de l'ordre dans lequel on considère x et y.

D'autre part, posons $x' = \frac{x}{\|x\|}$ et $y' = \frac{y}{\|y\|}$, ce qui est possible car on a supposé x et y non nuls.

Proposition III.3.3. *Soit E un espace euclidien de dimension 2 et $x', y' \in E$ tels que $\|x'\| = \|y'\| \neq 0$. Alors il existe une unique rotation $u \in \mathrm{SO}(E)$ telle que $u(x') = y'$.*

Démonstration. **Pour l'existence,** on considère $z' \in E$ tel que $\|z'\| = \|x'\| = \|y'\|$. D'après la Proposition II.4.8, il existe donc $s_1, s_2 \in \mathcal{O}^-(E)$ deux réflexions orthogonales telles que $s_1(x') = z'$ et $s_2(z') = y'$. Donc $u = s_2 \circ s_1 \in \mathrm{SO}(E)$ vérifie $u(x') = y'$.

Pour l'unicité, munissons E d'une orientation. Si $u, v \in \mathrm{SO}(E)$ sont telles que $u(x') = y' = v(x')$. Alors $x' = u^{-1}(v(x))$. Posons $w = u^{-1} \circ v \in \mathrm{SO}(E)$. Alors $x' \in \ker(w - \mathrm{id}_E)$. Soit θ l'angle de w dans une base orthonormée directe de E. Alors le polynôme caractéristique de w vérifie

$$\chi_w(X) = \begin{vmatrix} \cos\theta - X & -\sin\theta \\ \sin\theta & \cos\theta - X \end{vmatrix} = (X^2 - 2\cos\theta X + \cos\theta^2) - (-(\sin\theta)^2) = X^2 - 2\cos\theta X + 1$$

Comme $x' \neq 0$ est un vecteur propre de w de valeur propre 1, on a $\chi_w(1) = 2(1 - \cos\theta) = 0$. Donc $\cos\theta = 1$, ce qui impose que $\theta \equiv 0 \mod 2\pi$. Ainsi $w = \mathrm{id}_E$ donc $u = v$. ☐

Définition III.3.4. Si E est un espace euclidien **orienté** de dimension 2 et si $x, y \in E$ sont deux vecteurs **non nuls**, alors il existe une unique rotation $u \in \mathrm{SO}(E)$ telle que $u(x') = y'$ où $x' = \frac{x}{\|x\|}$ et $y' = \frac{y}{\|y\|}$. Dans une base orthonormée directe $\mathcal{E}$, il existe $\tau = \tau(x,y) \in \mathbb{R}$ tel que $\mathrm{Mat}_{\mathcal{E}}(u) = \begin{pmatrix} \cos\tau & -\sin\tau \\ \cos\tau & \sin\tau \end{pmatrix} = \Omega(\tau)$. Le réel $\tau(x,y)$ s'appelle *l'angle orienté* entre x et y. Il est uniquement déterminé modulo 2π.

Remarque III.3.5. Si $u(x') = y'$, alors $u^{-1}(y') = x'$. Or, on a vu que si u est la rotation d'angle τ, alors u^{-1} est la rotation d'angle $-\tau$. On a donc

$$\tau(x,y) \equiv -\tau(y,x) \mod 2\pi$$

Autrement dit, quand on échange l'ordre de x et y, cela renverse le signe de l'angle orienté.

Remarque III.3.6. Il y a un lien entre l'angle non orienté et l'angle orienté entre x et y.

Soit $x' = e_1$ qu'on complète en une base orthonormée directe (e_1, e_2) de E. Alors on a

$$y' = u(e_1) = \cos\tau(x,y)e_1 + \sin\tau(x,y)e_2.$$

Quand on calcule le produit scalaire entre x et y, on a donc

$$\langle x,y\rangle = \langle \|x\|e_1, \|y\|y'\rangle = \|x\|\|y\|\langle e_1, \cos\tau(x,y)e_1 + \sin\tau(x,y)e_2\rangle = \|x\|\|y\|\cos\tau(x,y)$$

Ainsi, on a

$$\cos\theta(x,y) = \cos\tau(x,y)$$

Donc, lorsqu'on choisit $\tau(x,y)\in[-\pi,\pi]$, on a

$$\theta(x,y) = |\tau(x,y)|$$

Par ailleurs, en écrivant la matrice de la famille (x,y) (qui n'est pas une base en général) dans la base orthonormée directe (e_1,e_2), c'est-à-dire qu'on écrit en colonne les matrices des vecteurs dans cette base, on a

$$\det_{(e_1,e_2)}(x,y) = \begin{vmatrix} \|x\| & \cos\tau(x,y)\|y\| \\ 0 & \sin\tau(x,y)\|y\| \end{vmatrix} = \sin\tau(x,y)\|x\|\|y\|$$

Ce déterminant ne dépend pas de la base orthonormée directe choisie.

Ainsi, pour déterminer l'angle orienté $\tau(x,y)$, il suffit de déterminer

— l'angle non orienté $\theta(x,y)\in[0,\pi]$ et

— le signe $\varepsilon(x,y)$ de $\det(x,y)$ dans une base orthonormée directe (avec $\varepsilon(x,y)=1$ lorsque ce déterminant est nul, mais ce choix est sans importance).

On a alors

$$\tau(x,y) = \varepsilon(x,y)\theta(x,y) \in]-\pi,\pi].$$

III.4 Structure supplémentaire en dimension 3

Dans le reste de cette section, $(E,\langle\cdot,\cdot\rangle,\mathcal{E}_0)$ est un espace euclidien orienté de dimension 3.

On va ici développer des outils de calculs propres à la dimension 3.

III.4.a Produit mixte

Proposition-définition III.4.1 (Produit mixte)**.** Soit $\mathcal{E}$ une base orthonormée directe de E et $\mathcal{F}=(x,y,z)$ une famille de 3 vecteurs de E (l'ordre de x,y,z est important !). Alors le déterminant $\det_{\mathcal{E}}(\mathcal{F})$ de la famille $\mathcal{F}$ dans la base $\mathcal{E}$ est indépendant de la base orthonormée directe $\mathcal{E}$ choisie (parmi les bases orthonormées **directes**).

On note $[x,y,z]=\det_{\mathcal{E}}(\mathcal{F})$ cette quantité et on l'appelle le *produit mixte* de $\mathcal{F}$.

Démonstration. Il faut seulement montrer que $\det_{\mathcal{E}}(\mathcal{F})$ ne dépend pas de la base orthonormée directe $\mathcal{E}$ choisie.

Soit $\mathcal{E}'$ une autre base orthonormée directe. Notons M la matrice de $\mathcal{F}$ dans la base $\mathcal{E}$ et M' la matrice de $\mathcal{F}$ dans la base $\mathcal{E}'$. Soit P la matrice de passage de la base $\mathcal{E}$ vers la base $\mathcal{E}'$. Alors $M=PM'$ et $P\in SO_3(\mathbb{R})$. Donc

$$\det(M) = \det(PM') = \det(M')$$

□

Exemple III.4.2. Soit $\mathcal{E}=(e_1,e_2,e_3)$ une base orthonormée directe de E. Posons $\mathcal{E}'=(e_1',e_2',e_3')$ la base orthonormée suivante :

$$\begin{cases} e_1' = \frac{1}{\sqrt{2}}(e_1+e_3) \\ e_2' = e_2 \\ e_3' = \frac{1}{\sqrt{2}}(e_1-e_3) \end{cases}$$

Considérons la famille $\mathcal{F}=(f_1,f_2,f_3)$ donnée par

$$\begin{cases} f_1 &= \sqrt{2}f_1 + f_2 &= e_1+e_2+e_3 \\ f_2 &= 2\sqrt{2}f_1 &= 2(e_1+e_3) \\ f_3 &= \sqrt{2}f_3 &= e_1-e_3 \end{cases}$$

On a alors

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & -1 \end{pmatrix} \qquad \text{et} \qquad M' = \begin{pmatrix} \sqrt{2} & 2\sqrt{2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

Ainsi

$$\det(M) = 2-(-2) = 4 \qquad \text{et} \qquad \det(M') = \sqrt{2}(-2\sqrt{2}) = -4$$

Comme ces déterminants sont différents, on comprend que $\mathcal{E}$ et $\mathcal{E}'$ n'ont pas la même orientation. Comme $\mathcal{E}$ est directe, cela nous dit donc que $\mathcal{E}'$ est indirecte. On a également

$$[f_1, f_2, f_3] = \det(M) = 4 = -\det(M')$$

Proposition III.4.3 (Propriétés du produit mixte)**.** *Soit $\mathcal{F} = (x, y, z)$ une famille de trois vecteurs de E.*

1. *Si $\mathcal{E}'$ est une base orthonormée indirecte, alors $[x, y, z] = -\det_{\mathcal{E}'}(x, y, z)$.*
2. $[x, y, z] = -[y, x, z] = -[x, z, y] = -[z, y, x] = [z, x, y] = [y, z, x]$.
3. *Les trois applications*

$$\left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ t & \mapsto & [t, y, z] \end{array}\right. \qquad \left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ t & \mapsto & [x, t, z] \end{array}\right. \qquad \left\{\begin{array}{ccc} E & \to & \mathbb{R} \\ t & \mapsto & [x, y, t] \end{array}\right.$$

   sont linéaires.

4. $[x, y, z] \neq 0$ *si, et seulement si, la famille (x, y, z) est libre (donc une base de E).*
5. *Si (x, y, z) est une base orthonormée, alors*

$$\begin{cases} [x, y, z] = 1 & \text{si } (x, y, z) \text{ est directe,} \\ [x, y, z] = -1 & \text{si } (x, y, z) \text{ est indirecte.} \end{cases}$$

Exercice 30. Démontrer cette proposition.
Indication : utiliser les propriétés du déterminant.

III.4.b Produit vectoriel

Définition et théorème III.4.4 (Produit vectoriel)**.** Soient $x, y \in E$ deux vecteurs. Alors il existe un unique $z \in E$ tel que

$$\forall v \in E,\ [x, y, v] = \langle z, v \rangle$$

On dit que $z \in E$ est le *produit vectoriel* de x et y, et on le note $x \wedge y = z$.

Remarque III.4.5. On peut donc retenir la formule

$$\forall x, y, z \in E,\ [x, y, z] = \langle x \wedge y, z \rangle$$

Démonstration. L'application

$$f : \begin{array}{ccc} E & \to & \mathbb{R} \\ v & \mapsto & [x, y, v] \end{array}$$

est une forme linéaire. D'après le théorème de dualité II.1.1, il existe donc un unique $z \in E$ tel que $\forall v \in E,\ f(v) = \langle z, v \rangle$. □

Remarque III.4.6 (Calcul explicite du produit vectoriel). Soit $\mathcal{E} = (e_1, e_2, e_3)$ une base orthonormée directe. On écrit $x = x_1e_1 + x_2e_2 + x_3e_3$ et $y = y_1e_1 + y_2e_2 + y_3e_3$. Alors, en développant par rapport à la troisième colonne, on a

$$[x, y, e_1] = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 0 \\ x_3 & y_3 & 0 \end{vmatrix} = x_2y_3 - x_3y_2$$

$$[x, y, e_2] = \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 0 \end{vmatrix} = x_3y_1 - x_1y_3$$

$$[x, y, e_3] = \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & 0 \\ x_3 & y_3 & 1 \end{vmatrix} = x_1y_2 - x_2y_1$$

Ainsi

$$\mathrm{Mat}_{(e_1,e_2,e_3)}(x \wedge y) = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

En pratique, pour faire le calcul, on peut compléter la matrice 3×2 de x et y écrit en colonne en une matrice 5×2 en recopiant les deux premières lignes et en calculant les trois mineurs du milieu dans l'ordre :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 & & y_1 \\ x_2 & & y_2 \\ & \times & \\ x_3 & & y_3 \\ & \times & \\ x_1 & & y_1 \\ & \times & \\ x_2 & & y_2 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} = \mathrm{Mat}_{(e_1,e_2,e_3)}(x \wedge y)$$

Proposition III.4.7. *Soient $x, y \in E$. Alors*

1. $x \wedge y = y \wedge x$;
2. *L'application* $\begin{array}{ccc} E & \to & E \\ v & \mapsto & x \wedge v \end{array}$ *est linéaire.*
3. *L'application* $\begin{array}{ccc} E & \to & E \\ v & \mapsto & v \wedge y \end{array}$ *est linéaire.*

Exercice 31. Démontrer cette proposition.

Proposition III.4.8. *Soient $x, y \in E$. Alors*

1. $x \wedge y = 0 \Leftrightarrow x$ *et* y *sont colinéaires.*
2. $x \wedge y$ *est orthogonal à* x *et à* y.
3. *si* $\langle x, y \rangle = 0$ *et* $\|x\| = \|y\| = 1$, *alors* $(x, y, x \wedge y)$ *est une base orthonormée **directe**.*

Démonstration. Montrons 1.

Si x et y sont colinéaires. Alors $\det_{\mathcal{E}}(x, y, z) = [x, y, z] = 0$ pour tout $z \in E$. En particulier $[x, y, x \wedge y] = \|x \wedge y\|^2 = 0$. Donc $x \wedge y = 0$.

Sinon, c'est que (x, y) est libre. Alors il existe $z \in E$ tel que (x, y, z) est une base de E. Donc

$$0 \neq \det_{\mathcal{E}}(x, y, z) = [x, y, z] = \langle x \wedge y, z \rangle.$$

Donc $x \wedge y \neq 0$.

Montrons 2. On a

$$\langle x \wedge y, x \rangle = [x, y, x] = \det_{\mathcal{E}}(x, y, x) = 0.$$

De même

$$\langle x \wedge y, y \rangle = [x, y, y] = \det_{\mathcal{E}}(x, y, y) = 0.$$

Montrons 3. Comme la famille (x, y) est orthonormée, elle se complète en une base orthonormée (x, y, z) qu'on peut supposer directe quitte à changer z en $-z$. Dans cette base, on a alors

$$\mathrm{Mat}_{(x,y,z)}(x \wedge y) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Donc $z = x \wedge y$. □

Exemple III.4.9. Si (e_1, e_2, e_3) est une base orthonormée directe de E, on a

$$e_3 = e_1 \wedge e_2 \qquad \text{et} \qquad e_2 = e_3 \wedge e_1 \qquad \text{et} \qquad e_1 = e_2 \wedge e_3$$

Remarque III.4.10. Le produit vectoriel n'est pas associatif :

$$e_1 \wedge (e_1 \wedge e_2) = e_1 \wedge e_3 = -e_2,$$

mais

$$(e_1 \wedge e_1) \wedge e_2 = 0 \wedge e_2 = 0.$$

Proposition III.4.11 (Lien entre produit vectoriel et angle orienté)**.** *Soit $x, y \in E$ deux vecteurs* ***non nuls*** *et F un plan (i.e.* $\dim(F) = 2$*) contenant x et y. On munit le sous-espace euclidien F d'une orientation, par le choix d'une base orthonormée (f_1, f_2). On pose $f_3 = f_1 \wedge f_2$ de sorte que (f_1, f_2, f_3) base orthonormée directe de E. On peut donc définir $\theta(x, y)$ l'angle orienté dans F entre x et y. Alors*

$$x \wedge y = \|x\|\|y\| \sin\big(\theta(x, y)\big) f_3.$$

Démonstration. Pour simplifier, notons $\theta = \theta(x, y)$. Comme (f_1, f_2) est une base orthonormée de F, il existe $\varphi \in \mathbb{R}$ tel que

$$\frac{x}{\|x\|} = \cos\varphi f_1 + \sin\varphi f_2.$$

Alors, dans F, on sait que

$$\mathrm{Mat}_{(f_1,f_2)}\left(\frac{y}{\|y\|}\right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} = \begin{pmatrix} \cos(\theta+\varphi) \\ \sin(\theta+\varphi \end{pmatrix}.$$

On a donc

$$\begin{aligned}\mathrm{Mat}_{(f_1,f_2,f_3)}\left(\frac{x}{\|x\|}\right) \wedge \mathrm{Mat}_{(f_1,f_2,f_3)}\left(\frac{y}{\|y\|}\right) &= \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \cos(\theta+\varphi) \\ \sin(\theta+\varphi) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ \cos\varphi\sin(\theta+\varphi) - \sin\varphi\cos(\theta+\varphi) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sin\big((\theta+\varphi) - \varphi\big) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sin(\theta) \end{pmatrix}\end{aligned}$$

Ainsi, par bilinéarité du produit vectoriel, on a bien

$$\frac{x}{\|x\|} \wedge \frac{y}{\|y\|} = \frac{1}{\|x\|\|y\|} x \wedge y = (\sin\theta) f_3$$

□

Exercice 32. Soient $u, v, w \in E$ trois vecteurs.

1. Montrer que $u \wedge (v \wedge w) \in \mathrm{Vect}(v, w)$.
2. Plus précisément, montrer que

$$u \wedge (v \wedge w) = \langle u, v\rangle w - \langle u, w \wedge v \qquad \text{et} \qquad (u \wedge v) \wedge w = \langle u, w\rangle v - \langle v, w\rangle u$$

   Inditation : on pourra considérer une base orthonormée directe (e_1, e_2, e_3) telle que $v \in \mathrm{Vect}(e_1)$ et $w \in \mathrm{Vect}(e_1, e_2)$.
3. Montrer que

$$(u \wedge v) \wedge (u \wedge w) = [u, v, w]u.$$

4. Montrer que

$$[u \wedge v, v \wedge w, w \wedge u] = [u, v, w]^2.$$

5. Montrer que

$$u \wedge (v \wedge w) + v \wedge (w \wedge u) + w \wedge (u \wedge v) = 0.$$

Exercice 33.[Identité de Lagrange] Soient $x, y, z, a, b, c \in \mathbb{R}$ six nombres réels. Sans faire de calculs, montrer que

$$(a^2 + b^2 + c^2)(x^2 + y^2 + z^2) = (ax + by + cz)^2 + (ay - bx)^2 + (cx - az)^2 + (bz - cx)^2.$$

Indication : On pourra chercher à écrire un produit scalaire à l'aide d'un angle.

III.5 Étude des isométries vectorielles de dimension 3

À nouveau, dans cette section, $(E, \langle\cdot,\cdot\rangle, \mathcal{E}_0)$ est un espace euclidien orienté de dimension 3.

Commençons par nous intéresser aux valeurs propres d'une isométrie.

Lemme III.5.1. *Soit $u \in \mathrm{O}(E)$. Si $\lambda \in \mathbb{R}$ est une valeur propre de u, alors $\lambda \in \{-1, 1\}$.*

Démonstration. Soit $x \in E \setminus \{0\}$ un vecteur propre associé à u. Comme u est une isométrie, on a

$$\|x\| = \|u(x)\| = \|\lambda x\| = |\lambda|\|x\|.$$

Donc $|\lambda| = 1$ car $\|x\| \neq 0$. □

Lemme III.5.2. *1. Si $u \in \mathrm{SO}(E)$, alors 1 est valeur propre de u.*

2. Si $u \in \mathrm{O}^-(E)$, alors -1 est valeur propre de u.

Autrement dit, si $u \in \mathrm{O}(E)$, alors $\det(u)$ est une valeur propre de u.

Démonstration. Soit $\chi_u \in \mathbb{R}[X]$ le polynôme caractéristique de u. Comme χ_u est de degré 3 qui est impair, il admet au moins une racine réelle λ_1 qui est donc une valeur propre.

Si χ_u admet une valeur propre $\mu \in \mathbb{C} \setminus \mathbb{R}$, alors la troisième valeur propre de u est nécessairement $\overline{\mu}$ et, dans $\mathbb{C}$, on a

$$\det(u) = \lambda_1 \mu\overline{\mu} = \lambda_1 |\mu|^2.$$

Donc $|\lambda_1| = 1$ et λ_1 est de même signe que $\det(u)$. Or $|\det(u)| = 1$. Ainsi $\lambda_1 = \det(u)$.

Si χ_u n'admet que des valeurs propres réelles, notons λ_2, λ_3 les deux autres valeurs propres. On sait que $\lambda_1, \lambda_2, \lambda_3 \in \{-1, 1\}$.

Si $\det(u) = -1$, alors il existe $i \in \{1, 2, 3\}$ tel que $\lambda_i = \det(u) = -1$ car les trois valeurs propres ne sont pas 1.

Si $\det(u) = 1$, alors il existe $i \in \{1, 2, 3\}$ tel que $\lambda_i = \det(u) = 1$ car les trois valeurs propres ne sont pas -1. □

Proposition-définition III.5.3. Soit $u \in \mathrm{O}(E)$ et $\lambda = \det(u)$. Alors il existe une base orthonormée **directe** (f_1, f_2, f_3) et un réel $\theta \in \mathbb{R}$ tels que

$$\mathrm{Mat}_{(f_1,f_2,f_3)}(u) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Si $u \in \mathrm{SO}(E)$, i.e. $\lambda = 1$, on dit alors que u est *la rotation* d'axe orienté par f_1 et d'angle θ.

Si $u \in \mathrm{O}^-(E)$, i.e. $\lambda = -1$, on dit alors que u est *l'antirotation* d'axe orienté par f_1 et d'angle θ.

Démonstration. Comme λ est une valeur propre de u, on peut choisir f_1 un vecteur propre de norme 1 associé à λ, i.e. $u(f_1) = \lambda f_1$. Complétons f_1 en une base orthonormée **directe** (f_1, f_2, f_3).

On remarque que $F = \mathrm{Vect}(f_2, f_3) = \mathrm{Vect}(f_1)^\perp$. Comme u préserve l'orthogonalité, on a donc $u(F) = F$. Donc u induit une isométrie $v = u_F \in \mathrm{O}(F)$ de F. Dans la base (f_1, f_2, f_3), on a alors

$$\lambda = \det(u) = \det\left(\mathrm{Mat}_{(f_1,f_2,f_3)}(u)\right) = \begin{vmatrix} \lambda & 0 \quad 0 \\ 0 & \\ 0 & \mathrm{Mat}_{(f_2,f_3)}(v) \end{vmatrix} = \lambda \det(v)$$

Donc $\det(v) = 1$, ce qui nous dit que $v \in \mathrm{SO}(F)$ est une rotation du plan F. Ainsi, il existe $\theta \in \mathbb{R}$ tel que

$$\mathrm{Mat}_{(f_2,f_3)}(v) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

D'où le résultat par l'écriture de u dans la base (f_1, f_2, f_3). □

Remarque III.5.4. Si u est une rotation d'angle 0, alors $u = \mathrm{id}_E$.

Si u est une rotation d'angle π, alors u est en fait le retournement (symétrie) orthogonale par rapport à l'axe $\mathrm{Vect}(f_1)$.

Si u est une antirotation d'angle 0, alors u est en fait la réflexion orthogonale par rapport au plan F.

Si u est une antirotation d'angle π, alors $u = -\mathrm{id}_E$.

On retrouve donc bien les symétries orthogonales de E comme cas particuliers des rotations et antirotations.

Remarque III.5.5. Si on change f_1 en $-f_1$, alors la base $(-f_1, f_2, f_3)$ est **indirecte**.

Complétons $-f_1$ en une nouvelle base orthonormée **directe** $(-f_1, -f_2, f_3)$. En considérant la matrice de passage entre ces bases, la matrice de u devient

$$\begin{aligned} \mathrm{Mat}_{(-f_1,-f_2,f_3)}(u) &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos(-\theta) & -\sin(-\theta) \\ 0 & \sin(-\theta) & \cos(-\theta) \end{pmatrix}. \end{aligned}$$

Ainsi, on voit que changer f_1 en $-f_1$ change également θ en $-\theta$.

On fera donc très attention au fait que l'angle d'une rotation dépend, au signe près, du choix de l'axe de la rotation. Ce choix n'est pas « canonique » et on pourra le faire comme bon nous semble.

Fait III.5.6 (Méthode de calcul de l'axe et de l'angle d'une isométrie vectorielle en dimension 3)**.** *Supposons qu'on sache déjà que $u \in \mathrm{O}(E)$, ce qu'on peut vérifier en calculant $u^*u = \mathrm{id}_E$, ou encore ${}^tAA = I_3$ lorsque A est la matrice de u dans une base orthonormée.*

On commence par calculer $\lambda = \det(u) \in \{\pm 1\}$ quand on ne le connait pas.

Pour calculer l'axe de u, il suffit de déterminer un vecteur propre $f_1 \in \ker(u - \lambda\,\mathrm{id}_E) \setminus \{0\}$ qu'on choisit de norme 1.

Pour calculer l'angle θ de u, on calcule d'abord la trace $\mathrm{Tr}(A)$. On en déduit alors $|\theta| \in [0, \pi]$ comme suit. On a $\mathrm{Tr}(A) = \lambda + 2\cos\theta$. Donc

$$\cos\theta = \begin{cases} \frac{\mathrm{Tr}(A)-1}{2} & \text{si } u \in \mathrm{SO}(E) \\ \frac{\mathrm{Tr}(A)+1}{2} & \text{si } u \in \mathrm{O}^-(E) \end{cases}$$

Pour déterminer le signe de $\sin\theta$, on considère g_2 linéairement indépendant avec f_1. Le procédé de Gram-Schmidt nous dit qu'on peut trouver f_2 tel que (f_1, f_2) est une famille orthonormée et $g_2 = xf_1 + yf_2$ pour $x, y \in \mathbb{R}$. On ne cherche pas à déterminer f_2, x, y. Si $f_3 = f_1 \wedge f_2$, alors (f_1, f_2, f_3) est donc une base orthonormée directe de E. Dans cette base, on a $u(g_2) = xu(f_1) + yu(f_2) = \lambda x f_1 + y\cos\theta f_2 + y\sin\theta f_3$. Écrivons le produit mixte de f_1, g_2 et $u(g_2)$ dans cette base. Alors

$$\langle f_1 \wedge g_2, u(g_2)\rangle = [f_1, g_2, u(g_2)] = \det{}_{(f_1,f_2,f_3)}(f_1, g_2, u(g_2)) = \begin{vmatrix} 1 & x & \lambda x \\ 0 & y & y\cos\theta \\ 0 & 0 & y\sin\theta \end{vmatrix} = y^2\sin\theta$$

Ainsi $\langle f_1 \wedge g_2, u(g_2)\rangle$ et $\sin\theta$ sont de même signe, ce qui permet de déterminer $\theta \in [-\pi, \pi]$.

Pour résumer :

étape 1 *Déterminer $\lambda = \det(u)$ et une matrice A de u dans une base orthonormée.*

étape 2 *Calculer $\cos\theta = \frac{\mathrm{Tr}(A)-\lambda}{2}$.*

étape 3 *Déterminer $f_1 \in \ker(u - \lambda\,\mathrm{id}_E)$ de norme 1.*

étape 4 *Choisir $g_2 \notin \mathrm{Vect}(f_1)$ pour lequel il est facile de calculer $f_1 \wedge g_2$ ou $u(g_2)$ (au choix).*

étape 5 *Calculer le signe de $\langle f_1 \wedge g_2, u(g_2)\rangle$Déterminer le signe de $\sin\theta$.*

étape 6 *Conclure sur la valeur de $\theta \in [-\pi, \pi]$ à partir de la donnée de $\cos\theta$ et du signe de $\sin\theta$.*

Exemple III.5.7. Soit $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. On constate A permute les termes de la base canonique donc envoie une base orthonormée sur une base orthonormée. Ainsi $A \in \mathrm{O}_3(\mathbb{R})$.

De plus $\det(A) = 1$ donc $A \in \mathrm{SO}_3(\mathbb{R})$.

On a

$$\ker(A - I_3) = \left\{ \begin{pmatrix} a \\ y \\ z \end{pmatrix}, \begin{matrix} -x + z = 0 \\ x - y = 0 \\ y - z = 0 \end{matrix} \right\} = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

Posons alors $f_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ qui nous donne l'axe orienté de la rotation u.

Alors l'angle de u vérifie

$$\cos\theta = \frac{\mathrm{Tr}(A) - 1}{2} = \frac{0 - 1}{2}$$

Donc $\theta = \pm\frac{2\pi}{3}$.

Prenons $g_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$. Alors $u(g_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ et $f_1 \wedge g_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Ainsi

$$[f_1, g_2, u(g_2)] = \langle \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 \cdot 0 + 1 \cdot 1 + (-1) \cdot 0 = 1$$

Donc $\sin\theta > 0$. Ainsi, u est la rotation d'axe orienté par $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et d'angle $\theta = +\frac{2\pi}{3}$.

Chapitre IV

Formes quadratiques

Dans tout ce chapitre, E est un $\mathbb{R}$-espace vectoriel de dimension finie $n > 0$.

IV.1 Définition et écriture des formes quadratiques

Au premier chapitre, on a étudié la notion de forme bilinéaire $\varphi : E \times E \to \mathbb{R}$ et on a vu qu'étant donnée une base $\mathcal{E} = (e_1, \dots, e_n)$ de E, on pouvait définir la matrice de φ dans $\mathcal{E}$ par

$$\mathrm{Mat}_{\mathcal{E}}(\varphi) = \begin{pmatrix} \varphi(e_1, e_1) & \cdots & \varphi(e_1, e_n) \\ \vdots & & \vdots \\ \varphi(e_n, e_1) & \cdots & \varphi(e_n, e_n) \end{pmatrix}$$

Certaines propriétés peuvent alors se lire sur la matrice :

— φ est symétrique (i.e. $\forall x, y \in E,\ \varphi(x, y) = \varphi(y, x)$) si la matrice de φ dans n'importe quelle base est symétrique.

— φ est un produit scalaire si la matrice de φ dans n'importe quelle base est dans $\mathcal{S}_n^{++}(\mathbb{R})$.

On a vu qu'on pouvait lire le fait qu'une application $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ est une forme bilinéaire ou non par sa formule, et que les produits scalaires partageaient de nombreuses *identités remarquables* en commun avec les polynômes de degré 2. On aimerait donc avoir un bon cadre pour étudier les polynômes de degré 2 lorsque ceux-ci sont en *plusieurs indéterminées.*

Dans ce chapitre, on va alors s'intéresser à des formes bilinéaires **symétriques** mais qui ne seront pas nécessairement définies positives. L'idée générale sera alors qu'au lieu, ou en plus de munir E d'un produit scalaire, on pourra le munir d'une forme bilinéaire symétrique et travailler dans cet espace, qu'on pourra appeler « espace quadratique ».

IV.1.a Définition

Définition IV.1.1. Une application $Q : E \to \mathbb{R}$ est une *forme quadratique* sur E s'il existe une application bilinéaire symétrique $\varphi : E \times E \to \mathbb{R}$ telle que

$$\forall x \in E,\ Q(x) = \varphi(x, x).$$

On dira alors que Q est *la forme quadratique de* φ.

Notation IV.1.2. On notera alors $\mathcal{Q}(E)$ l'ensemble des formes quadratiques sur E.

Exemple IV.1.3. Soit $E = \mathbb{R}^2$ muni du produit scalaire canonique et de la base canonique $\mathcal{E}$. Si φ est la forme bilinéaire symétrique qui a pour matrice $A = \mathrm{Mat}_{\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, on peut définir une forme quadratique Q sur $\mathbb{R}^2$ en posant

$$Q(x, y) = \varphi\big((x, y), (x, y)\big) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 2xy + 2yx + 3y^2 = x^2 + 4xy + 3y^2$$

On observe que Q est un polynôme en les variables x et y dont tous les monômes sont de degré total 2.

Exercice 34. Montrer que $\mathcal{Q}(E)$ est un sous-espace vectoriel de l'espace vectoriel des fonctions $E \to \mathbb{R}$.

Remarque IV.1.4. Les formes quadratiques qui ne sont pas des produits scalaires sont parfois utiles en physique.

En relativité restreinte, le cône de lumière consiste à s'intéresser à l'espace-temps définit par $Q(x,y,z,t) \geqslant 0$ avec $Q(x,y,z,t) = x^2+y^2+z^2-c^2t^2$ où c est la vitesse de la lumière dans le vide.

En mécanique du solide, on est amené à s'intéressé à l'ellipsoïde d'inertie qui correspond à l'équation $Q(x,y,z) = 1$ où Q est une forme quadratique de $\mathbb{R}^3$.

IV.1.b Forme polaire

Proposition-définition IV.1.5. Si Q est une forme quadratique sur E, alors la forme bilinéaire symétrique φ telle que $\forall x \in E,\ Q(x) = \varphi(x,x)$ est uniquement déterminée par Q. Elle est donnée par les formules de polarisation suivantes :

$$\forall x,y \in E,\ \varphi(x,y) = \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big) = \frac{1}{4}\Big(Q(x+y) - Q(x-y)\Big)$$

On appelle alors φ *la forme polaire de* Q.

On peut retenir ces formules en essayant d'écrire le produit xy à partir d'identités remarquables usuelles :

$$xy = \frac{1}{2}\big((x+y)^2 - x^2 - y^2\big) = \frac{1}{4}\big((x+y)^2 - (x-y)^2\big)$$

Démonstration. Soit Q une forme quadratique sur E et φ **une** forme bilinéaire symétrique vérifiant

$$\forall x \in E,\ Q(x) = \varphi(x,x)$$

Si $x,y \in E$, alors on a

$$\begin{aligned} Q(x+y) &= \varphi(x+y,x+y) = \varphi(x,x) + \varphi(x,y) + \varphi(y,x) + \varphi(y,y) \\ &= \varphi(x,x) + 2\varphi(x,y) + \varphi(y,y) \\ &= Q(x) + Q(y) + 2\varphi(x,y) \end{aligned}$$

Donc nécessairement $\varphi(x,y) = \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big)$.

Par un calcul similaire, on a

$$Q(x-y) = Q(x) + Q(y) - 2\varphi(x,y)$$

Donc

$$Q(x+y) - Q(x-y) = \Big(Q(x) + Q(y) + 2\varphi(x,y)\Big) - \Big(Q(x) + Q(y) - \varphi(x,y)\Big) = 4\varphi(x,y)$$

Ce qui donne la seconde formule de polarisation. ☐

On a donc une bijection entre les formes quadratiques sur E et les formes bilinéaires symétriques sur E, qui est en fait un isomorphisme d'espaces vectoriels

$$\begin{array}{ccc} \mathcal{Q}(E) & \overset{\simeq}{\longleftrightarrow} & \mathrm{Sym}(E) \\ Q & \mapsto & \Big((x,y) \mapsto \frac{1}{2}\big(Q(x+y) - Q(x) - Q(y)\big)\Big) \\ \Big(x \mapsto \varphi(x,x)\Big) & \hookleftarrow & \varphi \end{array}$$

IV.1.c Matrice d'une forme quadratique

Définition IV.1.6 (Matrice d'une forme quadratique)**.** Soit Q une forme quadratique sur E et $\mathcal{E} = (e_1, \dots, e_n)$ une base de E. On appelle *matrice de* Q *dans* $\mathcal{E}$ la matrice notée

$$\mathrm{Mat}(Q,\mathcal{E}) = \mathrm{Mat}(\varphi,\mathcal{E}) = \begin{pmatrix} \varphi(e_1,e_1) & \cdots & \varphi(e_1,e_n) \\ \vdots & & \vdots \\ \varphi(e_n,e_1) & \cdots & \varphi(e_n,e_n) \end{pmatrix}$$

Comme la forme polaire est uniquement déterminée par Q, cette matrice est donc bien définie.

Proposition IV.1.7 (Changement de base)**.** *Soient* $\mathcal{E}$ *et* $\mathcal{E}'$ *deux bases de* E *et* Q *une forme quadratique sur* E*. Si* P *est la matrice de passage de* $\mathcal{E}$ *à* $\mathcal{E}'$*, alors*

$$\mathrm{Mat}(Q,\mathcal{E}') = {}^tP\,\mathrm{Mat}(Q,\mathcal{E})P$$

Démonstration. C'est immédiat car on l'a déjà fait, en général, pour la matrice d'une forme bilinéaire. ☐

IV.1.d Écriture polynomiale

Soit $Q : E \to \mathbb{R}$ une forme quadratique de E de forme polaire $\varphi : E \times E \to \mathbb{R}$.

Soit $\mathcal{E} = (e_1, \ldots, e_n)$ une base de E et notons $A = \mathrm{Mat}(Q, \mathcal{E} = \big(a_{i,j}\big)_{1\leqslant i,j\leqslant n}$ la matrice de Q dans la base E. Si $x = \displaystyle\sum_{i=1}^{n} x_i e_i$, alors

$$Q(x) = \varphi\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} x_j e_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \varphi(e_i, e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j} x_i x_j$$

On a donc écrit Q comme un polynôme en les coordonnées x_i de x.

On remarque que :

- chaque terme du polynôme a exactement deux facteurs en x_i et un coefficient réel ;
- le terme $a_{i,j}x_ix_j$ est égal au terme $a_{j,i}x_jx_i$ car la matrice A est symétrique ;
- le terme $a_{i,i}x_ix_i$ se réécrit $a_{i,i}x_ix_i = \varphi(e_i, e_i)x_i^2 = Q(e_i)x_i^2$.

Donc on peut écrire plus simplement

$$Q(x) = \sum_{i=1}^{n} Q(e_i) + \sum_{1\leqslant i<j\leqslant n} 2a_{i,j}x_ix_j$$

et on l'appelle *l'écriture polynomiale de* Q dans la base $\mathcal{E}$.

On remarque alors que chaque terme de ce polynôme en n indéterminées est un monôme en les x_i de degré exactement 2. On dit qu'il s'agit d'un polynôme *homogène* de degré 2 pour dire que tous ses monômes ont le même degré 2.

Réciproquement, si on se donne un polynôme homogène de degré 2, disons

$$P(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{1\leqslant i<j\leqslant n} \beta_{i,j}x_ix_j$$

alors on peut lui associer une forme quadratique de forme polaire φ et de matrice A par identification des coefficients, en posant :

$$\begin{cases} a_{i,i} = \alpha_i & \forall 1 \leqslant i \leqslant n \\ 2a_{i,j} = \beta_{i,j} & \forall 1 \leqslant i < j \leqslant n \\ a_{j,i} = a_{j,i} & \forall 1 \leqslant j < i \leqslant n \end{cases}$$

On fera notamment attention au fait que $a_{i,j} = a_{j,i} = \frac{1}{2}\beta_{i,j}$ alors que $a_{i,i} = \alpha_i$.

Fait IV.1.8. *Par le choix d'une base $\mathcal{E}$ de E, on a une correspondance bijective entre les polynômes homogènes de degré 2 en n indéterminées et les formes quadratiques sur E.*

Exemple IV.1.9. (1) Soit $Q_1 : \mathbb{R}^2 \to \mathbb{R}$ donnée par $Q_1(x, y) = x^2 - y^2$. Tous ses monômes sont de degré 2 donc c'est une forme quadratique. Dans la base canonique, les coefficients de sa forme polaire sont $a_{1,1} = 1$, $a_{2,2} = -1$ et $a_{1,2} = a_{2,1} = 0$. Donc la matrice de Q_1 est $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(2) Soit $Q_2 : \mathbb{R}^2 \to \mathbb{R}$ donnée par $Q_2(x, y) = x^2 + 3y^2 + 4xy$. Tous ses monômes sont de degré 2 donc c'est une forme quadratique. Dans la base canonique, les coefficients de sa forme polaire sont $a_{1,1} = 1$, $a_{2,2} = 3$ et $a_{1,2} = a_{2,1} = \frac{4}{2} = 2$. Donc la matrice de Q_2 est $A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. On retrouve bien la forme polaire de l'exemple IV.1.3.

(3) Soit $Q_3 : \mathbb{R}^3 \to \mathbb{R}$ donnée par $Q_3(x, y, z) = x^2 - y^2 + xz$. Tous ses monômes sont de degré 2 donc c'est une forme quadratique. Dans la base canonique, la matrice de Q_3 est alors $A_3 = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}$.

(4) Soit $Q_4 : \mathbb{R}^3 \to \mathbb{R}$ donnée par $Q_4(x, y, z) = x^2 + y^2 + z^2 - xyz$. Bien qu'il y ait une symétrie en x, y, z, on remarque que le monôme $-xyz$ est de degré 3 donc Q_4 n'est pas une forme quadratique. Ca n'aura donc **aucun sens** de parler de sa matrice ou de sa forme polaire.

IV.2 Orthogonalité et bases orthogonales

Comme pour les produits scalaires, on va pouvoir définir une notion d'orthogonalité des formes quadratiques. A nouveau, l'orthogonalité dépendra de la forme quadratique choisie.

IV.2.a Orthogonalité

Définition IV.2.1. Soit Q une forme quadratique et φ sa forme polaire.

(1) Deux vecteurs $x, y \in E$ sont dits Q-*orthogonaux* (ou φ-orthogonaux) si $\varphi(x,y) = 0$.

(2) Deux parties $A, B \subseteq E$ sont dites Q-*orthogonales* (ou φ-orthogonames) si

$$\forall a \in A,\ \forall b \in B,\ \varphi(a,b) = 0.$$

On dispose ainsi d'une nouvelle notion d'orthogonalité pour une forme bilinéaire symétrique φ, qui étend la notion d'orthogonalité pour les produit scalaires. Cependant, des phénomènes nouveaux peuvent apparaître.

Exemple IV.2.2. Soit φ la forme bilinéaire symétrique sur $\mathbb{R}^2$ qui a pour matrice dans la base canonique $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Si on calcule

$$\varphi(e_1, e_2) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 + 0 = 1,$$

alors on constate que les deux vecteurs de la base canonique (e_1, e_2) ne sont pas φ-orthogonaux.

Plus surprenant, si on calcule

$$\varphi(e_1, e_1) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \times 0 + 0 \times 1 = 0,$$

on observe que le vecteur e_1 est φ-orthogonal à lui-même, ce qui n'était pas possible pour un produit scalaire.

On ne peut donc pas définir une notion de « norme » associée à une forme quadratique.

De même, le calcul donne que $\varphi(e_2, e_2) = 0$ donc e_2 est aussi φ-orthogonal à lui-même. Mais, même si (e_1, e_2) est une base de $\mathbb{R}^2$, cela ne veut pas dire que tous les vecteurs sont φ-orthogonaux à eux-mêmes. Par exemple

$$\varphi(e_1 + e_2, e_1 + e_2) = \varphi(e_1, e_1) + 2\varphi(e_1, e_2) + \varphi(e_2, e_2) = 0 + 2 \times 1 + 0 = 2$$

Donc le vecteur $e_1 + e_2$ n'est pas φ-orthogonal à lui-même.

On sera donc très prudent sur la notion de φ-orthogonalité lorsque φ n'est pas un produit scalaire.

Comme pour les produits scalaires, on peut définir la notion d'orthogonal d'une partie.

Définition IV.2.3. Soit Q une forme quadratique de forme polaire φ. Soit $A \subseteq E$ une partie non vide. On appelle Q-*orthogonal de* A (ou φ-orthogonal de A) l'ensemble noté

$$A^\circ = \{x \in E \mid \forall a \in A,\ \varphi(x,a) = 0\}.$$

Attention, même si ça n'apparait pas clairement dans la notation, l'espace A° dépend à la fois de A **et de** Q.

Proposition IV.2.4. *Avec les notations précédentes,* A° *est un sous-espace vectoriel de* E.

Démonstration. On a $\forall a \in A,\ \varphi(a, 0) = 0$ donc $0 \in A^\circ$.

De plus, si $x, y \in A^\circ$ et $\lambda \in \mathbb{R}$, alors on a

$$\forall a \in A,\ \varphi(x + \lambda y, a) = \varphi(x,a) + \lambda\varphi(y,a) = 0 + \lambda \times 0 = 0.$$

Donc $x + \lambda y \in A^\circ$. Donc A° est un sous-espace vectoriel de E. □

Définition IV.2.5. Soit Q une forme quadratique de forme polaire φ. On appelle *noyau de* Q (ou de φ) le Q-orthogonal de E, c'est-à-dire le sous-espace vectoriel

$$E^\circ = \{y \in E \mid \forall x \in E,\ \varphi(x,y) = 0\}$$

formé des vecteurs $y \in E$ qui sont φ-orthogonaux pour tous les éléments de E.

Le noyau d'une forme quadratique est donc un sous-espace vectoriel de E, cependant, ce n'est pas en général l'ensemble $Q^{-1}(\{0\})$ des éléments x tels que $Q(x) = 0$.

Exemple IV.2.6. Reprenons l'exemple précédent avec Q la forme quadratique de forme polaire φ ayant pour matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans la base canonique, c'est-à-dire que

$$\varphi\big((x,y),(u,v)\big) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix} = xv + yu$$

On a donc $Q(x,y) = 2xy$.

On observe que

$$\begin{aligned} Q^{-1}(\{0\}) &= \{(x,y) \in \mathbb{R}^2 \mid 2xy = 0\} \\ &= \{(x,y) \in \mathbb{R}^2 \mid x = 0 \text{ ou } y = 0\} \\ &= \{(x,0) \mid x \in \mathbb{R}\} \cup \{(0,y) \mid y \in \mathbb{R}\} \\ &= \operatorname{Vect}\big((1,0)\big) \cup \operatorname{Vect}\big((0,1)\big) \end{aligned}$$

est la réunion des deux droites de l'axe des abscisses et de l'axe des ordonnées.

Si, maintenant, on calcule E^0, alors on cherche l'ensemble des $(x,y) \in \mathbb{R}^2$ tels que quel que soit $(u,v) \in \mathbb{R}^2$, on a $\varphi\big((x,y),(u,v)\big) = 0$. Considérons un tel vecteur (x,y). En prenant $(u,v) = (y,x)$, on obtient $\varphi\big((x,y),(y,x)\big) = x^2 + y^2$. Comme $\varphi\big((x,y),(u,v)\big) = 0$, on a donc $x = 0$ et $y = 0$. Ainsi $(0,0)$ est le seul vecteur de $\mathbb{R}^2$ qui vérifie $\varphi\big((x,y),(u,v)\big) = 0$ pour tout $(u,v) \in \mathbb{R}^2$. Par conséquent $E^0 = \{(0,0)\} \subsetneq Q^{-1}(\{0\})$. Le noyau de Q n'est donc pas l'ensemble des vecteurs sur lesquels Q s'annule dans cet exemple.

Exercice 35. Soit Q une forme quadratique sur E Montrer que si $v \in E^0$ est dans le noyau de Q, alors on a $Q(v) = 0$.

L'exemple précédent montre que la réciproque est fausse !

Remarque IV.2.7. Lorsque ψ est un produit scalaire, on a vu que si F est un sous-espace vectoriel de E, alors $E = F \oplus F^\perp$. Cela permettait d'écrire

$$\dim F^\perp = \dim E - \dim F$$

Ici, pour une forme bilinéaire symétrique, on n'a plus de tel résultat en général. En effet, lorsque $E^\circ \neq 0$, pour n'importe quel sous-espace vectoriel F, on aura toujours $E^\circ \subseteq F^\circ$ par définition. En particulier, si on prend $F = E^\circ$, on a $F \cap F^\circ = E^\circ \neq 0$. Ici, en général,

$$\dim F^\circ \neq \dim E - \dim F$$

Par exemple, pour $F = E$, on a, en général

$$\dim E^\circ \neq \dim E - \dim E = 0$$

De même, pour $F = E^\circ$, on a, $(E^\circ)^\circ = E$ donc, en général

$$n = \dim(E^\circ)^\circ \neq \dim E - \dim E^\circ$$

IV.2.b Rang d'une forme quadratique et non-dégénérescence

Il semble donc utile de s'intéresser au cas où $E^\circ = 0$.

Définition IV.2.8. Une forme quadratique Q est dite *non dégénérée* si $E^0 = \{0\}$ et *dégénérée* sinon.

On appelle *rang de Q* le nombre entier noté

$$\operatorname{rg}(Q) = \dim E - \dim E^0.$$

Exemple IV.2.9. Si φ est un produit scalaire sur E et Q est la forme quadratique de φ, alors aucun vecteur n'est φ-orthogonal à lui-même. Donc, nécessairement, $E^0 = \{0\}$.

Exercice 36. Soit φ une forme bilinéaire symétrique positive sur E. Montrer que φ est un produit scalaire si, et seulement si, la forme quadratique associée à Q est non-dégénérée.

On dispose de critères matriciels pour déterminer le noyau, le rang et la (non-)dégénérescence de Q.

Proposition IV.2.10. *Soit Q une forme quadratique sur E de forme polaire φ. Soit $\mathcal{E}$ une base de E et $A = \mathrm{Mat}(Q, \mathcal{E})$ la matrice de Q dans la base $\mathcal{E}$. Alors*

(1) le noyau de Q est $E^0 = \ker A$;

(2) le rang de Q est $\mathrm{rg}(Q) = \mathrm{rg}(A)$;

(3) Q est non-dégénérée si, et seulement si, $\det(A) \neq 0$.

Démonstration. Soit $x, y \in E$ et X, Y leurs matrices respectives dans la base $\mathcal{E}$. On a $\varphi(x, y) = {}^tXAY$. Donc

$$\begin{aligned} y \in E^0 &\Leftrightarrow \forall x \in E,\ \varphi(x, y) = 0 \\ &\Leftrightarrow \forall X \in \mathcal{M}_{n,1}(\mathbb{R}),\ {}^tXAY = 0 \\ &\Leftrightarrow AY = 0 \end{aligned}$$

car cela signifie que pour le produit scalaire canonique $AY \in E^\perp = \{0\}$. Ainsi $y \in E^0 \Leftrightarrow Y \in \ker A$.

On a donc

$$\mathrm{rg}(Q) = \dim E - \dim \ker A = \mathrm{rg}(A)$$

Et, enfin, Q est non-dégénérée si, et seulement si, $\ker A = 0$, c'est-à-dire si, et seulement si, A est inversible. □

IV.2.c Bases orthogonales

Définition IV.2.11. Soit Q une forme quadratique sur E de forme polaire φ. Soit $\mathcal{F} = (f_1, \dots, f_d)$ une famille de d vecteurs de E. On dit que $\mathcal{F}$ est

— *Q-orthogonale* si $\forall 1 \leqslant i < j \leqslant d,\ \varphi(f_i, f_j) = 0$,

— *Q-orthonormée* si $\begin{cases} \forall 1 \leqslant i < j \leqslant d, & \varphi(f_i, f_j) = 0 \\ \forall 1 \leqslant i \leqslant d, & \varphi(f_i, f_i) = 1. \end{cases}$

Fait IV.2.12. *Soit Q une forme quadratique sur E. Une base $\mathcal{E}$ de E est Q-orthogonale (resp. Q-orthonormée) si, et seulement si, $\mathrm{Mat}(Q, \mathcal{E})$ est une matrice diagonale (resp. est la matrice I_n).*

Théorème IV.2.13. *Si Q est une forme quadratique sur E, alors il existe une base $\mathcal{E}$ de E qui est Q-orthogonale.*

Démonstration. Soit $\mathcal{F}$ une base quelconque de E. Alors $A = \mathrm{Mat}(Q, \mathcal{F})$ est une matrice symétrique réelle. D'après le théorème spectral matriciel, il existe donc une matrice diagonale D et une matrice orthogonale $P \in \mathrm{O}_n(\mathbb{R})$ telles que $D = {}^tPAP$. Considérons alors la base $\mathcal{E}$ de telle sorte que P est la matrice de passage de $\mathcal{F}$ à $\mathcal{E}$. On en déduit alors que $\mathrm{Mat}(Q, \mathcal{E}) = D$ donc $\mathcal{E}$ est une base Q-orthogonale. □

Remarque IV.2.14. On a en fait démontré un peu mieux. Supposons que E est un espace euclidien muni d'un produit scalaire ψ. Supposons que la base $\mathcal{F}$ choisie dans la preuve précédente est une base ψ-orthonormée. Alors, comme la matrice de passage de $\mathcal{F}$ à $\mathcal{E}$ est orthogonale, cela nous dit que $\mathcal{E}$ est une base ψ-orthonormée de E qui est aussi φ-orthogonal.

C'est ce qu'on appelle une *orthogonalisation simultanée* d'une forme bilinéaire symétrique φ par rapport à une forme bilinéaire symétrique définie positive ψ.

Remarque IV.2.15. Pour un produit scalaire ψ, on a vu par le procédé de Gram-Schmidt, qu'il est toujours possible de trouver une base ψ-orthonormée. Ce n'est plus le cas en général pour une forme quadratique.

Par exemple, la forme quadratique $Q(x, y) = 2xy$ sur $\mathbb{R}^2$ étudiée précédemment, de matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ dans la base canonique, n'admet aucune base Q-orthonormée.

En effet, supposons par l'absurde que $\mathcal{F}$ est une base Q-orthonormée. Cela signifie que $\mathrm{Mat}(Q, \mathcal{F}) = I_2$. Soit P la matrice de passage de la base canonique $\mathcal{E}$ vers $\mathcal{F}$. On a alors

$$I_2 = {}^tP\,\mathrm{Mat}(Q, \mathcal{E})P = {}^tP \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Mézalor

$$1 = \det(I_2) = \det({}^tP)\det(A)\det(P) = -\det(P)^2$$

Donc $\det(P)^2 = -1$, ce qui est absurde.

IV.2.d Isotropie

Concluons cette étude de l'orthogonalité en nous intéressant au cas des vecteurs qui sont Q-orthogonaux à eux-mêmes.

Définition IV.2.16. Soit Q une forme quadratique sur E. On dit que $x \in E$ est *un vecteur isotrope pour* Q si $Q(x) = 0$.

On appelle cône isotrope, l'ensemble noté $C(Q) = Q^{-1}(\{0\})$ constitué des vecteurs de E qui sont Q-isotropes.

Fait IV.2.17. *Un vecteur $x \in E$ est Q-isotrope si, et seulement si, x est Q-orthogonal à lui-même, c'est-à-dire que si φ désigne la forme polaire de Q, alors $\varphi(x, x) = 0$.*

Le cône isotrope d'une forme quadratique est un « cône » de E au sens où il vérifie la propriété suivante :

Proposition IV.2.18. *Soit Q une forme quadratique sur E. Alors*

$$\forall x \in E,\ \forall \lambda \in \mathbb{R},\ x \in C(Q) \Rightarrow \lambda x \in C(Q).$$

Démonstration. Soit $x \in C(Q)$ et $\lambda \in \mathbb{R}$, alors $Q(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda^2 \varphi(x, x) = \lambda^2 Q(x)$. Donc, si $Q(x) = 0$, alors $Q(\lambda x) = 0$. □

IV.3 Classification des formes quadratiques réelles

Soit Q une forme quadratique sur E. On a vu qu'il n'existait pas toujours de base orthonormée pour Q mais qu'il existait toujours des bases orthogonales. Il semble donc que dans certaines bases, la matrice de Q soit plus simple que dans d'autres. On se demande alors qu'est-ce qui est intrinsèquement caractérisé par Q, indépendamment de sa matrice.

Exemple IV.3.1. Soit $Q(x, y) = x^2 + 4xy + y^2$ la forme quadratique de matrice $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ dans la base canonique.

Si on choisit $f_1 = (3, 0)$ et $f_2 = (-2, 1)$. On considère alors la matrice de passage $P = \begin{pmatrix} 3 & -2 \\ 0 & 1 \end{pmatrix}$ de la base canonique vers la base (f_1, f_2) (qui est bien une base car échelonnée et génératrice de $\mathbb{R}^2$).

On sait alors que, dans cette nouvelle base, la matrice de Q est

$$\begin{aligned} B &= {}^tPAP \\ &= \begin{pmatrix} 3 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 6 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix} \end{aligned}$$

On voit que dans cette nouvelle base, la matrice de Q est plus simple et que Q s'écrit $Q(\alpha f_1 + \beta f_2) = 3\alpha^2 - 3\beta^2$, ce qui est une formule plus courte.

Cependant, on observe également que $\det(A) = -3 \neq \det(B) = -9$. De même $\mathrm{Tr}(A) = 2 \neq \mathrm{Tr}(B) = 0$.

Le déterminant et la trace d'une forme quadratique n'ont donc pas de sens car cette quantité dépend de la base choisie. On dit que ce ne sont pas des « invariants » de Q.

Mais alors, qu'est-ce qui dépend de Q ?

Proposition-définition IV.3.2. Soit Q une forme quadratique sur E. On appelle *discriminant* de Q le signe du déterminant de la matrice de Q dans une base $\mathcal{B}$ de E. Ce signe ne dépend pas de la base $\mathcal{B}$ choisie.

Démonstration. Si A et $B = {}^tPAP$ sont les matrices de Q dans deux bases de E et P la matrice de passage entre ces bases, alors on a

$$\det(B) = \det({}^tP) \det(A) \det(P) = \det(P)^2 \det(A).$$

Donc $\det(A)$ et $\det(B)$ sont de même signe. □

On observe que la forme quadratique d'un produit scalaire est de discriminant positif. Mais cela ne suffit pas pour que la forme polaire d'une forme quadratique soit un produit scalaire. Par exemple, dans $\mathbb{R}^2$, la matrice $-I_2$ est de déterminant positif mais sa forme bilinéaire symétrique associée n'est pas un produit scalaire puisqu'elle est définie négative.

On va donc chercher une donnée qui ne dépend que de Q et pas du choix de la base, mais qui soit plus riche que le discriminant et permettra de récupérer de nombreuses informations sur Q.

IV.3.a Loi d'inertie de Sylvester

Définition IV.3.3. Soit Q une forme quadratique sur E. On dit que Q est

— *positive* si $\forall v \in E,\ Q(v) \geqslant 0$;

— *définie positive* si $\forall v \in E \setminus \{0\},\ Q(v) > 0$;

— *négative* si $\forall v \in E,\ Q(v) \leqslant 0$;

— *définie négative* si $\forall v \in E \setminus \{0\},\ Q(v) < 0$;

Exercice 37. Si Q est définie positive ou définie négative, alors Q est non-dégénérée.
Indication : Comparer le noyau de Q et le cône isotrope de Q.

Théorème IV.3.4 (Loi d'inertie de Sylvester)**.** *Soit Q une forme quadratique sur E. Alors*

(1) Il existe un couple d'entiers naturels (s,t) vérifiant $s+t \leqslant n$ et une base $(f_1,\dots,f_n)$ de E telle que

$$Q\left(x_1 f_1 + \cdots + x_n f_n\right) = \sum_{i=1}^{s} x_i - \sum_{j=s+1}^{s+t} x_j$$

(2) Un tel couple (s,t) ne dépend pas de la base $(f_1,\dots,f_n)$ de E ainsi choisie. Plus précisément

$$s = \max\left\{\dim(F) \mid F \text{ est un sous-espace vectoriel de } E \text{ tel que } Q_{|F} \text{ est } \textbf{définie positive}.\right\}$$
$$t = \max\left\{\dim(F) \mid F \text{ est un sous-espace vectoriel de } E \text{ tel que } Q_{|F} \text{ est } \textbf{définie négative}.\right\}$$

De plus, si (s,t) est un couple d'entiers naturels vérifiant $s+t \leqslant n$, alors il existe une forme quadratique Q associée à un tel couple.

Remarque IV.3.5. Dire que Q est de signature (s,t) signifie donc qu'il existe une certaine base $\mathcal{B}$de E qui permet d'écrire Q sous la forme diagonale :

$$\mathrm{Mat}(Q,\mathcal{B}) = \left(\begin{array}{c|c|c} I_s & 0 & 0 \\ \hline 0 & -I_t & 0 \\ \hline 0 & 0 & 0 \end{array}\right) = \mathrm{diag}\left(\underbrace{1,\dots,1}_{s \text{ fois}}, \underbrace{-1,\dots,-1}_{t \text{ fois}}, 0,\dots,0\right)$$

L'indépendance de s et t signifie que dès qu'une base permet d'écrire Q sous cette forme, alors les valeurs de s et t sont entièrement déterminées.

Définition IV.3.6. Le couple (s,t) du théorème précédent s'appelle *la signature* de Q.

Exemple IV.3.7. La forme quadratique $Q(x,y,z,t) = x^2 + y^2 + z^2 - c^2t^2$ est de signature $(3,1)$ et de rang $r = s+t = 4$.

Démonstration. Pour (1), on sait déjà que Q admet une base $\mathcal{E} = (e_1,\dots,e_n)$ qui est Q-orthogonale. Ainsi, lorsqu'on écrit Q dans cette base, on a

$$Q(x_1e_1 + \cdots + x_ne_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2.$$

Quitte à changer l'ordre des e_i, on peut supposer que les λ_i sont rangés de telle sorte que $\lambda_1,\dots,\lambda_s$ sont strictement positifs, $\lambda_{s+1},\dots,\lambda_{s+t}$ sont strictement négatifs, et $\lambda_{s+t+1},\dots,\lambda_n$ sont nuls.

Posons alors $f_i = \frac{1}{\sqrt{|\lambda_i|}}$ pour tout $1 \leqslant i \leqslant s+t$ et $f_i = e_i$ pour $s+t+1 \leqslant i \leqslant n$. Dans la base des f_i, on a alors

$$Q\left(x_1f_1 + \cdots + x_nf_n\right) = Q\left(\frac{x_1}{\sqrt{|\lambda_1|}}e_1 + \cdots + \frac{x_{s+t}}{\sqrt{|\lambda_{s+t}|}}e_{s+t} + x_{s+t+1}e_{s+t+1} + \cdots + x_ne_n\right) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2.$$

Dans la base $\mathcal{F} = (f_1,\dots,f_n)$, on a alors

$$\mathrm{Mat}_{\mathcal{F}}(Q) = \left(\begin{array}{c|c|c} I_s & 0 & 0 \\ \hline 0 & -I_t & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

On en déduit alors que le noyau de Q est $E^0 = \mathrm{Vect}(f_{s+t+1},\dots,f_n)$. Notons également $E^+ = \mathrm{Vect}(f_1,\dots,f_s)$ et $E^- = \mathrm{Vect}(f_{s+1},\dots,f_{s+t})$. Par construction à partir de la base $\mathcal{F}$, on a immédiatement que $E = E^+ \oplus E^- \oplus E^0$.

Notons

$$s' = \max\left\{\dim(F) \mid F \text{ est un sous-espace vectoriel de } E \text{ tel que } Q_{|F} \text{ est } \textbf{définie positive}.\right\}$$

et

$$t' = \max\left\{\dim(F) \mid F \text{ est un sous-espace vectoriel de } E \text{ tel que } Q_{|F} \text{ est } \textbf{définie négative}.\right\}$$

Par lecture sur la matrice de Q on constate que $Q_{|E^+}$ est définie positive et que $Q_{|E^-}$ est définie négative. Donc $s' \geqslant s$ et $t' \geqslant t$.

Réciproquement, soit F un sous-espace vectoriel de E tel que $Q_{|F}$ est définie positive. Si $x \in F \cap (E^- \oplus E^0)$, qu'on écrit $x = y + z$ avec $y \in E^-$ et $z \in E^0$. Alors

$$Q(x) = Q(y+z) = \varphi(y,y) + 2\varphi(y,z) + \varphi(z,z) = Q(y) + 0 + 0 \leqslant 0$$

car $z \in E^0$. Mais comme $x \in F$ et Q est définie positive sur F, on a également $Q(x) \geqslant 0$ avec égalité si et seulement si $x = 0$. On est donc dans le cas d'égalité.

Ainsi $F \cap (E^- \oplus E^0) = \{0\}$ Donc

$$\dim(F) \leqslant \dim E - (\dim E^- \oplus E^0) = n - (t + (n - s - t)) = s.$$

Ainsi $s' \leqslant s$.

Par double inégalité, on a donc $s' = s$. Par un raisonnement analogue, on peut montrer que $t' = t$, ce qui démontre (2).

Enfin, si (s,t) est un couple d'entiers tels que $s + t \leqslant n$, alors la forme quadratique

$$Q(x_1, \dots, x_n) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2$$

a bien pour signature (s,t) par lecture sur la base canonique. □

Fait IV.3.8. *Si Q est une forme quadratique sur E de signature (s,t) alors le rang de Q est $\mathrm{rg}(Q) = s + t$.*

Le théorème de Sylvester nous permet de distinguer les formes quadratiques suivant leur signature, qui ne dépend pas de la base choisie. Mais, lorsque Q est une forme quadratique sur E, comment choisir une base $\mathcal{F}$ dans laquelle la matrice de Q est diagonale et de coefficients diagonaux dans $\{-1, 0, 1\}$?

Supposons qu'on ait trouvé une telle base $\mathcal{F} = (f_1, \dots, f_n)$ de E. Alors on sait qu'on peut écrire

$$Q(x_1 f_1 + \cdots + x_n f_n) = \sum_{i=1}^{s} x_i^2 - \sum_{j=s+1}^{s+t} x_j^2$$

Mais supposons qu'on souhaite écrire Q dans une autre base $\mathcal{E} = (e_1, \dots, e_n)$ qui serait, par exemple, la base canonique de $E = \mathbb{R}^2$. En écrivant la base $\mathcal{E}$ dans la base $\mathcal{F}$ sous la forme

$$e_k = \sum_{\ell=1}^{n} \lambda_{k,\ell} f_\ell,$$

avec $\lambda_{k,\ell}$ des coefficients dans $\mathbb{R}$, on a alors

$$\begin{aligned}
Q\Big(x_1 e_1 + \cdots + x_n e_n\Big) &= Q\Big(\sum_{k=1}^{n} x_k \underbrace{\sum_{\ell=1}^{n} \lambda_{k,\ell} f_\ell}_{=e_k}\Big) \\
&= Q\Big(\sum_{\ell=1}^{n} \left(\sum_{k=1}^{n} \lambda_{k,\ell} x_k\right) f_\ell\Big) \\
&= \sum_{i=1}^{s} \left(\sum_{k=1}^{n} \lambda_{k,i} x_k\right)^2 - \sum_{j=s+1}^{s+t} \left(\sum_{k=1}^{n} \lambda_{k,j} x_k\right)^2
\end{aligned}$$

On voit donc que les polynômes en n-variables $P_\ell(x_1, \dots, x_n) := (\sum_{k=1}^{n} \lambda_{k,\ell} x_k)$ homogènes de degré 1 occupent une place clé dans cette écriture. On va donc chercher à construire la base $\mathcal{F}$ à partir d'une écriture de Q dans la base $\mathcal{E}$ sous la forme de tels polynômes P_ℓ, en déduisant $\mathcal{F}$ des $P_\ell(x_1, \dots, x_n)$.

IV.3.b Interlude : formes linéaires et bases antéduales

On rappelle que

Définition IV.3.9. Si E est un espace vectoriel, on appelle *forme linéaire* sur E une application linéaire $\varphi : E \to \mathbb{R}$.

Le mot « forme » peut être retenu comme « à valeurs dans $\mathbb{R}$ ».

Définition IV.3.10. On appelle *espace dual* de E et on note E^* l'ensemble des formes linéaires sur E. C'est un sous-espace vectoriel des fonctions $E \to \mathbb{R}$.

Fait IV.3.11. *Si E est de dimension finie, alors E^* est de même dimension que E.*

Remarque IV.3.12. Attention ! Il peut y avoir des difficultés en dimension infinie.

Démonstration. Soit $\mathcal{E}$ une base de E de dimension n, alors $E^* = \mathrm{L}(E, \mathbb{R})$ s'identifie à l'espace vectoriel des matrices lignes $\mathcal{M}_{1,n}$ dans la base E, donc est de dimension n. □

Proposition-définition IV.3.13 (Rappel : base duale)**.** Soit $(e_1, \dots, e_n)$ une base de E. Alors il existe une unique famille $(\varphi_1, \dots, \varphi_n)$ de formes linéaires sur E telle que

$$\forall 1 \leqslant i, j \leqslant n \quad \varphi_i(e_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

La famille $(\varphi_1, \dots, \varphi_n)$ est alors une base de E^* et on l'appelle la *base duale* de la base $(e_1, \dots, e_n)$.

Proposition-définition IV.3.14 (base duale)**.** Soit E un $\mathbb{R}$-espace vectoriel de dimension finie. Si $(\varphi_1, \dots, \varphi_n)$ est une base de E^*, alors il existe une unique famille $(e_1, \dots, e_n)$ de E telle que

$$\forall 1 \leqslant i, j \leqslant n \quad \varphi_i(e_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

La famille $(e_1, \dots, e_n)$ est alors une base de E et on l'appelle la *base antéduale* de la base $(\varphi_1, \dots, \varphi_n)$.

Démonstration. Soit $(\varphi_1, \dots, \varphi_n)$ une base de E^*. On considère l'application

$$\begin{array}{cccc} \Psi : & E & \to & \mathbb{R}^n \\ & x & \mapsto & \big(\varphi_1(x), \dots, \varphi_n(x)\big) \end{array}$$

Comme les φ_i sont linéaires, Ψ est une application linéaire. Montrons que Ψ est bijective.

Supposons par l'absurde qu'il existe $x \in \ker \Psi \setminus \{0\}$. Comme $x \neq 0$, on peut compléter la famille libre (x) en une base $(f_1, \dots, f_n)$ telle que $x = f_1$. Sur cette base, on peut définir la forme linéaire $\varphi : E \to \mathbb{R}$ telle que $\varphi(f_i) = \begin{cases} 1 & \text{si } i = 1 \\ 0 & \text{sinon.} \end{cases}$. Comme $(\varphi_1, \dots, \varphi_n)$ est une base de E^*, il existe donc des réels $\alpha_1, \dots, \alpha_n$ tels que

$$\varphi = \alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n$$

Mais $\Psi(x) = \big(\varphi_1(x), \dots, \varphi_n(x)\big) = 0$ dans $\mathbb{R}^n$ donc $\varphi_1(x) = \dots = \varphi_n(x) = 0$. Ainsi

$$\varphi(x) = 0 + \dots + 0 = 0$$

Ce qui contredit $\varphi(x) = \varphi(f_1) = 1$. Donc Ψ est injective. Comme Ψ est injective entre espaces de même dimension, elle est bijective.

Notons alors $(E_1, \dots, E_n)$ la base canonique de $\mathbb{R}^n$. Alors la base $(e_1, \dots, e_n)$ donnée par $e_i = \Psi^{-1}(E_i)$ est une base de E comme image par un isomorphisme d'espaces vectoriels. De plus, pour tout $1 \leqslant j \leqslant n$, on a

$$\Psi(e_j) = E_j = (0, \dots, 0, 1, 0 \dots, 0) = \big(\varphi_1(e_j), \dots, \varphi_n(e_j)\big)$$

Donc pour tout $1 \leqslant i \leqslant n$, on a

$$\varphi_i(e_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

□

IV.3.c Procédé d'orthogonalisation de Gauss

Proposition IV.3.15. *Soient* $(\varphi_1, \dots, \varphi_r)$ *une famille libre de formes linéaires sur* E *et* $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ *des nombres réels. Alors l'application*

$$\begin{array}{rccl} Q: & E & \to & \mathbb{R} \\ & x & \mapsto & \alpha_1\varphi_1(x)^2 + \cdots + \alpha_r\varphi_r(x)^2 \end{array}$$

est une forme quadratique sur E *de signature* (s, t) *avec*

$$s = \operatorname{Card}\big\{i \in [\![1, r]\!] \mid \alpha_i > 0\big\} \qquad et \qquad sT = \operatorname{Card}\big\{j \in [\![1, r]\!] \mid \alpha_j < 0\big\}.$$

Démonstration. Dans le cas particulier où $r = 1$, on voit que

$$\begin{aligned} Q(x+y) - Q(x) - Q(y) &= \alpha_1\varphi_1(x+y)^2 - \alpha_1\varphi_1(x)^2 - \alpha_1\varphi_1(xy)^2 \\ &= \alpha_1\Big(\varphi(x)^2 + 2\varphi_1(x)\varphi_1(y) + \varphi_1(y)^2 - \varphi_1(x)^2 - \varphi_1(y)^2\Big) \\ &= 2\alpha_1\varphi_1(x)\varphi_1(y) \end{aligned}$$

Dans le cas général, par linéarité, on observe donc que la forme polaire de Q est

$$Phi(x, y) = \sum_{k=1}^{r} \alpha_k\varphi_k(x)\varphi_k(y)$$

qui est bilinéaire donc Q est bien une forme quadratique de forme polaire Φ.

Complétons alors la famille libre $(\varphi_1, \dots, \varphi_r)$ en une base $\mathcal{F} = (\varphi_1, \dots, \varphi_n)$ de E^* et considérons sa base antéduale $(f_1, \dots, f_n)$. Dans cette base, on a

$$Phi(f_i, f_j) = \sum_{k=1}^{r} \alpha_k\varphi_k(f_i)\varphi_k(f_j) = \sum_{k=1}^{r} \alpha_k\delta_{i,k}\delta_{j,k} = \begin{cases} \alpha_i & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Donc

$$\operatorname{Mat}(Q, \mathcal{F}) = \operatorname{Mat}(\Phi, \mathcal{F}) = \operatorname{diag}(\alpha_1, \dots, \alpha_r, 0, \dots, 0).$$

Ainsi, quitte à changer chaque f_k en $\frac{1}{\sqrt{|\alpha_k|}}$ lorsque $\alpha_k \neq 0$, on voit que Q est bien de signature (s, t) en comptant le nombre de α_k de chaque signe. □

Théorème IV.3.16 (Théorème de réduction de Gauss)**.** *Soit* Q *une forme quadratique sur* E*. Alors il existe une base* $\mathcal{F}$ *qui est* Q*-orthogonale.*

La preuve de ce théorème va également donner un algorithme que nous pourrons appliquer en pratique.

Démonstration. Soit $(e_1, \dots, e_n)$ une base quelconque E.

On a vu précédemment qu'on peut identifier les formes linéaires sur E aux polynômes homogènes de degré 1 en n variables et les formes quadratiques aux polynômes homogènes de degré 2.

Il suffit de montrer que tout polynôme homogène Q de degré 2 en n variables $X_1, \dots, X_n$ s'écrit sous la forme

$$Q = \sum_{i=1}^{n} \alpha_i L_i^2, \tag{IV.1}$$

où $L_1, \cdots, L_n$ sont des polynômes homogènes de degré 1 en $X_1, \cdots, X_n$ linéairement indépendants et $\alpha_1, \cdots, \alpha_n \in \mathbb{R}$. En effet, si cette assertion est vraie, alors les formes linéaires φ_i associées aux polynômes L_i forment une base de E^*, et il suffit de considérer la base $\mathcal{F} = (f_1, \cdots, f_n)$ antéduale de la base $(\varphi_1, \dots, \varphi_n)$. On a alors

$$Q(x) = \sum_{i=1}^{n} \alpha_i(x_i')^2$$

où x est de coordonnées $(x_1', \cdots, x_n')$ dans $\mathcal{F}$ par construction. On en déduit immédiatement que $\mathcal{F}$ est orthogonale pour Q.

Nous allons donc démontrer le théorème de Gauss pour les polynômes homogènes de degré 2 en n variables par récurrence sur n.

Si $n = 0$ ou 1, c'est immédiat. Supposons maintenant que le théorème est vrai jusqu'au rang $n - 1$ avec $n \geq 2$, nous allons le montrer pour n.

Soit $Q \in K[X_1, \dots, X_n]$ homogène de degré 2. On l'écrit sous la forme

$$Q = \sum_{i=1}^{n} a_i X_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{i,j} X_i X_j.$$

Premier cas : il existe i tel que $a_i \neq 0$.

Quitte à permuter les variables, on peut supposer que $a_1 \neq 0$ puis même que $a_1 = 1$ (car si Q/a_1 vérifie (IV.1), Q également en multipliant les coefficients λ_i par a_1).

On peut alors écrire

$$\begin{aligned} Q &= X_1^2 + 2X_1 \sum_{j=2}^{n} a_j X_j + 2 \sum_{2 \leq i < j \leq n} a_{ij} X_i X_j \\ Q &= X_1^2 + 2X_1 L + Q_1 \end{aligned}$$

avec L et Q_1 des formes respectivement linéaire et quadratique en $n-1$ variables $X_2, \ldots, X_n$. On écrit alors

$$Q = (X_1 + L)^2 + (Q_1 - L^2),$$

et on choisit $alpha_1 = 1, L_1 = x_1 + L$. Par hypothèse de récurrence, comme $Q_1 - L^2$ est une forme quadratique en $n-1$ variables $X_2, \ldots, X_n$, on a

$$Q_1 - L^2 = \sum_{i=2}^{n} \alpha_i L_i^2,$$

et il ne reste plus qu'à montrer que $L_1, \cdots, L_n$ sont bien linéairement indépendantes. Les formes $L_2, \cdots, L_n$ le sont par construction, mais elles s'annulent toutes en $(1, 0, \cdots, 0)$ alors que L_1 non. On a donc bien l'indépendance linéaire.

Deuxième cas : pour tout i, on a $a_i = 0$.

À moins que Q soit nulle (cas immédiatement résolu), on peut supposer quitte à permuter les variables et normaliser que $a_{1,2} = tfrac12$. Ceci permet d'écrire (suivant les mêmes idées qu'auparavant)

$$Q = x_1 x_2 + x_1 L^{(1)} + x_2 L^{(2)} + Q'$$

avec L, L', Q' respectivement linéaire, linéaire et quadratique, qui dépendent seulement des variables $X_3, \ldots, X_n$. On peut réécrire ceci

$$\begin{aligned} Q &= (x_1 + L^{(2)})(x_2 + L^{(1)}) + (Q_1 - L^{(1)} L^{(2)}) \\ &= \frac{1}{4}\left(L_1^2 - L_2^2\right) + Q_2 \end{aligned}$$

avec

$$L_1 = x_1 + x_2 + L^{(1)} + L^{(2)}, \qquad L_2 = x_1 - x_2 - L^{(1)} + L^{(2)}, \qquad Q_2 = Q_1 - L^{(1)} L^{(2)}.$$

Par hypothèse de récurrence, on peut écrire

$$Q_2 = \sum_{i=3}^{n} \alpha_i L_i^2,$$

avec $L_3, \cdots, L_n$ linéaires en $X_3, \ldots, X_n$ et linéairement indépendantes. En insérant cette formule dans l'expression ci-dessus, on obtient (IV.1) pour Q. Pour l'indépendance linéaire, on observe que $L_3, \cdots, L_n$ engendrent en fait toutes les formes linéaires en $x_3, \cdots, x_n$, en particulier les formes $X_3, \ldots, X_n$. En conséquence, vu les définitions de L_1 et L_2, les formes linéaires $L_1, \cdots, L_n$ engendrent $X_1 + X_2$ et $X_1 - X_2$, donc X_1, X_2. Ceci prouve que cette famille engendre les polynômes linéaires en n variables, de dimension n, donc la famille est une base, en particulier libre. □

Exemple IV.3.17. Considérons la forme quadratique $Q(x, y, z, t) = x^2 - 2xt + t^2 + yz$.

On voit que le coefficient en x^2 est non nul donc on est dans le premier cas. On écrit alors Q sous la forme

$$\begin{array}{rcccccc} Q(x,y,z,t) & = & x^2 + 2x(ay + bz + ct) & & & + & Q'(y,z,t) \\ & = & x^2 + 2x(ay + bz + ct) + \big(ay + bz + ct\big)^2 & - & \big(ay + bz + ct\big)^2 & + & Q'(y,z,t) \\ & = & \big(x + (ay + bz + ct)\big)^2 & - & \big(ay + bz + ct\big)^2 & + & Q'(y,z,t). \end{array}$$

Ici, on prend $a = 0, b = 0, c = -1$ et $Q'(y, z, t) = (t^2 + yz)$. Ainsi

$$Q(x, y, z, t) = (x - t)^2 - t^2 + t^2 + yz = (x + t)^2 + yz.$$

On a alors une nouvelle forme quadratique en trois variables $\widetilde{Q}(y, z, t) = yz$ telle que

$$Q(x, y, z, t) = (x - t)^2 + \widetilde{Q}(y, z, t).$$

On va appliquer à nouveau notre algorithme à $\widetilde{Q}(y,z,t) = yz$. Ici on voit qu'il n'y a aucun terme au carré, on est donc dans le deuxième cas. On écrit ainsi

$$\begin{array}{rcccccc} \widetilde{Q}(y,z,t) & = & (z-at)(y-bt) & + & z(bt)+y(at) & + & Q''(t) \\ & = & \frac{1}{4}\big((z-at)+(y-bt)\big)^2 - \frac{1}{4}\big((z-at)-(y-bt)\big)^2 & + & z(bt)+y(at) & + & Q''(t) \end{array}$$

Ici, on prend $a = b = 0$ et $Q''(t) = 0$. Donc

$$\widetilde{Q}(y,z,t) = \frac{1}{4}(y+z)^2 - \frac{1}{4}(y-z)^2.$$

Finalement

$$Q(x,y,z,t) = (x-t)^2 + \frac{1}{4}(y+z)^2 - \frac{1}{4}(y-z)^2 + 0.t^2$$

On a donc écrit quatre formes linéaires

$$\begin{array}{ll} \varphi_1(x,y,z,t) = x-t & \varphi_2(x,y,z,t) = y+z \\ \varphi_3(x,y,z,t) = y-z & \varphi_4(x,y,z,t) = t \end{array}$$

qui forment une base de $(\mathbb{R}^4)^*$ (cette famille est automatiquement échelonnée donc libre par l'algorithme de Gauss).

En comptant les signes positifs, négatifs et nuls des coefficients devant ces quatre formes linéaires, à savoir $(1, \frac{1}{4}, -\frac{1}{4}, 0)$, on voit que Q est de signature $(s,t) = (2,1)$ et de rang $s+t = 3$.

On peut également trouver une base antéduale de $(\varphi_1, \varphi_2, \varphi_3, \varphi_4)$ donnée par

$$\begin{array}{ll} f_1 = (1,0,0,0) & f_2 = (0, \frac{1}{2}, \frac{1}{2}, 0) \\ f_3 = (0, \frac{1}{2}, -\frac{1}{2}, 0) & f_4 = (1,0,0,1) \end{array}$$

De cette base, on peut trouver :

— un sous-espace de dimension $s = 2$ sur lequel Q est définie positive $E^+ = \mathrm{Vect}(f_1, f_2)$;

— un sous-espace de dimension $t = 1$ sur lequel Q est définie négative $E^- = \mathrm{Vect}(f_3)$;

— le noyau de Q, qui est $E^0 = \mathrm{Vect}(f_4)$.

IV.4 Application aux coniques planes et aux quadriques (culturel)

Étant donné un polynôme en n variables, on peut s'intéresser au lieu d'annulation de ce polynôme dans $\mathbb{R}^n$.

Définition IV.4.1. Soit $P(x_1, \dots, x_n)$ un polynôme en n variables. On dit que P est de *degré total* d si le maximum des degrés des monômes de P vaut d.

On appelle *conique* de $\mathbb{R}^2$ le lieu d'annulation d'un polynôme $P(x,y) = ax^2 + bxy + cy^2 + dx + ey + f$ de degré 2 en 2 variables.

On appelle *quadrique* de $\mathbb{R}^2$ le lieu d'annulation d'un polynôme $P(x,y,z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j$ de degré 2 en 3 variables.

Cela correspond à une écriture en base canonique de $\mathbb{R}^n$ pour $n = 2$ ou 3. Plus généralement, si on a un espace euclidien E, on peut appeler quadrique de E un sous-ensemble de la forme

$$\{x \in E \mid Q(x) + \varphi(x) + c = 0\}$$

où $Q \in \mathcal{Q}(E)$ est une forme quadratique sur E, $\varphi \in E^*$ est une forme linéaire sur E et $c \in \mathbb{R}$ est une constante.

Si on a une base $(e_1, \dots, e_n)$ de E, l'écriture $Q(x) + \varphi(x) + c = 0$ correspond à un polynôme de degré inférieur à 2 :

$$P(x_1, \dots, x_n) = Q(x_1e_1 + \cdots + x_ne_n) + \varphi(x_1e_1 + \cdots + x_ne_n) + c$$

Motivation : Historiquement, l'étude des coniques et quadriques présente divers intérêts en physique.

Par exemple, le mouvement de deux corps célestes est solution d'une équation différentielle telle que la courbe suivie par chacun des deux corps est une conique (e.g. le mouvement d'une planète autour d'une étoile, d'un astéroïde autour d'une planète, d'une comète autour du Soleil, d'un satellite autour de la Terre, etc.)

Également, certaines quadriques de l'espaces ont la particularité d'être des « surfaces réglées », c'est-à-dire qu'on peut recouvrir ces surfaces par des familles de droites. C'est par exemple ainsi qu'on construit de grandes cheminées industrielles en béton (e.g. tours de refroidissement en béton d'une centrale nucléaire). Il est beaucoup plus commode de construire une surface réglée industriellement puisque les formes à utiliser sont des segments de droites.

La littérature étudiant les coniques est très riche. Dans ce cours, on va se contenter seulement de distinguer différents types de coniques, sans chercher à en identifier des éléments métriques remarquables (paramètre, excentricité) qui correspondent à certaines quantité de mouvement ou d'énergie en physique.

Moralement, en deçà d'une certaine quantité d'énergie cinétique initiale, un satellite suivra une trajectoire elliptique autour de la Terre et sera donc « en orbite » ; Au-delà d'une certaine quantité d'énergie cinétique initiale, il suivra une trajectoire hyperbolique et s'éloignera indéfiniment ; à moins que sa trajectoire ne passe par la Terre elle-même dans l'un de ces deux cas.

IV.4.a Classification des coniques affines

Soit $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 \mid Q(x,y) + \varphi(x,y) + c = 0\}$.

Si on s'intéresse à la signature de Q, il y a six situations :

$$(s,t) \in \{(2,0),(1,1),(0,2),(1,0),(0,1),(0,0)\}.$$

Si $(s,t) = (0,0)$, cela signifie $Q = 0$ et l'algèbre linéaire nous assure que $\mathcal{C}$ est soit vide, soit une droite affine. Sinon, dans le cas où $(s,t) = (0,2)$ ou $(s,t) = (0,1)$, on peut se ramener aux autres cas en changeant les signes car $-Q$ sera de signature $(2,0)$ ou $(1,0)$ et

$$\mathcal{C} = \{(x,y) \in \mathbb{R}^2 \mid -Q(x,y) - \varphi(x,y) - c = 0\}.$$

Il n'y a donc que 3 cas à étudier pour la signature de Q, qui sont $(2,0),(1,1),(1,0)$.

Étudier $\mathcal{C}$, c'est étudier la ligne de niveau $-c$ de la fonction en 2 variables

$$f(x,y) = Q(x,y) + \varphi(x,y)$$

Suivant la valeur de c, on sait que cette ligne de niveau peut être vide ou réduite à un point. Dans ces cas là, on dit que $\mathcal{C}$ est dégénérée. On va chercher à étudier les coniques non-dégénérées.

Théorème IV.4.2. *Soit $\mathcal{C}$ une ellipse d'équation $Q(x,y) + \varphi(x,y) + c = 0$. On suppose que $\mathcal{C} \neq \emptyset$ et $\mathcal{C} \neq \{(x_0,y_0)\}$. Alors*

1. *si Q est de signature $(s,t) = (2,0)$, alors $\mathcal{C}$ est une* ***ellipse****;*
2. *si Q est de signature $(s,t) = (1,1)$, alors $\mathcal{C}$ est soit une* ***hyperbole****, soit une réunion de* ***deux droites****;*
3. *si Q est de signature $(s,t) = (1,0)$, alors $\mathcal{C}$ est soit une* ***parabole****, soit* ***une droite****, soit une réunion de* ***deux droites****.*

Démonstration. On va distinguer les cas suivant la signature de Q.

Si Q est de signature $(s,t) = (2,0)$, alors on peut trouver une base (f_1,f_2) orthonormée de $\mathbb{R}^2$ telle que $Q(Xf_1 + Yf_2) = \frac{X^2}{a^2} + \frac{Y^2}{b^2}$. On peut alors écrire $\varphi(Xf_1 + Yf_2) = -rX + -sY$ avec $r,s \in \mathbb{R}$. Donc

$$\begin{aligned}\mathcal{C} &= \{Xf_1 + Yf_2 \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} - rX - sY + c = 0\} \\ &= \{Xf_1 + Yf_2 \mid (\frac{X}{a} - \frac{ra}{2})^2 + (\frac{Y}{b} - \frac{sb}{2})^2 = \frac{r^2a^2}{4} + \frac{s^2b^2}{4} - c\}\end{aligned}$$

On voit donc que :

- si la quantité $h = \frac{r^2a^2}{4} + \frac{s^2b^2}{4} - c$ est strictement négative, $\mathcal{C}$ est vide ;
- si $h = 0$, alors $\mathcal{C}$ est le point $Xf_1 + Yf_2$ donné par $X = \frac{ra^2}{2}$ et $Y = \frac{sb^2}{2}$;
- si $h > 0$, alors $\mathcal{C}$ est l'ensemble des points $Xf_1 + Yf_2$ où $(\frac{X}{a}, \frac{Y}{b}) \in \mathbb{R}^2$ décrit le cercle de centre $(\frac{ra^2}{2}, \frac{sb^2}{2})$ et de rayon $\sqrt{h}$. Ici, on a donc déformé ce cercle d'un facteur a dans la direction f_1 et d'un facteur b dans la direction f_2, ce qui fait de $\mathcal{C}$ une ellipse, comme déformation affine d'un cercle.

Si Q est de signature $(s,t)=(1,1)$, alors on peut trouver une base (f_1,f_2) orthonormée de $\mathbb{R}^2$ telle que $Q(Xf_1+Yf_2)=\frac{X^2}{a^2}-\frac{Y^2}{b^2}$. On peut alors écrire $\varphi(Xf_1+Yf_2)=-rX+sY$ avec $r,s\in\mathbb{R}$. Donc

$$\begin{aligned}\mathcal{C}&=\{Xf_1+Yf_2\mid \frac{X^2}{a^2}-\frac{Y^2}{b^2}-rX+sY+c=0\}\\&=\left\{Xf_1+Yf_2\;\middle|\;\Big(\frac{X}{a}-\frac{ra}{2}\Big)^2-\Big(\frac{Y}{b}-\frac{sb}{2}\Big)^2=\frac{r^2a^2}{4}-\frac{s^2b^2}{4}-c\right\}\\&=\left\{Xf_1+Yf_2\;\middle|\;\Big(\frac{X}{a}-\frac{ra}{2}+\frac{Y}{b}-\frac{sb}{2}\Big)\Big(\frac{X}{a}-\frac{ra}{2}-\frac{Y}{b}+\frac{sb}{2}\Big)=\frac{r^2a^2}{4}-\frac{s^2b^2}{4}-c\right\}\\&=\left\{Xf_1+Yf_2\;\middle|\;X'Y'=h\right\}\end{aligned}$$

où on a posé

$$X'=\Big(\frac{X}{a}-\frac{ra}{2}+\frac{Y}{b}-\frac{sb}{2}\Big)\qquad Y'=\Big(\frac{X}{a}-\frac{ra}{2}-\frac{Y}{b}+\frac{sb}{2}\Big)\qquad h=\frac{r^2a^2}{4}-\frac{s^2b^2}{4}-c.$$

- Si $h\neq 0$, alors $\mathcal{C}$ est une hyperbole d'asymptotes les droites $X'=0$ et $Y'=0$;
- Si $h=0$, alors $\mathcal{C}$ est la réunion des deux droites d'équations respectivement $X'=0$ et $Y'=0$.

Si Q est de signature $(s,t)=(1,0)$, alors on peut trouver une base (f_1,f_2) orthonormée de $\mathbb{R}^2$ telle que $Q(Xf_1+Yf_2)=\frac{X^2}{a^2}$. On peut alors écrire $\varphi(Xf_1+Yf_2)=-rX+sY$ avec $r,s\in\mathbb{R}$. Donc

$$\begin{aligned}\mathcal{C}&=\{Xf_1+Yf_2\mid \frac{X^2}{a^2}-rX+sY+c=0\}\\&=\left\{Xf_1+Yf_2\;\middle|\;\Big(\frac{X}{a}-\frac{ra}{2}\Big)^2+sY=\frac{r^2a^2}{4}-c\right\}\\&=\left\{Xf_1+Yf_2\;\middle|\;(X')^2+sY=h\right\}\end{aligned}$$

où on a posé

$$X'=\Big(\frac{X}{a}-\frac{ra}{2}\Big)\qquad h=\frac{r^2a^2}{4}-c.$$

- Si $s\neq 0$, alors $\mathcal{C}$ est une parabole 'équation $Y=\frac{-1}{s}(X')^2+\frac{h}{s}$ de direction asymptotique la droite d'équation $X'=0$;
- si $s=0$ et $h<0$, alors $\mathcal{C}$ est vide ;
- si $s=0$ et $h=0$, alors $\mathcal{C}$ est la droite d'équation $X'=0$;
- si $s=0$ et $h>0$, alors $\mathcal{C}$ est la réunion des deux droites d'équation $X'=\sqrt{h}$ et $X'=-\sqrt{h}$.

□

IV.4.b Esquisse de classification des quadriques affines

Définition IV.4.3. On appelle *quadrique affine* le lieu d'annulation $\mathcal{Q}$ dans $\mathbb{R}^3$ d'un polynôme f de degré total 2 en 3 variables, qu'on peut voir comme une fonction en 3 variables qui s'écrit sous la forme :

$$f(x,y,z)=Q(x,y,z)+\varphi(x,y,z)+c$$

où Q est une forme quadratique sur $\mathbb{R}^3$, φ est une forme linéaire sur $\mathbb{R}^3$ et $c\in\mathbb{R}$ est une constante.

On dit que $\mathcal{Q}$ est *non-dégénérée* (resp. *propre*) si la forme quadratique $\widetilde{Q}(x,y,z,t)=t^2f(\frac{x}{t},\frac{y}{t},\frac{z}{t})$, appelée *homogénéisée* de f, est une forme quadratique non-dégénérée (resp. non dégénérée et non-définie).

Fait IV.4.4. *Avec les notations précédentes, notons Φ la forme polaire de Q. Alors la quadrique $\mathcal{Q}$ a un centre de symétrie qui est le point $(x_0,y_0,z_0)\in\mathbb{R}^3$ si, et seulement si, on a*

$$\forall v\in\mathbb{R}^3,\ \Phi\big(v,(x_0,y_0,z_0)\big)=-\frac{1}{2}\varphi(v).$$

Lorsque c'est le cas, on dit que $\mathcal{Q}$ est une quadrique à centre.

Théorème IV.4.5. *La quadrique $\mathcal{Q}$ est à centre si, et seulement si, Q est non-dégénérée.*

Démonstration. Soit $\widetilde{\Phi}\in E^{**}$ la forme linéaire donnée par $v\in\mathbb{R}^3\mapsto\Phi(x,\cdot)\in E^*$. Alors v_0 est un centre de $\mathcal{Q}$ si, et seulement si, $\widetilde{\Phi}(v_0)=\frac{\varphi}{2}$. Donc v_0 est unique si, et seulement si, $\widetilde{\Phi}$ est injective. Ceci valant si, et seulement si, Q est non-dégénérée. □

On distingue alors 3 familles de quadriques affines non-triviales sur $\mathbb{A}$:

les cônes : ce sont les quadriques dont l'image contient l'unique centre (de symétrie). Il s'agit donc des cônes isotropes de formes quadratiques, à translation près. Leur classification se ramène donc à la classification des formes quadratiques de dimension 3.

les quadriques à centre pur : ce sont les quadriques $\mathcal{Q}$ qui admettent un centre de symétrie sans le contenir, par exemple les sphères, les cylindres, etc. Cela revient à étudier les équations $Q(x, y, z) = 1$ où Q est une forme quadratique.

les quadriques paraboliques : ce sont les quadriques qui n'ont pas de centre.

Elles se ramènent à une équation du type

$$z = Q'(x, y)$$

où Q' est une forme quadratique en deux variables.

Leur classification se ramène à la classification de Q'.

Exemple IV.4.6. Si $n = 3$, il y a alors 5 quadriques propres à transformation affine près :

$Q(x, y, z) = 1$ **avec** Q **de signature** $(3, 0)$ **:** ellipsoïde ;

$Q(x, y, z) = 1$ **avec** Q **de signature** $(2, 1)$ **:** hyperboloïde à une nappe ;

$Q(x, y, z) = 1$ **avec** Q **de signature** $(1, 2)$ **:** hyperboloïde à deux nappes ;

$z = Q'$ **avec** Q' **de signature** $(2, 0)$ **:** paraboloïde elliptique ;

$z = Q'$ **avec** Q' **de signature** $(1, 1)$ **:** paraboloïde hyperbolique.

Exercice 38.[Difficile] Justifier ce résultat et dessiner les images des quadriques correspondantes. Justifier la nomenclature.

On notera que les quadriques affines propres ne sont pas des cônes. Il y a également des quadriques affines dégénérées (cône à base elliptique, cylindre elliptique, cylindre hyperbolique, cylindre parabolique).

Chapitre V

Espaces hermitiens

Jusqu'à présent, on a toujours travaillé dans un $\mathbb{R}$-espace vectoriel. Mais on sait que, bien souvent, il est plus commode de résoudre les équations (linéaires, différentielles, etc.) sur $\mathbb{C}$. On aimerait donc définir des notions analogues à celles vues pour les espaces euclidiens dans le cadre d'un espace vectoriel complexe.

V.1 Formes sesquilinéaires et espaces hermitiens

Dans toute la suite, on se fixe E un $\mathbb{C}$-espace vectoriel de dimension finie n (c'est donc un $\mathbb{R}$-espace vectoriel de dimension $2n$ dont on aurait oublié la structure complexe).

V.1.a Formes sesquilinéaire

Les formes sesquilinéaires jouent un rôle analogue aux formes bilinéaires à ceci près qu'elles tiennent, en plus, compte de la « conjugaison complexe » sur E.

Définition V.1.1. Une *forme sesquilinéaire* sur E est une application $\varphi : E \times E \to \mathbb{C}$ telle que :
— $\forall x \in E,\ \varphi_d : y \mapsto \varphi(x, y)$ est $\mathbb{C}$-linéaire ;
— $\forall y \in E,\ \varphi_g : x \mapsto \varphi(x, y)$ est semi-linéaire, c'est-à-dire que

$$\forall z, z' \in E,\ \forall \lambda, \mu \in \mathbb{C},\ \varphi(\lambda z + \mu z', y) = \overline{\lambda}\varphi(z, y) + \overline{\mu}\varphi(z', y).$$

Remarque V.1.2. Le choix semi-linéaire à gauche plutôt qu'à droite est arbitraire mais a une incidence sur les notations. Si on fait un choix différent, il faudra inverser la gauche et la droite dans tout ce qui va suivre.

Soit $\varphi : E \times E \to \mathbb{C}$ une forme sesquilinéaire où E est de dimension finie. Ceci permet de définir une application semi-linéaire $\varphi_g : E \to E^*$ donnée par $\varphi_g(y) = (x \mapsto \varphi(x, y))$.

Définition V.1.3. On appelle *noyau* de φ, noté $\ker \varphi = \{y \in E \mid \varphi_g(y) \equiv 0\}$.

On dit que φ est *non-dégénérée* si $\ker \varphi = 0$ ou, ce qui est équivalent, si φ_g est injective.

Remarque V.1.4. Attention : a priori, cette définition n'est pas symétrique en x et en y et on n'a a priori pas de structure $\mathbb{C}$-linéaire dans l'autre sens (il n'y a pas de structure $\mathbb{C}$-linéaire des formes semi-linéaires).

V.1.b Formes hermitiennes et antihermitiennes

Puisqu'une forme sesquilinéaire n'est pas linéaire des deux côtés, ça n'aurait pas de sens de parler de symétrie. Néanmoins, on peut définir un analogue à la symétrie qui tient en plus compte de la structure complexe, comme suit :

Définition V.1.5. Une forme sesquilinéaire φ est dite *hermitienne* si

$$\forall x, y \in E,\ \varphi(x, y) = \overline{\varphi(y, x)}.$$

Une forme sesquilinéaire φ est dite *antihermitienne* si

$$\forall x, y \in E,\ \varphi(x, y) = -\overline{\varphi(y, x)}.$$

Proposition V.1.6. *(1) L'ensemble des formes sesquilinéaire, noté* $\mathrm{Sesq}(V)$ *est un* $\mathbb{R}$*-espace vectoriel. L'ensemble des formes hermitiennes* $\mathcal{H}(V)$ *(resp. antihermitiennes* $\mathcal{AH}(V)$*) en sont des* $\mathbb{R}$*-sous-espace vectoriels.*

(2) Toute forme sesquilinéaire se décompose de manière unique en somme d'une forme hermitienne et d'une forme antihermitienne. En particulier, comme $\mathbb{R}$*-espace vectoriel,* $\mathrm{Sesq}(V)$ *est somme directe de* $\mathcal{H}(V)$ *et de* $\mathcal{AH}(V)$*.*

Démonstration. (1) est évident.

(2) **Existence :** Les formes sesquilinéaires $\varphi_h : (x,y) \mapsto \varphi(x,y) + \overline{\varphi(y,x)}$ et $\varphi_a : (x,y) \mapsto \varphi(x,y) - \overline{\varphi(y,x)}$ sont respectivement hermitienne et antihermitienne (attention à l'ordre de x et y !), ce qui donne l'existence.

Unicité : Si $\varphi = \varphi'_h + \varphi'_a$ pour φ'_h forme hermitienne et φ'_h forme antihermitienne, alors $\psi = \varphi_h - \varphi'_h = \varphi'_a - \varphi_a$ est simultanément hermitienne et antihermitienne. Mais alors, pour tout $(x,y) \in V \times V$, on a

$$\psi(x,y) = (\varphi_h - \varphi'_h)(x,y) = \overline{(\varphi_h - \varphi'_h)(x,y)} = \overline{(\varphi'_a - \varphi_a))(x,y)} = -(\varphi'_a - \varphi_a)(x,y) = -\psi(x,y).$$

Ainsi $\psi = 0$. □

V.1.c Matrices des formes sesquilinéaires

On a vu que certains ensembles de matrices naturels permettaient de bien paramétrer les formes bilinéaires symétriques, les endomorphismes auto-adjoints, les isométries, les bases orthonormées. Proposons ici des définitions analogues qui serviront prochainement.

Notation V.1.7. La matrice d'une forme sesquilinéaire φ dans une $\mathbb{C}$-base $\mathcal{B} = (e_1, \dots, e_n)$ de E, est

$$\mathrm{Mat}(\varphi, \mathcal{B}) = [\varphi(e_i, e_j)]_{1 \le i,j \le n}.$$

Pour toute matrice $A = [a_{i,j}]_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathcal{M}_{m,n}(\mathbb{C})$, on note $\overline{A} = [\overline{a_{i,j}}]_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathcal{M}_{n,m}(\mathbb{C})$ et $A^* = {}^t\overline{A}$.

Fait V.1.8. *Soit φ une forme sesquilinéaire sur E et $\mathcal{B}$ une base de E. Pour $x, y \in E$, si $X, Y \in \mathcal{M}_{n,1}(\mathbb{C})$ représentent x, y respectivement dans la base $\mathcal{B}$, alors*

$$\varphi(x,y) = X^* \,\mathrm{Mat}(\varphi, \mathcal{B}) Y.$$

Fait V.1.9. *Matriciellement,*

$$\varphi \text{ est hermitienne } \iff \mathrm{Mat}(\varphi, \mathcal{B}(\varphi))^* = \mathrm{Mat}(\varphi, \mathcal{B})$$

et

$$\varphi \text{ est antihermitienne } \iff \mathrm{Mat}(\varphi, \mathcal{B}(\varphi))^* = -\mathrm{Mat}(\varphi, \mathcal{B}).$$

On est donc amené à poser la définition :

Définition V.1.10. Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite hermitienne si $A^* = A$ et antihermitienne si $A^* = -A$.

Notation V.1.11. On note $\mathcal{H}_n(\mathbb{C})$ l'ensemble des matrices hermitiennes et $\mathcal{AH}_n(\mathbb{C})$ l'ensemble des matrices antihermitiennes.

On a également besoin de définir un analogue à l'ensemble des matrices orthogonales.

Définition V.1.12. Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite *unitaire* si $A^*A = I_n$.

Notation V.1.13. On note $\mathrm{U}_n(\mathbb{C})$ l'ensemble des matrices unitaires.

V.1.d Orthogonalité

Pour définir une notion robuste d'orthogonalité, on a besoin d'une certaine symétrie en les deux paramètres d'une forme sesquilinéaire. On va donc se restreindre aux formes hermitiennes mais, comme l'a montré le chapitre sur les formes quadratiques, il pourrait exister un cadre plus général permettant de définir une notion d'orthogonalité (voir dans la littérature la notion de forme sesquilinéaire réflexive).

Définition V.1.14. Soit $\varphi : E \times E \to \mathbb{C}$ une forme hermitienne. On dit que deux vecteurs x, y sont orthogonaux si $\varphi(x,y) = 0$.

Si $A, B \subseteq E$ sont deux parties de E, on dit que A et B sont orthogonaux si

$$\forall a \in A,\ \forall b \in B,\ \varphi(a,b) = 0.$$

On appelle orthogonal d'une partie $A \subseteq E$ le sous-$\mathbb{C}$-espace vectoriel de E donné par

$$A^\perp = \{x \in E,\ \forall a \in A,\ \varphi(x,a) = 0\} = \{x \in E,\ \forall a \in A,\ \varphi(a,x) = 0\}$$

Fait V.1.15. *Une forme hermitienne φ est non-dégénérée si, et seulement si, $E^\perp = 0$.*

Proposition V.1.16. *Soit E de dimension finie et $\varphi : E \times E \to \mathbb{C}$ une forme hermitienne supposée* **non-dégénérée***. On a les résultats usuels sur les espaces orthogonaux, à savoir, pour F, F_1, F_2 des sous-espaces vectoriels de E :*

1. $\dim F^{\perp} = \dim E - \dim F$;
2. $0^{\perp} = E$ *et* $E^{\perp} = 0$*, donc* $(E^{\perp})^{\perp} = E$;
3. $(F_1 + F_2)^{\perp} = F_1^{\perp} \cap F_2^{\perp}$;
4. $(F_1 \cap F_2)^{\perp} = F_1^{\perp} + F_2^{\perp}$.

Démonstration. Par définition, un jeu de réécriture donne $F^{\perp} = (\varphi_g(F))^{\top}$ où l'application semi-linéaire $\varphi : E \to E^*$ est vue comme application $\mathbb{R}$-linéaire, qui est un isomorphisme car φ est non-dégénérée. □

V.1.e Espaces hermitiens

Définition V.1.17. On dit qu'une forme sesquilinéaire φ est :

— *positive* si $\forall x \in E,\ \varphi(x,x) \geqslant 0$;

— *négative* si $\forall x \in E,\ \varphi(x,x) \geqslant 0$;

— *définie positive* si $\forall x \in E \setminus \{0\},\ \varphi(x,x) > 0$;

— *définie négative* si $\forall x \in E \setminus \{0\},\ \varphi(x,x) < 0$.

Une forme hermitienne définie positive s'appelle *un produit scalaire* (préhilbertien).

Exemple V.1.18 (Exemples fondamentaux).

(1) Sur $\mathbb{C}^n$, le produit scalaire canonique est l'application

$$\varphi : \big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) \mapsto \overline{x_1} y_1 + \cdots + \overline{x_n} y_n$$

Matriciellement, cela s'écrit $\varphi(X, Y) = X^* Y$ lorsque X, Y sont des matrices colonnes.

(2) Si $E = \mathcal{C}^0([a,b], \mathbb{C})$ est un espace de fonctions à valeurs complexes sur un intervalle $[a,b]$, alors on peut définir sur E un produit scalaire par

$$\varphi(f,g) = \int_a^b \overline{f(t)} g(t) dt.$$

(3) Si $E = \mathcal{M}_n(\mathbb{C})$, on peut définir sur E le produit scalaire

$$\varphi(A, B) = \mathrm{Tr}(A^* B).$$

Proposition V.1.19 (Inégalité de Cauchy-Schwarz : cas général)**.** *Soit φ une forme hermitienne* **positive** *sur E. Alors pour tous $x, y \in E$, on a*

$$\varphi(x,y)\overline{\varphi(x,y)} \leqslant \varphi(x,x)\varphi(y,y).$$

Démonstration. On considère l'application $P : \mathbb{C} \to \mathbb{C}$ suivante :

$$\begin{aligned} P(t) =& \varphi(x+ty, x+ty) && \\ =& \varphi(x, x+ty) + \bar{t}\varphi(y, x+ty) && \text{par semi-linéarité à gauche} \\ =& \varphi(x,x) + t\varphi(x,y) + \bar{t}\varphi(y,x) + \bar{t}t\varphi(y,y) && \text{par linéarité à droite} \\ =& \varphi(x,x) + t\varphi(x,y) + \overline{t\varphi(x,y)} + \bar{t}t\varphi(y,y) && \text{par hermitianité de } \varphi \end{aligned}$$

Distinguons deux cas.

1er cas : On suppose que $\varphi(x,x) = \varphi(y,y) = 0$. Alors la restriction de P à $\mathbb{R}$ est une application linéaire $\mathbb{R} \to \mathbb{R}$ donnée par $t \mapsto t(\varphi(x,y) + \overline{\varphi(x,y)}) = 2t\mathfrak{Re}(\varphi(x,y))$ qui garde un signe positif. C'est donc l'application nulle, ce qui montre que $\mathfrak{Re}(\varphi(x,y)) = 0$.

Considérons également la restriction de P à $i\mathbb{R}$. C'est donc une application $\mathbb{R}$-linéaire $i\mathbb{R} \to \mathbb{R}$ donnée par $it \mapsto it\varphi(x,y) - it\overline{\varphi(x,y)} = -2t\mathfrak{Im}(\varphi(x,y))$ qui garde un signe constant positif. C'est donc l'application nulle, ce qui montre que $\mathfrak{Im}(\varphi(x,y)) = 0$.

Ainsi, dans ce premier cas, on a $\varphi(x,y) = 0$ d'où les deux termes de l'inégalité de Cauchy-Schwarz sont nuls.

2nd cas : Quitte à échanger les rôles de x et y, on peut supposer que $\varphi(y,y) > 0$. Posons astucieusement $t = -\frac{\overline{\varphi(x,y)}}{\varphi(y,y)}$ qui correspondrait au lieu du minimum de ce polynôme dans le cas réel. On a alors

$$\begin{aligned} P(t) =& \varphi(x,x) - \frac{\overline{\varphi(x,y)}}{\varphi(y,y)}\varphi(x,y) - \frac{\varphi(x,y)}{\varphi(y,y)}\overline{\varphi(x,y)} + \frac{\overline{\varphi(x,y)}}{\varphi(y,y)}\frac{\varphi(x,y)}{\varphi(y,y)}\varphi(y,y) \\ =& \varphi(x,x) - \frac{\varphi(x,y)\overline{\varphi(x,y)}}{\varphi(y,y)} \geqslant 0. \end{aligned}$$

Ce qui conclut. □

Définition V.1.20. On appelle *espace hermitien* un couple (H, φ) formé d'un $\mathbb{C}$-espace vectoriel H de dimension finie et d'un produit scalaire φ.

Définition V.1.21. Sur un espace hermitien (H, φ), on définit une norme par

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

Proposition V.1.22 (Inégalité de Cauchy-Schwarz)**.** *Soit (H, φ) un espace hermitien. Alors pour tous $x, y \in H$, on a*

$$|\varphi(x,y)| \leqslant \|x\|\|y\|$$

avec égalité si, et seulement si, la famille (x, y) est liée.

Démonstration. Le cas d'égalité correspond à $P(t) = 0$, donc à $x + ty = 0$ par non-dégénérescence. □

Corollaire V.1.23. *Une forme hermitienne positive est définie si, et seulement si, elle est non dégénérée.*

Proposition V.1.24. *Pour résumé, dans un espace hermitien (H, φ), si $x, y \in H$ et $\lambda \in \mathbb{C}$, on a les propriétés :*

1. $\|x\| \geqslant 0$ ***(positivité)*** *et* $\|x\| = 0 \Leftrightarrow x = 0$ ***(séparation)****;*
2. $\|\lambda x\| = |\lambda|\|x\|$ ***(positive homogénéité)****;*
3. $\|x + y\| \leqslant \|x\| + \|y\|$ ***(inégalité triangulaire directe)****;*
4. $\big|\|x\| - \|y\|\big| \leqslant \|x - y\|$ ***(inégalité triangulaire renversée)****;*
5. $\big|\varphi(x,y)\big| \leqslant \|x\|\|y\|$ ***(inégalité de Cauchy-Schwarz)****;*
6. $\|x + y\|^2 = \|x\|^2 + 2\mathfrak{Re}\big(\varphi(x,y)\big) + \|y\|^2$ ***(identité de polarisation 1)****;*
7. $\|x + y\|^2 - \|x - y\|^2 = 4\mathfrak{Re}\big(\varphi(x,y)\big)$ ***(identité de polarisation 2)****;*
8. $\|x\|^2 - \|y\|^2 = \|x + y\| + 2\mathfrak{Im}\big(\varphi(x,y)\big)$ ***(identité de polarisation 3)****;*
9. $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ ***(identité de la médiane*** *ou* ***du parallélogramme)****.*

Démonstration. Démonstrations analogues à celles des espaces euclidiens. □

Remarque V.1.25. Attention ! Dans les identités de polarisation, on n'a plus le produit scalaire mais seulement des parties réelles ou imaginaire. Il faut donc en utiliser deux différentes pour reconstituer φ à partir de $\|\cdot\|$.

V.1.f Bases orthogonales

Définition V.1.26. Soit φ une forme hermitienne sur E. Une base $(e_1, \ldots, e_n)$ de E est dite *orthogonale* pour φ (ou φ-orthogonale) si $\forall i \neq j,\ \varphi(e_i, e_j) = 0$.

Si, de plus, $\forall i,\ \varphi(e_i, e_i) = 1$, alors la base est dite *orthonormée* pour φ.

Fait V.1.27. *(1) La matrice de φ dans une base orthogonale est diagonale.*

(2) La matrice de φ dans une base orthonormée est l'identité.

Proposition V.1.28. *Soit φ une forme hermitienne non-dégénérée. Alors il existe une base φ-orthogonale de E.*

Démonstration. On procède par récurrence sur $n = \dim(E)$. Si $n = 1$, c'est évident.

Sinon, soit e_1 un vecteur tel que $\varphi(e_1, e_1) \neq 0$ et $F = \mathrm{Vect}(e_1)$. Alors φ restreinte à $F \times F$ est non-dégénérée donc F et $F^\perp$ sont en somme directe. On conclut en appliquant l'hypothèse de récurrence sur $F^\perp$. □

Proposition V.1.29 (Procédé d'orthonormalisation de Gram-Schmidt)**.** *Soit (H, φ) un espace hermitien et $(e_1, \ldots, e_r)$ une famille libre de H. Alors il existe une famille libre $(f_1, \ldots, f_r)$ de H qui est φ-orthonormée et telle que* $\mathrm{Vect}(f_1, \ldots, f_m) = \mathrm{Vect}(e_1, \ldots, e_m)$ *pour tout $1 \leq m \leq r$.*

Démonstration. Pour $m = 1$, on pose $f_1 = \frac{e_1}{\sqrt{\varphi(e_1,e_1)}}$.

Supposons $(f_1, \ldots, f_m)$ construite de sorte que :

- $\varphi(f_i, f_i) = 1$ et $\varphi(f_i, f_j) = 0$ pour $1 \leq i \neq j \leq m$
- et $\mathrm{Vect}(f_1, \ldots, f_i) = \mathrm{Vect}(e_1, \ldots, e_i)$ pour $1 \leqslant i \leqslant m$.

On considère alors

$$g_{m+1} = e_{m+1} - \sum_{i=1}^{m} \varphi(e_{m+1}, f_i) f_i \neq 0$$

car $e_{m+1} \notin \mathrm{Vect}(f_1, \ldots, f_m)$.

On pose $f_{m+1} = \frac{g_{m+1}}{\|g_{m+1}, g_{m+1}\|}$. On vérifie alors par linéarité que f_{m+1} convient. □

V.1.g Lien avec les matrices unitaires.

On va maintenant voir le lien entre les bases φ-orthonormées et les bases unitaires.

Lemme V.1.30 (Coordonnées des vecteurs pour un produit scalaire hermitien)**.** *Soit* $(e_1, \ldots, e_n)$ *une base φ-orthonormée. Alors*

$$x = \sum_{i=1}^{n} \varphi(e_i, x)e_i = \sum_{i=1}^{n} \overline{\varphi(x, e_i)}e_i.$$

Démonstration. En effet,

$$\begin{aligned}
\varphi(e_i, x) &= \varphi\left(e_i, \sum_{j=1}^{n} \lambda_j e_j\right) \\
&= \sum_{j=1}^{n} \lambda_j \varphi(e_i, e_j) \text{ par linéarité à droite,} \\
&= \lambda_i
\end{aligned}$$

Donc

$$x = \sum_{i=1}^{n} \lambda_i e_i = \sum_{i=1}^{n} \varphi(e_i, x)e_i.$$

□

Remarque V.1.31. Attention : comme φ n'est pas symétrique, la position de x dans le produit scalaire est importante !

Lemme V.1.32. *Soient $\mathcal{E}$ et $\mathcal{E}'$ deux bases orthonormées de H et P la matrice de passage de $\mathcal{E}$ à $\mathcal{E}'$. Alors $P \in \mathrm{U}_n(\mathbb{C})$.*

Démonstration. Notons $\mathcal{E} = (e_1, \ldots, e_n)$ et $\mathcal{E}' = (e'_1, \ldots, e'_n)$. Notons également Q la matrice de passage de la base $\mathcal{E}'$ à la base $\mathcal{E}$, de sorte que $Q = P^{-1}$.

Comme $\mathcal{E}$ est orthonormée, on sait qu'on peut écrire

$$\forall 1 \leqslant j \leqslant n,\ e'_j = \sum_{i=1}^{n} \varphi(e_i, e'_j)e_i$$

Le (i, j)-ème coefficient de P est donc $P_{i,j} = \varphi(e_i, e'_j)$.

De même, comme $\mathcal{E}'$ est orthonormée, on sait qu'on peut écrire

$$\forall 1 \leqslant i \leqslant n,\ e_i = \sum_{j=1}^{n} \varphi(e'_j, e_i)e'_j$$

Le (j, i)-ème coefficient de Q est donc $Q_{j,i} = \varphi(e'_j, e_i) = \overline{\varphi(e_i, e'_j)} = \overline{P_{i,j}}$.

On constate ainsi que $P^{-1} = Q = {}^t\overline{P} = P^*$. Donc P est une matrice unitaire. □

D'une manière analogue au cas euclidien, on peut démontrer :

Théorème V.1.33. *Soit $u \in \mathrm{End}(H)$ un endomorphisme. S'équivalent :*

(i) u est unitaire ;
(ii) il existe une base φ-orthonormée $\mathcal{E}$ telle que $\mathrm{Mat}_{\mathcal{E}}(u) \in \mathrm{U}_n(\mathbb{V})$;
(iii) pour toute base orthonormée $\mathcal{E}$, on a $\mathrm{Mat}_{\mathcal{E}}(u) \in \mathrm{U}_n(\mathbb{C})$;
(iv) il existe une base orthonormée $\mathcal{E} = (e_1, \ldots, e_n)$ telle que $\big(u(e_1), \ldots, u(e_n)\big)$ est une base orthonormée ;
(v) pour toute base orthonormée $\mathcal{E} = (e_1, \ldots, e_n)$ la famille $\big(u(e_1), \ldots, u(e_n)\big)$ est une base orthonormée.

Remarque V.1.34. Si $\mathcal{E}$ est une base orthonormée de $\mathcal{E}$ (n'importe laquelle), on peut donc définir une fonction

$$\begin{array}{rccc}\Phi : & \mathrm{U}(E) & \to & \mathrm{U}_n(\mathbb{C}) \\ & u & \mapsto & \mathrm{Mat}_{\mathcal{E}}(u)\end{array}$$

C'est en fait une bijection, et même un isomorphisme de groupes.

Cela permet alors de définir naturellement

$$\mathrm{SU}(E) = \Phi^{-1}\big(\mathrm{SU}_n(\mathbb{C})\big) \qquad \text{et} \qquad \mathrm{O}^-(E) = \Phi^{-1}\big(\mathrm{U}_n^-(\mathbb{C})\big)$$

V.2 Endomorphismes remarquables d'un espace hermitien

Dans cette section, on se donne (H, φ) un espace hermitien, où H est de dimension finie n.

V.2.a Adjoint d'un endomorphisme

Fait V.2.1 (Dualité du produit scalaire)**.** *Pour toute forme linéaire $f \in H^*$, il existe un unique $x = x_f \in H$ tel que $\varphi(x_f, \cdot) = f$.*

Démonstration. Comme φ est non-dégénérée, on a un morphisme semi-linéaire injectif $\varphi_g : H \to H^*$ donné par $x \mapsto \varphi(x, \cdot)$. Comme H et H^* sont des $\mathbb{R}$-espaces vectoriels de même dimension $2n$, c'est donc un isomorphisme entre $\mathbb{R}$-espaces vectoriels donc, en particulier, une bijection. □

Proposition-définition V.2.2. Soit $u \in \mathrm{End}(H)$. Il existe un unique endomorphisme $u^* \in \mathrm{End}(H)$ tel que

$$\forall x, y \in H,\ \varphi(u^*(x), y) = \varphi(x, u(y)).$$

On l'appelle *endomorphisme adjoint* de u.

Démonstration. Matriciellement, on a $A = \mathrm{Mat}(\varphi, \mathcal{B}) \in \mathrm{GL}_n(\mathbb{C})$ car φ est non-dégénérée. Donc

$$\mathrm{Mat}_{\mathcal{B}}(u^*) = (A^{-1})^* \,\mathrm{Mat}_{\mathcal{B}}(u)^* A^*$$

convient. □

Corollaire V.2.3. *Soit $u, v \in \mathrm{End}(H)$ et $\lambda \in \mathbb{C}$. On a les égalités :*

1. $(\lambda u)^* = \overline{\lambda} u^*$;
2. $(u_1 + u_2)^* = u_1^* + u_2^*$;
3. $(u_1 \circ u_2)^* = u_2^* \circ u_1^*$;
4. $(u^*)^{-1} = (u^{-1})^*$.
5. $(u^*)^* = u$.

Démonstration. Preuve analogue au cas euclidien. □

Définition V.2.4 (Endomorphismes remarquables)**.** Soit $u \in \mathrm{End}(H)$. On dit que u est :

- *normal* si $u^* \circ u = u \circ u^*$;
- *hermitien* si $u^* = u$;
- *antihermitien* si $u^* = -u$;
- *unitaire* si $u^* \circ u = \mathrm{id}_V$.

Fait V.2.5. *Les endomorphismes hermitiens, antihermitiens, unitaires sont normaux.*

Remarque V.2.6. On utilise aussi la terminologie *endomorphisme auto-adjoint* au lieu de hermitien, et *isométrie* au lieu de unitaire.

Définition V.2.7. L'ensemble $\mathcal{U}(\varphi)$ (resp. $\mathcal{O}(\varphi)$) des endomorphismes unitaires pour φ est un sous-groupe de $\mathrm{GL}(V)$ appelé *groupe unitaire*, c'est-à-dire que la composée d'endomorphismes unitaires et l'inverse d'un endomorphisme unitaire sont des endomorphismes unitaires.

On note $\mathcal{SU}(\varphi)$ le sous-groupe des endomorphismes unitaires de déterminant 1, appelé *groupe spécial unitaire.*

V.2.b Réduction des endomorphismes normaux

Lemme V.2.8. *Si u est un endomorphisme normal et $P \in \mathbb{C}[X]$, alors $P(u)$ est normal.*

Démonstration. $P(u)^* = \overline{P}(u^*)$. Donc si u et u^* commutent, alors $P(u)$ et $P(u^*)$ aussi. □

Proposition V.2.9. *Si u est un endomorphisme normal et $P \in \mathbb{C}[X]$, alors $(\ker P(u))^{\perp}$ est stable par u.*

Démonstration. Soit $y \in (\ker P(u))^{\perp}$ et $x \in \ker P(u)$. On veut montrer que $\varphi(x, u(y)) = 0$. On a $\varphi(x, u(y)) = \varphi(u^*(x), y)$. Mais $P(u) \circ u^*(x) = u^* \circ P(u)(x) = 0$. Donc $u^*(x) \in \ker P(u)$. Donc $\varphi(u^*(x), y) = 0$. □

Proposition V.2.10. *On a* $\ker u^* = (\mathrm{im}\, u)^{\perp}$ *et* $\mathrm{im}\,(u^*) = (\ker u)^{\perp}$.

Démonstration.
$$\begin{aligned} y \in \ker u^* &\Leftrightarrow \forall x \in V,\ \varphi(u^*(y), x) = 0 \\ &\Leftrightarrow \forall x \in V,\ \varphi(y, u(x)) = 0 \\ &\Leftrightarrow y \in (\operatorname{im} u)^{\perp} \end{aligned}$$

En échangeant u et u^*, on a $\ker u = (\operatorname{im} u^*)^{\perp}$. On conclut car, comme φ est anisotrope donc non-dégénérée, en dimension finie on a $\left(W^{\perp}\right)^{\perp} = W$ pour tout sous-espace vectoriel W de V. □

Proposition V.2.11. *Si u est normal, alors* $\operatorname{im} u = (\ker u)^{\perp}$.

Démonstration. Soit $x \in (\operatorname{im} u)^{\perp}$. On a

$$\begin{aligned} \varphi(u(x), u(x)) &= \varphi(u^*(u(x)), x) \\ &= \varphi(u(u^*(x)), x) && \text{car } u \text{ est normal} \\ &= 0 && \text{car } x \in (\operatorname{im} u)^{\perp} \end{aligned}$$

Comme φ est anisotrope, on a $u(x) = 0$ donc $x \in \ker u$. Ainsi $(\operatorname{im} u)^{\perp} \subset \ker u$. On a égalité par égalité des dimensions en appliquant le théorème du rang : $\dim \operatorname{im} u = \dim V - \dim \ker u = \dim (\ker u)^{\perp}$. □

Théorème V.2.12 (Théorème spectral des endomorphismes normaux)**.** *Soit u un endomorphisme normal et*

$$\chi_u = \prod_{i=1}^{m} P_i^{m_i} = \prod_{i=1}^{m} (X - \lambda_i)^{m_i}$$

la décomposition du polynôme caractéristique de u en puissances de polynômes irréductibles unitaires deux à deux premiers entre eux, c'est-à-dire que les $\lambda_i \in \mathbb{C}$ sont deux à deux distincts. Alors

$$V = \overset{\perp}{\bigoplus_{i \in [\![1,m]\!]}} \ker(u - \lambda_i \operatorname{id}_H).$$

En particulier, u est diagonalisable en base orthogonale pour φ.

Démonstration. On a vu que $P_i(u)$ est normal et que $\operatorname{im} P_i(u) = (\ker P_i(u))^{\perp}$. De plus, $\operatorname{im} P_i(u) \cap \ker P_i(u) = 0$ car φ est anisotrope. Donc on en déduit par récurrence immédiate que $\ker P_i(u)^j = \ker P_i(u)$ pour tous $i, j \geq 1$. Le lemme des noyaux nous donne $V = \bigoplus_{i \in [\![1,m]\!]} \ker P_i(u)^{m_i} = \bigoplus_{i \in [\![1,m]\!]} \ker P_i(u)$. Il suffit de vérifier que la somme directe est orthogonale.

Par Bézout, il existe des polynômes $U, V \in C[X]$ tels que $UP_i + VP_j = 1$. Soit $x \in \ker P_j(u)$ qu'on écrit $x = U(u) \circ P_i(u)(x) + V(u) \circ P_j(u)(x) = P_i(u) \circ U(u)(x) \in \operatorname{im} P_i(u) = (\ker P_i(u))^{\perp}$. Ainsi $\ker P_j(u) \subset (\ker P_i(u))^{\perp}$, donc la somme directe est orthogonale. □

Corollaire V.2.13 (Théorème spectral hermitien)**.**
Soit $u \in \operatorname{End}(H)$. Alors

1. *u est normal si, et seulement si, u est diagonalisable en base φ-orthonormée ;*
2. *u est φ-autoadjoint si, et seulement si, u est diagonalisable en base φ-orthonormée et ses valeurs propres sont réelles ;*
3. *u est unitaire si, et seulement si, u est diagonalisable en base φ-orthonormée et ses valeurs propres sont de module* 1*.*

Démonstration. Pour les sens directs, il suffit d'étudier le spectre complexe de u dans chaque cas.

Si u est auto-adjoint ou unitaire, on sait que u est normal donc il existe une base φ-orthonormée $\mathcal{B}$ qui diagonalise u. Notons

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \operatorname{Mat}_{\mathcal{B}}(u)$$

avec $\lambda_i \in \mathbb{C}$ qui sont les valeurs propres de u.

Si u est auto-adjoint, alors $D^* = D$. Donc pour tout $1 \leqslant i \leqslant n$, on a $\lambda_i = \overline{\lambda_i}$. Donc les valeurs propres de u sont réelles.

Si u est unitaire, alors $D^*D = I_n$. Donc pour tout $1 \leqslant i \leqslant n$, on a $\lambda_i \overline{\lambda_i} = |\lambda_i|^2 = 1$. Donc les valeurs propres de u sont de module 1.

Ce qui démontre les sens directs.

Pour la réciproque, il suffit de remarque que si u s'écrit

$$D = \operatorname{Mat}_{\mathcal{B}}(u) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

en base φ-orthonormée $\mathcal{B}$, alors

— dans tous les cas, on a $D^*D = DD^*$ donc u est normal ;

— si les λ_i sont réels, on a $D^* = D$ donc u est auto-adjoint ;

— si les λ_i sont de module 1, on a $D^*D = I_n$ donc u est unitaire.

$\square$

Le théorème spectral nous dit, en particulier, que si u est auto-adjoint, alors $\mathrm{Sp}_{\mathbb{C}}(u) \subset \mathbb{R}$.
Dans $\mathbb{R}$, la notion de signe prend sens, contrairement à $\mathbb{C}$.

Définition V.2.14. On dit qu'un endomorphisme u auto-adjoint est :

— *positif* si $\mathrm{Sp}_{\mathbb{C}}(u) \subset \mathbb{R}_+ = [0, +\infty[$;

— *défini positif* si $\mathrm{Sp}_{\mathbb{C}}(u) \subset \mathbb{R}_+^* =]0, +\infty[$.

Corollaire V.2.15 (Théorème spectral hermitien matriciel)**.** *Soit $A \in \mathcal{H}_n(\mathbb{C})$. Alors il existe une matrice $U \in \mathrm{U}_n(\mathbb{C})$ et une matrice diagonale réelle D telles que $A = U^*DU$.*

Notation V.2.16. Dans l'écriture précédente, les coefficients de D correspondent au spectre de A avec multiplicité des valeurs propres. On note alors

— $\mathcal{H}_n^+(\mathbb{C})$ l'ensemble des matrices hermitiennes $A \in \mathcal{H}_n(\mathbb{C})$ telles que $\mathrm{Sp}_{\mathbb{C}}(A) \subset [0, +\infty[$;

— $\mathcal{H}_n^{++}(\mathbb{C})$ l'ensemble des matrices hermitiennes $A \in \mathcal{H}_n(\mathbb{C})$ telles que $\mathrm{Sp}_{\mathbb{C}}(A) \subset]0, +\infty[$.

On retrouve également l'ensemble des résultats qu'on a étudié pour les espaces euclidiens à partir du théorème spectral hermitien, et même un peu davantage :

Corollaire V.2.17 (Théorème spectral euclidien)**.**
On suppose que (V, φ) est un espace euclidien. Soit $u \in \mathrm{End}_{\mathbb{R}}(V)$. Alors

$$u \text{ est symétrique} \iff u \text{ est diagonalisable en base } \varphi\text{-orthonormée},$$

$$u \text{ est normal} \iff u \text{ est diagonalisable par blocs en base } \varphi\text{-orthonormée},$$

$$\text{avec des blocs de la forme } M_i = \begin{cases} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} & \text{avec } b \in \mathbb{R}^* \text{ et } a \in \mathbb{R} \\ (a) & \text{avec } a \in \mathbb{R} \end{cases}, \text{ et}$$

$$u \text{ est orthogonal} \iff u \text{ est diagonalisable par blocs en base } \varphi\text{-orthonormée}$$

$$\text{avec des blocs de la forme } M_i = \begin{cases} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} & \text{avec } \theta \in \mathbb{R} \\ (a) & \text{avec } a \in \{\pm 1\} \end{cases},$$

Les démonstrations de ces deux corollaires sont laissées en exercice au lecteur qui pourra alors s'assurer avoir bien compris le théorème de réduction des endomorphismes normaux.

That's all folks !