Université de Nantes

Algèbre linéaire et bilinéaire

Licence troisième année - Parcours Maths - X5M0040

2016-2017

Sylvain Gervais

Table des matières

Liste des notations

Dans ce cours, E sera toujours un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ (sur $\mathbb{R}$ dans les chapitres 3 et 4) de dimension finie n.

Notations prérequises

$\mathscr{L}(E, F)$	L'espace vectoriel des applications linéaires de l'espace vectoriel E vers l'espace vectoriel F.
$\mathscr{L}(E)$	Désigne $\mathscr{L}(E, E)$, l'espace vectoriel des endomorphismes de E ; c'est un espace vectoriel sur $\mathbb{K}$ de dimension n^2.
$\mathrm{GL}(E)$	L'ensemble des isomorphismes de E dans E.
Id_E	L'application *identité* de E.
$\mathscr{M}_n(\mathbb{K})$	L'ensemble des matrices carrées de taille $n \times n$ à coefficients dans $\mathbb{K}$; c'est un espace vectoriel sur $\mathbb{K}$ de dimension n^2.
$\mathrm{GL}_n(\mathbb{K})$	L'ensemble des matrices de taille $n \times n$ à coefficients dans $\mathbb{K}$, inversibles (*ie* de déterminant non nul).
I_n	La matrice *identité* de taille $n \times n$.
$[f]_b$	Matrice de l'endomorphisme f de E dans la base b. Rappelons que si l'on note $b = (e_1, \ldots, e_n)$ et $[f]_b = (a_{i,j})_{1 \leq i,j \leq n}$, alors la $j^{\text{ème}}$ colonne $\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$ est la colonne des coordonnées de $f(e_j)$ dans la base b : $$f(e_j) = \sum_{i=1}^{n} a_{i,j} e_i.$$ De plus, si $b' = (e'_1, \ldots, e'_n)$ est une deuxième base de E et P la matrice de passage de b à b' (c'est-à-dire que les colonnes de P sont les coordonnées de $e'_1, \ldots, e'_n$ dans la base b), alors $[f]_{b'} = P^{-1}AP$.
$\mathrm{Tr}(A)$	Trace de la matrice $A \in \mathscr{M}_n(\mathbb{K})$.
$\mathbb{K}[X]$	L'espace vectoriel des polynômes à coefficients dans $\mathbb{K}$.
$[\![p, q]\!]$	L'ensemble $\{p, p+1, \ldots, q-1, q\}$ pour deux entiers naturels p et q, $p < q$.

Notations introduites dans ce cours

χ_f	Polynôme caractéristique de l'endomorphisme f, p. 5
$m_f(\lambda)$	Multiplicité de la valeur propre λ pour l'endomorphisme f, p. 6
$C(P)$	Matrice compagnon du polynôme P, p. 6
I_f	Ensemble des polynômes annulateurs de l'endomorphisme f, p. 11
μ_f	Polynôme minimal de l'endomorphisme f, p. 11
I_A	Ensemble des polynômes annulateurs de la matrice A, p. 11
μ_A	Polynôme minimal de la matrice A, p. 11
f_F	Endomorphisme du sous-espace vectoriel F induit par f, p. 14
F_λ	Sous-espace caractéristique de l'endomorphisme f associé à la valeur propre λ, p. 16
E^*	Le dual de l'espace vectoriel E, p. 19
E^{**}	Le bidual de l'espace vectoriel E, c'est-à-dire le dual de E^*, p. 19
$b^* = (e_1^*, \dots, e_n^*)$	Base duale de la base $b = (e_1, \dots, e_n)$ de E, p. 19
${}^t u$	Application transposée de l'endomorphisme u, p. 21
F°	Orthogonal pour la dualité du sous-espace vectoriel F de E, p. 22
${}^\circ G$	Orthogonal pour la dualité du sous-espace vectoriel G de E^*, p. 22
$\mathscr{L}_2(E)$	Espace vectoriel des formes bilinéaires sur E, p. 25
$\mathscr{L}_2^s(E)$	Espace vectoriel des formes bilinéaires symétriques sur E, p. 25
$\mathcal{Q}(E)$	Espace vectoriel des formes quadratiques sur E, p. 26
$A^\perp$	Orthogonal (pour une forme quadratique donnée) de la partie A de E, p. 27
$\ker q$	Noyau de la forme quadratique q, p. 28
$\mathrm{rg}(q)$	Rang de la forme quadratique q, p. 28
$\mathrm{O}(n)$	Groupe orthogonal : ensemble des matrices orthogonales de taille $n \times n$, p. 36
f^*	Adjoint de l'endomorphisme f, p. 37
$\mathrm{O}(E)$	Groupe orthogonal de l'espace vectoriel euclidien E, p. 39
$\mathrm{SO}(E)$	Groupe spécial orthogonal de l'espace vectoriel euclidien E, p. 40
$\mathscr{S}(E)$	Sous-espace vectoriel des endomorphismes symétriques de E, p. 43
φ_f, q_f	Forme bilinéaire symétrique et forme quadratique associées à l'endomorphisme symétrique f, p. 44

1

Réduction des endomorphismes

I _ Rappels

a) Valeurs propres - Vecteurs propres

Définition *Soit f un endomorphisme de E.*

1) Un scalaire $\lambda \in \mathbb{K}$ est une valeur propre *de f s'il existe un élément non nul x de E tel que $f(x) = \lambda x$. Le spectre de f, noté $\mathrm{Sp}(f)$, est l'ensemble des valeurs propres de f.*

2) Un tel vecteur x est appelé vecteur propre *de f associé à (la valeur propre) λ.*

3) Le sous-espace propre *associé à λ, noté E_λ, est la réunion de $\{0\}$ et de l'ensemble des vecteurs propres associés à λ : c'est donc le sous-espace vectoriel $\ker(f - \lambda \mathrm{Id}_E)$.*

Lemme 1 *Les sous-espaces propres d'un endomorphisme f sont en somme directe : si $\lambda_1, \ldots, \lambda_k$ sont les valeurs propres deux à deux distinctes de f, alors toute famille $(x_1, \ldots, x_k)$ d'éléments non nuls de E, tels que $x_i \in E_{\lambda_i}$, est libre.*

Démonstration. On montre par récurrence la propriété $\mathcal{P}(l)$ suivante :

si $x_1, \ldots, x_l$ vérifient $x_i \in E_{\lambda_i}$ pour tout $i \in [\![1, l]\!]$ et $x_1 + \cdots + x_l = 0$, alors $x_i = 0$ pour tout $i \in [\![1, l]\!]$.

La propriété $\mathcal{P}(1)$ est clairement vérifiée. Montrons que $\mathcal{P}(l-1)$ implique $\mathcal{P}(l)$. On considère donc, pour $i \in [\![1, l]\!]$, $x_i \in E_{\lambda_i}$ tels que $x_1 + \cdots + x_l = 0$. En appliquant f à cette égalité, on obtient

$$0 = f(0) = f(x_1) + \cdots + f(x_l) = \lambda_1 x_1 + \cdots + \lambda_l x_l.$$

Combinée avec la première, cette dernière égalité implique $(\lambda_1 - \lambda_l)x_1 + \cdots + (\lambda_{l-1} - \lambda_l)x_{l-1} = 0$.

Si $y_i = (\lambda_i - \lambda_l)x_i$, alors $y_i \in E_{\lambda_i}$ et $y_1 + \cdots + y_{l-1} = 0$. Par hypothèse de récurrence, on obtient donc $y_i = 0$ pour chaque $i \in [\![1, l-1]\!]$, soit $(\lambda_i - \lambda_l)x_i = 0$; puisque $\lambda_i \neq \lambda_l$, ceci implique $x_i = 0$. Ainsi, $x_1 = \cdots = x_{l-1} = 0$, et alors $x_l = 0$.

□

Définition *Soit f un endomorphisme de E.*

1) f est dit diagonalisable *s'il existe une base de E formée de vecteurs propres, c'est-à-dire dans laquelle la matrice de f est diagonale.*

2) f est dit trigonalisable *s'il existe une base de E dans laquelle la matrice de f est triangulaire supérieure.*

Proposition 1.1 *Un endomorphisme f de E est diagonalisable si, et seulement si, $E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda$.*

Démonstration. Notons $\lambda_1, \dots, \lambda_k$ les valeurs propres deux à deux distinctes de f.

$\Longrightarrow$ Soit $(e_1, \dots, e_n)$ une base de E formée de vecteurs propres de f. On suppose que $e_1, \dots, e_{\alpha_1}$ sont éléments de E_{λ_1}, $e_{\alpha_1+1}, \dots, e_{\alpha_1+\alpha_2}$ sont éléments de E_{λ_2} etc. Ainsi, pour $i \in [\![1, k]\!]$, on a

$$\alpha_i = \dim \operatorname{Vect}\big(e_{\alpha_1+\cdots+\alpha_{i-1}+1}, \cdots, e_{\alpha_1+\cdots+\alpha_i}\big) \leq \dim E_{\lambda_i},$$

et donc

$$\dim E = n = \alpha_1 + \cdots + \alpha_k \leq \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = \dim\big(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}\big)$$

d'après le lemme 1. Par conséquent, $\dim\big(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}\big) = n$ et $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

$\Longleftarrow$ On suppose que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Alors, si $(e_1, \dots, e_n)$ est une base adaptée à cette décomposition, la matrice de f dans cette base est diagonale puisque, pour chaque $i \in [\![1, n]\!]$, $f(e_i) = \lambda e_i$ pour une certaine valeur propre λ de f.

□

Point de vue matriciel. Nous avons les définitions similaires suivantes pour $A \in \mathscr{M}_n(\mathbb{K})$:

(i) une valeur propre de A est un scalaire $\lambda \in \mathbb{K}$ tel qu'il existe $X \in \mathscr{M}_{n,1}(\mathbb{K}) \setminus \{0\}$ vérifiant $AX = \lambda X$;

(ii) A est dite diagonalisable s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit diagonale ;

(iii) A est dite trigonalisable s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit triangulaire supérieure.

Remarque : si $f \in \mathscr{L}(E)$ et b est une base de E, alors f est diagonalisable (resp. trigonalisable) si, et seulement si $[f]_b$ l'est.

b) Polynôme caractéristique

Définition *Si $A \in \mathscr{M}_n(\mathbb{K})$, on définit l'application $\chi_A : \mathbb{K} \to \mathbb{K}$ par $\chi_A(\lambda) = \det(A - \lambda \mathrm{I}_n)$.*

Proposition 1.2 *L'application χ_A est une fonction polynomiale de degré n ; plus précisément, on a*

$$\chi_A(\lambda) = \sum_{i=1}^{n} a_i \lambda^i \quad \textit{avec} \quad a_n = (-1)^n, \quad a_{n-1} = (-1)^{n-1}\mathrm{Tr}(A) \quad \textit{et} \quad a_0 = \det A.$$

Lemme 2 *Soient p et n deux entiers naturels vérifiant $1 \leq p \leq n$, $A_1, \dots, A_n$ n éléments de $\mathbb{K}^n$ et $B_1, \dots, B_p$ p éléments de $\mathbb{K}^n$.*
Alors, l'application de $\mathbb{K}$ dans $\mathbb{K}$, qui, à x associe $\det(A_1 + xB_1, \dots, A_p + xB_p, A_{p+1}, \dots, A_n)$, est polynomiale de degré inférieur ou égal à p, le coefficient de x^p étant égal à $\det(B_1, \dots, B_p, A_{p+1}, \dots, A_n)$.

Démonstration. On démontre ce lemme par récurrence sur p. Pour $p = 1$, on a

$$\det(A_1 + xB_1, A_2, \dots, A_n) = \det(A_1, A_2, \dots, A_n) + x \det(B_1, A_2, \dots, A_n)$$

ce qui montre la propriété dans ce cas. Supposons à présent la propriété démontrée au rang p et montrons qu'elle est vraie au rang $p + 1$. On a

$$\begin{aligned}\det(A_1 + xB_1, \dots, A_p + xB_p, A_{p+1} + xB_{p+1}, A_{p+2}, \dots, A_n) &= \det(A_1 + xB_1, \dots, A_p + xB_p, A_{p+1}, \dots, A_n) \\ &\quad + x \det(A_1 + xB_1, \dots, A_p + xB_p, B_{p+1}, A_{p+2}, \dots, A_n).\end{aligned}$$

Par hypothèse de récurrence, les deux déterminants du second membre sont polynomiaux en x, le premier de degré inférieur ou égal à p et le second de degré inférieur ou égal à $p+1$ avec pour coefficient de degré $p+1$ $\det(B_1, \dots, B_p, B_{p+1}, A_{p+2}, \dots, A_n)$. Ceci nous donne bien la propriété au rang $p + 1$.

□

Démonstration de la proposition 1.2. Notons $A_1, \dots, A_n$ les vecteurs colonnes de A et $E_1, \dots, E_n$ les vecteurs colonnes de I_n.

• On a, pour $\lambda \in \mathbb{K}$, $\chi_A(\lambda) = \det(A - \lambda \mathrm{I}_n) = \det(A_1 - \lambda E_1, \dots, A_n - \lambda E_n)$. Le lemme 2 montre donc que χ_A est un polynôme de degré n, le coefficient dominant étant égal à $\det(-E_1, \dots, -E_n) = (-1)^n$.

• Le coefficient constant de χ_A est $\chi_A(0) = \det A$.

• On montre par récurrence sur n que le coefficient du terme λ^{n-1} est égal à $(-1)^{n-1}\mathrm{Tr}(A)$.

Si $n = 1$, alors $A = (a)$ et $\chi_A(\lambda) = \det(A - \lambda \mathrm{I}_1) = a - \lambda$: $a = (-1)^{1-1}\mathrm{Tr}(A)$.
Supposons la propriété vraie au rang n. Écrivons $A \in \mathscr{M}_{n+1}(\mathbb{K})$ sous la forme

$$A = \left(\begin{array}{c|c} A' & A'_{n+1} \\ \hline b_1 \cdots b_n & b_{n+1} \end{array}\right)$$

avec $A' \in \mathscr{M}_n(\mathbb{K})$ et $A'_{n+1} \in \mathscr{M}_{n,1}(\mathbb{K})$. Si $A'_1, \dots, A'_n$ sont les vecteurs colonnes de A', $A_1, ..., A_{n+1}$ ceux de A et $C = \left(\begin{smallmatrix} A'_{n+1} \\ 0 \end{smallmatrix}\right)$, on a :

$$\begin{aligned} \chi_A(\lambda) &= \det(A_1 - \lambda E_1, \dots, A_{n+1} - \lambda E_{n+1}) \\ &= \underbrace{\det(A_1 - \lambda E_1, \dots, A_n - \lambda E_n, C)}_{\Delta_1} + (b_{n+1} - \lambda) \underbrace{\det(A_1 - \lambda E_1, \dots, A_n - \lambda E_n, E_{n+1})}_{\Delta_2} \end{aligned}$$

(où $E_1, \cdots, E_{n+1}$ sont les vecteurs colonnes de I_{n+1}). D'après le lemme 2, Δ_1 est un polynôme de degré inférieur ou égal à n, le coefficient de λ^n étant égal à $\det(-E_1, \dots, -E_n, C)$. Or ce dernier déterminant est nul puisque $C \in \mathrm{Vect}(E_1, \dots, E_n)$. Par conséquent, Δ_1 est de degré inférieur ou égal à $n - 1$. Ensuite, on a, si $E'_1, \cdots, E'_n$ sont les vecteurs colonnes de I_n :

$$\Delta_2 = \begin{vmatrix} A'_1 - \lambda E'_1 & \cdots & A'_n - \lambda E'_n & [0] \\ b_1 & \cdots & b_n & 1 \end{vmatrix} = \det(A'_1 - \lambda E'_1, \dots, A'_n - \lambda E'_n) = \chi_{A'}(\lambda).$$

Donc Δ_2 est un polynôme de degré n, son coefficient dominant étant égal à $(-1)^n$ et celui de degré $n-1$ égal à $(-1)^{n-1}\mathrm{Tr}(A')$ par hypothèse de récurrence. En conséquence, le terme de degré n de χ_A vaut

$$b_{n+1}(-1)^n - (-1)^{n-1}\mathrm{Tr}(A') = (-1)^n\Big(\mathrm{Tr}(A') + b_{n+1}\Big) = (-1)^n \mathrm{Tr}(A).$$

□

Définition *Le polynôme χ_A est appelé* polynôme caractéristique *de la matrice A.*

Pour définir la notion de polynôme caractéristique d'un endomorphisme, nous avons besoin du lemme suivant.

Lemme 3 *Pour toutes matrices $A \in \mathscr{M}_n(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$, on a $\chi_{P^{-1}AP} = \chi_A$.*

Démonstration. Pour tout λ dans $\mathbb{K}$, on a

$$\chi_{P^{-1}AP}(\lambda) = \det(P^{-1}AP - \lambda \mathrm{I}_n) = \det(P^{-1}AP - \lambda P^{-1}P) = \det\big(P^{-1}(A - \lambda \mathrm{I}_n)P\big) = \det(A - \lambda \mathrm{I}_n) = \chi_A(\lambda).$$

□

Ainsi, si f est un endomorphisme de E et b, b' deux bases de E, on a $\chi_{[f]_b} = \chi_{[f]_{b'}}$. Ceci permet de donner la définition suivante :

Définition *On appelle* polynôme caractéristique *d'un endomorphisme f de E, noté χ_f, le polynôme caractéristique de la matrice de f dans une base quelconque de E.*

Proposition 1.3 *Soit f un endomorphisme de E et λ un élément de $\mathbb{K}$. Alors :*

λ est valeur propre de f si, et seulement si, $\chi_f(\lambda) = 0$.

Remarque : bien sûr, il découle immédiatement de ce résultat et des définitions que λ est valeur propre d'une matrice carrée A si, et seulement si, $\chi_A(\lambda) = 0$.

Démonstration. Par définition, λ est valeur propre de f si, et seulement si, $\ker(f - \lambda \mathrm{Id}_E)$ n'est pas réduit à $\{0\}$. Par conséquent, on a :

$$\begin{aligned}
\lambda \in \mathrm{Sp}(f) &\Longleftrightarrow f - \lambda\mathrm{Id}_E \quad \text{non injectif} \\
&\Longleftrightarrow f - \lambda\mathrm{Id}_E \quad \text{non bijectif} \quad (\text{car } \dim E \text{ finie}) \\
&\Longleftrightarrow \det(f - \lambda\mathrm{Id}_E) = 0 \\
&\Longleftrightarrow \chi_f(\lambda) = 0.
\end{aligned}$$

□

Définition *Soit f un endomorphisme de E et $\lambda \in Sp(f)$. On appelle* multiplicité *de λ, notée $m_f(\lambda)$, sa multiplicité comme racine de χ_f :*

$$m_f(\lambda) = \max\left\{k \in \mathbb{N} \,/\, (X-\lambda)^k \textit{ divise } \chi_f\right\}.$$

Proposition 1.4 *Pour f endomorphisme de E et $\lambda \in \mathrm{Sp}(f)$, on a $1 \leq \dim E_\lambda \leq m_f(\lambda)$.*

Démonstration. Notons α la dimension de E_λ. Puisque λ est valeur propre, $E_\lambda \neq \{0\}$ donc $\alpha \geq 1$. Choisissons une base $(e_1, \ldots, e_\alpha)$ de E_λ, que l'on complète en une base $b = (e_1, \ldots, e_n)$ de E. Puisque $f(e_i) = \lambda e_i$ pour $i \in [\![1, \alpha]\!]$, la matrice de f dans b est de la forme

$$[f]_b = \left(\begin{array}{c|c} \lambda \mathrm{I}_\alpha & B \\ \hline 0 & C \end{array}\right)$$

et on a donc

$$\chi_f(X) = \det([f]_b - X\mathrm{I}_n) = \left|\begin{array}{c|c} \lambda \mathrm{I}_\alpha - X\mathrm{I}_\alpha & B \\ \hline 0 & C - X\mathrm{I}_{n-\alpha} \end{array}\right| = (\lambda - X)^\alpha \chi_C(X).$$

Ainsi, $(X-\lambda)^\alpha$ divise χ_f et $\alpha \leq m_f(\lambda)$.

□

Un exemple : matrice compagnon d'un polynôme unitaire.

Soit $P = X^n - a_{n-1}X^{n-1} - \cdots - a_1 X - a_0$ un polynôme unitaire de degré n à coefficients dans $\mathbb{K}$. On appelle *matrice compagnon* de P la matrice $C(P) \in \mathscr{M}_n(\mathbb{K})$ donnée par

$$C(P) = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & \vdots \\ \vdots \; \ddots & 0 & 1 & a_{n-1} \end{pmatrix}.$$

Proposition 1.5 *Pour tout polynôme unitaire $P \in \mathbb{K}[X]$, on a $\chi_{C(P)} = (-1)^n P$.*

Démonstration. On a

$$\chi_{C(P)}(\lambda) = \begin{vmatrix} -\lambda & 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & -\lambda & 0 & \vdots \\ \vdots & \ddots & 1 & -\lambda & a_{n-2} \\ 0 & \cdots & 0 & 1 & a_{n-1} - \lambda \end{vmatrix}.$$

En ajoutant à la première ligne la $i^{\text{ème}}$ multipliée par λ^{i-1}, ceci pour tout $i \in [\![2, n]\!]$, on obtient

$$\chi_{C(P)}(\lambda) = \begin{vmatrix} 0 & 0 & \cdots & \cdots & 0 & -P(\lambda) \\ 1 & -\lambda & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & 0 & \vdots \\ \vdots & & \ddots & 1 & -\lambda & a_{n-2} \\ 0 & \cdots & \cdots & 0 & 1 & a_{n-1} - \lambda \end{vmatrix}$$

ce qui conduit à

$$\chi_{C(P)}(\lambda) = -(-1)^{n-1} P(\lambda) \begin{vmatrix} 1 & -\lambda & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & 0 \\ \vdots & & \ddots & 1 & -\lambda \\ 0 & \cdots & \cdots & 0 & 1 \end{vmatrix} = (-1)^n P(\lambda).$$

□

Nous arrivons à présent aux premiers critères de trigonalisation et diagonalisation.

Théorème 1.1 *Un endomorphisme de E est trigonalisable si, et seulement si, son polynôme caractéristique est scindé sur $\mathbb{K}$.*
De même, une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est trigonalisable si, et seulement si, χ_A est scindé sur $\mathbb{K}$.

Démonstration. $\Longrightarrow$ Supposons qu'il existe une base b de E dans laquelle la matrice de l'endomorphisme f est triangulaire supérieure. Cette matrice étant de la forme

$$[f]_b = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

son polynôme caractéristique vaut

$$\chi_f(X) = \det\Big([f]_b - X\mathrm{I}_n\Big) = \det \begin{pmatrix} \lambda_1 - X & & * \\ & \ddots & \\ 0 & & \lambda_n - X \end{pmatrix} = \prod_{i=1}^{n} (\lambda_i - X).$$

Ainsi, le polynôme caractéristique de f est scindé sur $\mathbb{K}$.

$\Longleftarrow$ On montre par récurrence sur $n = \dim E$ la propriété $\mathcal{P}(n)$: tout endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension n dont le polynôme caractéristique est scindé sur $\mathbb{K}$ est trigonalisable.
$\mathcal{P}(1)$ est vraie car tout endomorphisme est trigonalisable (même diagonalisable) en dimension 1. Supposons donc que $\mathcal{P}(n)$ soit vraie et considérons un endomorphisme f d'un espace vectoriel E de dimension $n+1$ tel que χ_f soit scindé sur $\mathbb{K}$.

• Puisque χ_f est scindé, il admet au moins une racine λ dans $\mathbb{K}$. D'après la proposition 1.3, λ est une valeur propre de f : il existe donc e_1 non nul dans E tel que $f(e_1) = \lambda e_1$. Complétons ce vecteur en une base

$b = (e_1, e_2, ..., e_{n+1})$ de E et notons F le sous-espace vectoriel de E engendré par $\{e_2, \ldots, e_{n+1}\}$. On a :

$$[f]_b = \left(\begin{array}{c|c} \lambda & L \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \quad \text{et} \quad \chi_f(X) = (\lambda - X)\chi_B(X).$$

Puisque χ_f est scindé sur $\mathbb{K}$, il en est de même pour χ_B.

Par hypothèse de récurrence, la matrice B est trigonalisable : il existe $P \in \mathrm{GL}_n(\mathbb{K})$ tel que $T := P^{-1}BP$ soit triangulaire supérieure. Alors, si Q désigne la matrice $(n+1) \times (n+1)$ définie par $Q = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right)$, Q est inversible et

$$Q^{-1}[f]_b Q = \left(\begin{array}{c|c} \lambda & L' \\ \hline 0 & \\ \vdots & T \\ 0 & \end{array}\right) \quad \text{est triangulaire supérieure,}$$

ce qui montre que f est trigonalisable.

□

Remarque : nous venons de voir que si $A = (a_{i,j})_{1 \leq i,j \leq n}$ est une matrice triangulaire, son polynôme caractéristique vaut $\chi_A(X) = \prod_{i=1}^{n}(a_{i,i} - X)$.

Puisque tout polynôme à coefficients complexes est scindé sur $\mathbb{C}$, on déduit du théorème 1.1

Corollaire 1.1 *Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie est trigonalisable ; toute matrice à coefficients complexes est trigonalisable sur $\mathbb{C}$.*

Théorème 1.2 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Un endomorphisme f de E est diagonalisable si, et seulement si, les deux assertions suivantes sont satisfaites :*

(i) χ_f *est scindé sur* $\mathbb{K}$ *;*

(ii) pour toute valeur propre λ de f, $\dim E_\lambda = m_f(\lambda)$.

Démonstration. $\Longrightarrow$ Si f est diagonalisable, alors E est somme directe des sous-espaces propres (proposition 1.1). Par conséquent, en notant $b = (e_1, \ldots, e_n)$ une base de E adaptée à cette décomposition, $\lambda_1, \ldots, \lambda_k$ les valeurs propres deux à deux distinctes de f et α_i la dimension du sous-espace propre E_{λ_i} ($i \in [\![1, k]\!]$), la matrice de f dans b est de la forme

$$[f]_b = \left(\begin{array}{c|c|c} \lambda_1 \mathrm{I}_{\alpha_1} & (0) & (0) \\ \hline (0) & \ddots & (0) \\ \hline (0) & (0) & \lambda_k \mathrm{I}_{\alpha_k} \end{array}\right).$$

On en déduit que $\chi_f(X) = \prod_{i=1}^{k} (\lambda_i - X)^{\alpha_i}$ est scindé sur $\mathbb{K}$ avec, pour $i \in [\![1, k]\!]$, $m_f(\lambda_i) = \alpha_i = \dim E_{\lambda_i}$.

$\Longleftarrow$ D'après le lemme 1, les sous-espaces propres de f sont en somme directe. On a donc, si $\lambda_1, \ldots, \lambda_k$ désignent les valeurs propres deux à deux distinctes de f :

$$\begin{aligned} \dim\left(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}\right) &= \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} \\ &= m_f(\lambda_1) + \cdots + m_f(\lambda_k) \quad \text{par hypothèse} \\ &= \text{degré}(\chi_f) \quad \text{car } \chi_f \text{ est scindé} \\ &= n, \end{aligned}$$

donc E est somme directe des sous-espaces propres de f. Par conséquent, f est diagonalisable d'après la proposition 1.1.

□

II _ Polynômes d'endomorphismes - Polynôme minimal

a) Calcul dans $\mathscr{L}(E)$

Rappelons que l'on peut munir $\mathscr{L}(E)$, l'ensemble des endomorphismes de E, des trois opérations suivantes :

- somme : pour $f, g \in \mathscr{L}(E)$, $(f+g) : E \to E$ est défini par $(f+g)(x) = f(x) + g(x)$;
- produit extérieur : si $f \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$, $(\lambda f) : E \to E$ est défini par $(\lambda f)(x) = \lambda\, f(x)$;
- composition : pour $f, g \in \mathscr{L}(E)$, $f \circ g : E \to E$ est défini par $(f \circ g)(x) = f\Big(g(x)\Big)$.

Les deux premières font de $\mathscr{L}(E)$ un $\mathbb{K}$-espace vectoriel ; il est de dimension finie égale à n^2.

Notation : pour $f \in \mathscr{L}(E)$, on note $f^0 = \mathrm{Id}_E$, $f^1 = f$, $f^2 = f \circ f$, ..., $f^k = \underbrace{f \circ f \circ \cdots \circ f}_{k \text{ fois}}$ $(k \in \mathbb{N}^*)$.

Bien sûr, on a $f^k \circ f^l = f^{k+l}$.

Ainsi, si $P = \sum\limits_{i=0}^{d} a_i X^i = a_0 + a_1 X + \cdots + a_d X^d$ est un polynôme à coefficients dans $\mathbb{K}$, on peut définir $P(f)$ en posant

$$P(f) = \sum_{i=0}^{d} a_i f^i = a_0 \mathrm{Id}_E + a_1 f + \cdots + a_d f^d.$$

$P(f)$ est un endomorphisme de E.

Définition *Soient f et g deux endomorphismes de E. On dit que g est un* polynôme en f *s'il existe $P \in \mathbb{K}[X]$ tel que $g = P(f)$.*

Remarque : on peut faire la même chose avec les matrices. Si $P = \sum\limits_{i=0}^{d} a_i X^i \in \mathbb{K}[X]$ et $A \in \mathscr{M}_n(\mathbb{K})$ alors

$$P(A) = \sum_{i=0}^{d} a_i A^i = a_0 \mathrm{I}_n + a_1 A + a_2 A^2 + \cdots + a_d A^d.$$

Proposition 1.6 *Si b est une base de E, f un endomorphisme de E, $A = [f]_b$ la matrice de f dans b et P un élément de $\mathbb{K}[X]$, alors $\big[P(f)\big]_b = P(A)$.*

Démonstration. C'est une conséquence directe du fait que, si b est une base de E, $f, g \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$, alors

$$[f+g]_b = [f]_b + [g]_b, \qquad [\lambda f]_b = \lambda [f]_b \quad \text{et} \quad [f \circ g]_b = [f]_b \cdot [g]_b.$$

$\square$

Proposition 1.7 *Pour tout $f \in \mathscr{L}(E)$, $\lambda \in \mathbb{K}$ et $(P, Q) \in \mathbb{K}[X]^2$, on a*

(i) $(P+Q)(f) = P(f) + Q(f)$;

(ii) $(\lambda P)(f) = \lambda\big(P(f)\big)$;

(iii) $(PQ)(f) = P(f) \circ Q(f)$.

Démonstration. Notons $P = \sum a_i X^i$ et $Q = \sum b_i X^i$. Alors, $P + Q = \sum (a_i + b_i) X^i$, $\lambda P = \sum (\lambda a_i) X^i$, et on obtient :

$$(P+Q)(f) = \sum (a_i + b_i) f^i = \sum \big(a_i f^i + b_i f^i\big) = \sum a_i f^i + \sum b_i f^i = P(f) + Q(f) \ ;$$

$$(\lambda P)(f) = \sum (\lambda a_i) f^i = \lambda \sum a_i f^i = \lambda\big(P(f)\big).$$

D'autre part, on a $PQ = \sum c_i X^i$ où $c_i = \sum_{j+k=i} a_j b_k$ et :

$$\begin{aligned} P(f) \circ Q(f) &= \left(\sum a_i f^i\right) \circ \left(\sum b_j f^j\right) = \sum_{i,j} a_i b_j f^i \circ f^j = \sum_{i,j} a_i b_j f^{i+j} \\ &= \sum_k \Big(\sum_{i+j=k} a_i b_j\Big) f^k = \sum_k c_k f^k \\ &= (PQ)(f). \end{aligned}$$

$\square$

Corollaire 1.2 *Pour tout endomorphime f de E et tout couple (P, Q) de polynômes à coefficients dans $\mathbb{K}$, $P(f)$ et $Q(f)$ commutent : $P(f) \circ Q(f) = Q(f) \circ P(f)$.*
En particulier, f commute avec $P(f)$ pour tout polynôme P.

Démonstration. D'après la proposition 1.7, on a $P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f)$. Le cas particulier s'obtient en prenant $Q = X$.

$\square$

Proposition 1.8 *Soient f et g deux endomorphismes de E tels que $f \circ g = g \circ f$, et $P \in \mathbb{K}[X]$.*
Alors $f \circ P(g) = P(g) \circ f$.

Démonstration. • On montre que $f \circ g^k = g^k \circ f$ pour tout $k \in \mathbb{N}$ par récurrence sur k, cette propriété étant clairement vraie pour $k = 0$ ($g^0 = \mathrm{Id}_E$) et $k = 1$ (c'est l'hypothèse). Si $f \circ g^k = g^k \circ f$, alors

$$f \circ g^{k+1} = (f \circ g^k) \circ g = (g^k \circ f) \circ g = g^k \circ (f \circ g) = g^k \circ (g \circ f) = g^{k+1} \circ f.$$

• Si $P = \sum a_i X^i$, alors

$$f \circ P(g) = f \circ \left(\sum a_i g^i\right) = \sum a_i (f \circ g^i) = \sum a_i (g^i \circ f) = \left(\sum a_i g^i\right) \circ f = P(g) \circ f.$$

$\square$

Proposition 1.9 *Soient f un endomorphisme de E, P un élément de $\mathbb{K}[X]$ et x un vecteur propre de f pour la valeur propre λ.*
Alors, x est un vecteur propre de $P(f)$ pour la valeur propre $P(\lambda)$.

Démonstration. • Remarquons d'abord que pour tout $k \in \mathbb{N}$, on a $f^k(x) = \lambda^k x$. Cette propriété se montre par récurrence sur k :

- ∗ pour $k = 0$, on a $f^0(x) = \mathrm{Id}_E(x) = x = \lambda^0 x$;
- ∗ si $f^k(x) = \lambda^k x$, alors $f^{k+1}(x) = f\big(f^k(x)\big) = f(\lambda^k x) = \lambda^k f(x) = \lambda^k(\lambda x) = \lambda^{k+1} x$.

• Maintenant, si $P = \sum_{i=0}^{d} a_i X^i$, on obtient :

$$\big(P(f)\big)(x) = \sum_{i=0}^{d} a_i f^i(x) = \sum_{i=0}^{d} a_i \lambda^i x = \left(\sum_{i=0}^{d} a_i \lambda^i\right) \cdot x = P(\lambda) \cdot x.$$

$\square$

b) Polynômes annulateurs

Définition *Soit f un endomorphisme de E. On dit qu'un polynôme $P \in \mathbb{K}[X]$* annule *f si l'endomorphisme $P(f)$ est nul, c'est-à-dire que pour tout x dans E, $P(f)(x) = 0$.*
On notera I_f l'ensemble des polynômes de $\mathbb{K}[X]$ annulateurs de f.

Remarque : notons que I_f n'est jamais vide, ni même réduit à $\{0\}$. En effet, $\mathscr{L}(E)$ étant de dimension finie égale à n^2, la famille $(f^0 = \mathrm{Id}_E, f, f^2, \ldots, f^{n^2})$ est liée. Il existe donc n^2+1 éléments non tous nuls $a_0, a_1, \ldots, a_{n^2}$ de $\mathbb{K}$ tels que $\sum_{i=0}^{n^2} a_i f^i = 0$; le polynôme $P = \sum_{i=0}^{n^2} a_i X^i$ est donc un polynôme non nul annulateur de f.

Exemple. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorphisme défini par $f(x,y) = (-x,y)$. On $(f \circ f)(x,y) = f(-x,y) = (x,y)$ donc $f \circ f = \mathrm{Id}_{\mathbb{R}^2}$. Par conséquent, si $P = X^2 - 1$, alors $P(f) = f^2 - \mathrm{Id}_{\mathbb{R}^2} = f \circ f - \mathrm{Id}_{\mathbb{R}^2} = 0$.

Lemme 4 *Soit f un endomorphisme de E. Alors :*

(i) I_f *est un sous-espace vectoriel de* $\mathbb{K}[X]$ *;*

(ii) $\forall P \in I_f$, $\forall Q \in \mathbb{K}[X]$, *on a* $PQ \in I_f$.

Démonstration. *(i)* Si P, Q sont deux éléments de I_f et λ un élément de $\mathbb{K}$, alors

$$(P + \lambda Q)(f) = P(f) + \lambda Q(f) = 0$$

donc $P + \lambda Q \in I_f$.
(ii) Pour $P \in I_f$ et $Q \in \mathbb{K}[X]$, on a $(PQ)(f) = (QP)(f) = Q(f) \circ P(f) = 0$ puisque $P(f) = 0$; donc $PQ \in I_f$.
□

Théorème 1.3 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et f un endomorphisme de E. Il existe un unique polynôme unitaire $\mu_f \in \mathbb{K}[X]$ vérifiant, pour tout $P \in \mathbb{K}[X]$:*

$$P \in I_f \textit{ si, et seulement si, } P \textit{ est un multiple de } \mu_f.$$

Démonstration. Nous avons vu que I_f n'est pas réduit à $\{0\}$. Par conséquent, l'ensemble

$$\left\{\deg P \,/\, P \in I_f \setminus \{0\}\right\}$$

est une partie non vide de $\mathbb{N}$ et admet donc un plus petit élément. Ceci montre qu'il existe donc dans I_f un polynôme unitaire de degré minimal : notons μ un tel polynôme.

• Puisque μ est dans I_f, il en est de même de tout multiple de μ d'après le lemme 4.

• Pour $P \in I_f$, on effectue la division euclidienne de P par μ : $P = \mu Q + R$ avec $R = 0$ ou $d^\circ R < d^\circ \mu$. Mais $\mu \in I_f$ donc $\mu Q \in I_f$ et, puisque $P \in I_f$, $R = P - \mu Q \in I_f$. Par conséquent, si $R \neq 0$, nous disposons dans I_f d'un élément de degré strictement inférieur au degré de μ : c'est impossible. Par conséquent, $R = 0$ et P est un multiple de μ.

• Soit à présent $\mu' \in \mathbb{K}[X]$ unitaire tel qu'un polynôme P soit élément de I_f si, et seulement, μ' divise P. Alors, puisque μ et μ' sont éléments de I_f, μ divise μ' et μ' divise μ donc $\mu' = \mu$ car ces deux polynômes sont unitaires.

□

Définition *L'unique polynôme μ_f ainsi défini est appelé* polynôme minimal *de f.*

Corollaire 1.3 *On a donc $I_f = \mu_f \mathbb{K}[X] = \left\{\mu_f Q \,/\, Q \in \mathbb{K}[X]\right\}$ (ensemble des multiples de μ_f). En particulier, μ_f annule f.*

Remarque : μ_f est de degré au moins 1 car si μ est constant, $\mu(f) = \mu \mathrm{Id}_E$.

Exemples : $\mu_{\mathrm{Id}_E} = X - 1$, et plus généralement, pour $\lambda \in \mathbb{K}$, $\mu_{\lambda \mathrm{Id}_E} = X - \lambda$.
Si $f : \mathbb{R}^2 \to \mathbb{R}^2$ est l'endomorphisme défini par $f(x,y) = (-x,y)$, on a $\mu_f = X^2 - 1$.

Il existe une version matricielle de ce qui précède. Un polynôme $P \in \mathbb{K}[X]$ annule une matrice $A \in \mathscr{M}_n(\mathbb{K})$ si $P(A) = 0$. On note I_A l'ensemble des polynômes annulateurs de A : il vérifie les propriétés du lemme 4. De même que pour un endomorphisme, il existe un unique polynôme unitaire $\mu_A \in \mathbb{K}[X]$, appelé *polynôme minimal* de A, tel que I_A soit exactement l'ensemble des multiples de μ_A. Ces deux versions (endomorphismes et matrices) coïncident au sens suivant :

Proposition 1.10 *Si b est une base de E et f un endomorphisme de E, alors $I_f = I_{[f]_b}$. En particulier, f et $[f]_b$ ont même polynôme minimal.*

Démonstration. C'est une conséquence directe du fait que, pour tout $P \in \mathbb{K}[X]$, on a $P\big([f]_b\big) = \big[P(f)\big]_b$ (proposition 1.6). En effet, cette propriété montre que $P \in \mathbb{K}[X]$ annule f si, et seulement si, P annule $[f]_b$. □

c) Liens avec les valeurs propres

Lemme 5 *Soit f un endomorphisme de E, λ une valeur propre de f et $P \in \mathbb{K}[X]$.*
Si $P(f) = 0$, alors $P(\lambda) = 0$.

Démonstration. D'après la proposition 1.9, $P(\lambda)$ est une valeur propre de $P(f)$. Par conséquent, si $P(f) = 0$, alors $P(\lambda) = 0$.

□

Théorème 1.4 *Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie et λ un élément de $\mathbb{K}$. Alors :*

$$\lambda \text{ est une valeur propre de } f \text{ si, et seulement si, } \mu_f(\lambda) = 0.$$

Remarque : nous disposons bien sûr de la version matricielle de ce théorème. Pour $A \in \mathscr{M}_n(\mathbb{K})$, $\mathrm{Sp}(A)$ coïncide avec l'ensemble des racines de μ_A dans $\mathbb{K}$.

Démonstration. $\Longrightarrow$ Puisque $\mu_f(f) = 0$, le lemme 5 montre que si $\lambda \in \mathrm{Sp}(f)$, alors $\mu_f(\lambda) = 0$.

$\Longleftarrow$ Soit $\lambda \in \mathbb{K}$ une racine de μ_f : il existe $Q \in \mathbb{K}[X]$ tel que $\mu_f(X) = (X - \lambda)Q(X)$. Puisque $d^\circ Q = d^\circ \mu_f - 1$, μ_f ne divise pas Q donc $Q(f) \neq 0$: il existe un élément non nul x de E tel que $Q(f)(x) \neq 0$. Si $y = Q(f)(x)$, alors :

$$0 = \mu_f(f)(x) = (f - \lambda \mathrm{Id}_E) \circ Q(f)(x) = (f - \lambda \mathrm{Id}_E)(y) = f(y) - \lambda y$$

donc λ est valeur propre de f.

□

Théorème 1.5 (Cayley[1]-Hamilton[2]) *Pour tout endomorphisme f d'un espace vectoriel de dimension finie, le polynôme caractéristique de f annule f : $\chi_f(f) = 0$, ou encore : μ_f divise χ_f.*

Démonstration. • Soit x un élément de E et montrons que $\chi_f(f)(x) = 0$. Puisque $\chi_f(f)(0) = 0$, on suppose dans ce qui suit que $x \neq 0$. L'espace vectoriel E étant de dimension finie, il existe $k \in [\![0, n-1]\!]$ tel que $\big(x, f(x), \ldots, f^k(x)\big)$ soit libre et $\big(x, f(x), \ldots, f^{k+1}(x)\big)$ soit liée. Complétons la famille $\big(x, f(x), \ldots, f^k(x)\big)$ en une base b de E. Déterminons la forme de la matrice de f dans cette base :

(i) si $k \geq 1$, on a, pour $i \in [\![0, k-1]\!]$, $f\big(f^i(x)\big) = f^{i+1}(x)$;

(ii) puisque $\big(x, f(x), \ldots, f^k(x)\big)$ est libre et $\big(x, f(x), \ldots, f^{k+1}(x)\big)$ liée, il existe $(a_0, a_1, \ldots, a_k) \in \mathbb{K}^{k+1}$ tel que $f\big(f^k(x)\big) = f^{k+1}(x) = a_0 x + a_1 f(x) + \cdots + a_k f^k(x)$.

La matrice $[f]_b$ est donc de la forme

$$[f]_b = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right) \quad \text{où} \quad A = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots\, 0 & 1 & a_k \end{pmatrix} \in \mathcal{M}_{k+1}(\mathbb{K}),$$

ce qui donne $\chi_f = \chi_A \chi_C$ et donc $\chi_f(f)(x) = \chi_A(f) \circ \chi_C(f)(x) = \chi_C(f) \circ \chi_A(f)(x)$.

[1] Arthur Cayley, mathématicien britannique, 1821-1895.
[2] William Rowan Hamilton, mathématicien, physicien et astronome irlandais, 1805-1865.

• Pour conclure, nous allons montrer que $\chi_A(f)(x) = 0$. La matrice A est la matrice compagnion du polynôme $P = X^{k+1} - a_k X^k - \cdots - a_0$ donc $\chi_A = (-1)^{k+1} P$ (proposition 1.5). Il suffit donc de démontrer que $P(f)(x) = 0$. Or

$$P(f)(x) = f^{k+1}(x) - a_k f^k(x) - \cdots - a_1 f(x) - a_0 x = 0 \quad \text{par définition des coefficients } a_k, \ldots, a_0.$$

□

Le théorème de Cayley-Hamilton a la conséquence pratique suivante :

Corollaire 1.4 *Soit $f \in \mathscr{L}(E)$. Si $\chi_f = (-1)^n (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_k)^{\alpha_k}$, alors $\mu_f = (X - \lambda_1)^{\beta_1} \cdots (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_i \leq \alpha_i$ pour tout $i \in [\![1, k]\!]$.*

Démonstration. Puisque μ_f divise χ_f, μ_f s'écrit bien comme dans l'énoncé avec $\beta_i \leq \alpha_i$. Les β_i sont non nuls car les racines de χ_f sont les valeurs propres de f (proposition 1.3), donc les racines de μ_f (théorème 1.4).

□

d) Lemme des noyaux et critère de diagonalisation

Théorème 1.6 (Lemme des noyaux) *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie , f un endomorphisme de E, $P_1, \ldots, P_k$ k polynômes à coefficients dans $\mathbb{K}$ deux à deux premiers entre eux et $P = P_1 \cdots P_k$.*

1) *On a $\ker P(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_k(f)$.*

2) *On suppose de plus que $E = \ker P(f)$. On note, pour $i \in [\![1, k]\!]$, $Q_i = \frac{P}{P_i}$: d'après le point 1), on a $E = \ker P_i(f) \oplus \ker Q_i(f)$. Si π_i désigne la projection sur $\ker P_i(f)$ dans la direction $\ker Q_i(f)$, alors π_i est un polynôme en f.*

Démonstration. 1) On raisonne par récurrence sur k et on suppose dans un premier temps que $k = 2$. Puisque P_1 et P_2 sont premiers entre eux, il existe deux polynômes U et V tels que $P_1 U + P_2 V = 1$ (identité de Bézout). D'après la proposition 1.7, cette égalité implique $U(f) \circ P_1(f) + V(f) \circ P_2(f) = \mathrm{Id}_E$.

• De $P(f) = P_1(f) \circ P_2(f) = P_2(f) \circ P_1(f)$ (corollaire 1.2), on déduit que $\ker P_i(f) \subset \ker P(f)$ pour $i \in \{1, 2\}$; ainsi, $\ker P_1(f) + \ker P_2(f) \subset \ker P(f)$.

• Pour $x \in E$, on a $x = \mathrm{Id}_E(x) = \underbrace{U(f) \circ P_1(f)(x)}_{y} + \underbrace{V(f) \circ P_2(f)(x)}_{z}$. Mais (corollaire 1.2 et proposition 1.7)

$$P_2(f)(y) = P_2(f) \circ U(f) \circ P_1(f)(x) = U(f) \circ P_2(f) \circ P_1(f)(x) = U(f) \circ P(f)(x)$$

donc si $x \in \ker P(f)$, alors $y \in \ker P_2(f)$, et de même, $z \in \ker P_1(f)$. Ceci montre que

$$\ker P(f) \subset \ker P_1(f) + \ker P_2(f).$$

• Enfin, considérons un élément x de $\ker P_1(f) \cap \ker P_2(f)$. Alors

$$x = \mathrm{Id}_E(x) = U(f) \circ \underbrace{P_1(f)(x)}_{=0} + V(f) \circ \underbrace{P_2(f)(x)}_{=0} = 0$$

donc $\ker P_1(f)$ et $\ker P_2(f)$ sont en somme directe.

• Supposons à présent le résultat vrai pour k polynômes et considérons $k+1$ polynômes $P_1, \ldots, P_{k+1}$ deux à deux premiers entre eux. Notons Q le produit $Q = P_1 \cdots P_k$. Alors Q est premier avec P_{k+1} et $P = Q P_{k+1}$. D'après ce qui précède (cas $k = 2$), on a $\ker P(f) = \ker Q(f) \oplus \ker P_{k+1}(f)$ et par hypothèse de récurrence, on a $\ker Q(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_k(f)$, ce qui montre le résultat au rang $k+1$.

2) Comme $P = P_i Q_i$ avec P_i et Q_i premiers entre eux, il existe $(U, V) \in \mathbb{K}[X]^2$ tel que $UP_i + VQ_i = 1$. On a vu ci-dessus (cas $k = 2$) que tout $x \in \ker P(f)$ s'écrit sous la forme $x = y + z$ avec $y = V(f) \circ Q_i(f)(x) \in \ker P_i(f)$ et $z = U(f) \circ P_i(f)(x) \in \ker Q_i(f)$. Sous l'hypothèse $E = \ker P(f)$, on obtient ainsi la décomposition de tout élément de E selon la somme directe $E = \ker P_i(f) \oplus \ker Q_i(f)$ et on a donc $\pi_i(x) = y$. Par conséquent, on obtient $\pi_i = V(f) \circ Q_i(f) = (VQ_i)(f)$: c'est un polynôme en f.

□

Théorème 1.7 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et f un endomorphisme de E. Les trois assertions suivantes sont équivalentes :*

(i) *f est diagonalisable ;*

(ii) *le polynôme minimal μ_f de f est scindé à racines simples sur $\mathbb{K}$;*

(iii) *f admet un polynôme annulateur scindé à racines simples sur $\mathbb{K}$.*

Démonstration. Puisque μ_f divise tout polynôme annulateur de f et qu'il annule lui-même f, les assertions *(ii)* et *(iii)* sont équivalentes.

Notons $\lambda_1, \ldots, \lambda_k$ les valeurs propres deux à deux distinctes de f.

• Supposons que f soit diagonalisable. Alors (théorème 1.2 et corollaire 1.4) χ_f est scindé sur $\mathbb{K}$ donc μ_f est de la forme $(X-\lambda_1)^{\beta_1}\cdots(X-\lambda_k)^{\beta_k}$ avec $\beta_i \geq 1$. On va montrer que $P=(X-\lambda_1)\cdots(X-\lambda_k)$ annule f (alors, μ_f divise P donc $\beta_1=\cdots=\beta_k=1$).

Puisque f est supposé diagonalisable, E est somme directe des sous-espaces propres de f : $E=\overset{k}{\underset{i=1}{\oplus}} E_{\lambda_i}$. Fixons x dans E et écrivons x sous la forme $x=x_1+\cdots+x_k$ avec $x_i \in E_{\lambda_i}$. Alors :

$$P(f)(x)=\sum_{i=1}^{k} P(f)(x_i)=\sum_{i=1}^{k}(f-\lambda_1 \mathrm{Id}_E)\circ\cdots\circ(f-\lambda_k \mathrm{Id}_E)(x_i) \ ;$$

mais, pour $i \in [\![1,k]\!]$, on a

$$\begin{aligned}(f-\lambda_1\mathrm{Id}_E)\circ\cdots\circ(f-\lambda_k\mathrm{Id}_E)(x_i) &= (f-\lambda_1\mathrm{Id}_E)\circ\cdots\circ(f-\lambda_{i-1}\mathrm{Id}_E)(f-\lambda_{i+1}\mathrm{Id}_E)\circ\cdots \\ &\qquad\qquad\cdots\circ(f-\lambda_k\mathrm{Id}_E)(f-\lambda_i\mathrm{Id}_E)(x_i) \\ &= 0 \quad \text{car } (f-\lambda_i\mathrm{Id}_E)(x_i)=0\end{aligned}$$

donc $P(f)(x)=0$, ceci pour tout $x \in E$.

• Réciproquement, si $\mu_f=(X-\lambda_1)\cdots(X-\lambda_k)$, alors, d'après le lemme des noyaux (théorème 1.6), on a

$$E=\ker \mu_f(f)=\bigoplus_{i=1}^{k}\ker(f-\lambda_i \mathrm{Id}_E)$$

donc f est diagonalisable (proposition 1.1).

□

III _ Sous-espaces stables - Sous-espaces caractéristiques

a) Endomorphisme induit sur un sous-espace stable

Définition *Soit f un endomorphisme d'un espace vectoriel E. Un sous-espace vectoriel F de E est dit* stable *par f si $f(F)\subset F$, c'est-à-dire $f(x)\in F$ pour tout $x \in F$.*

Dans ce cas, f induit un endomorphisme $f_F : F \to F$ défini par $f_F(x)=f(x)$ pour $x\in F$.

Proposition 1.11 *Soient f et g deux endomorphismes de E tels que $f\circ g=g\circ f$. Alors, pour tout polynôme $P\in\mathbb{K}[X]$, $\mathrm{Im}\,P(g)$ et $\ker P(g)$ sont stables par f.*

Nous soulignons le cas où $g=f$:

Corollaire 1.5 *Pour tout endomorphisme f de E et tout polynôme P de $\mathbb{K}[X]$, $\mathrm{Im}\,P(f)$ et $\ker P(f)$ sont stables par f.*

Démonstration. • Soit y un élément de $\operatorname{Im} P(g)$: il existe $x \in E$ tel que $y = P(g)(x)$. En utilisant la proposition 1.8, on obtient

$$f(y) = f\Big(P(g)(x)\Big) = \big(f \circ P(g)\big)(x) = \big(P(g) \circ f\big)(x) = P(g)\Big(f(x)\Big)$$

donc $f(y) \in \operatorname{Im} P(g)$.

• Si $x \in \ker P(g)$, on a, toujours en utilisant la proposition 1.8 :

$$P(g)\Big(f(x)\Big) = \big(P(g) \circ f\big)(x) = \big(f \circ P(g)\big)(x) = f\Big(P(g)(x)\Big) = f(0) = 0$$

donc $f(x) \in \ker P(g)$.

□

Proposition 1.12 *Soit f un endomorphisme de E et F un sous-espace vectoriel de E stable par f. Si f_F désigne l'endomorphisme de F induit par f, alors :*

1) χ_{f_F} *divise* χ_f **2)** μ_{f_F} *divise* μ_f.

Démonstration. (1) Soit b_F une base de F que l'on complète en une base b de E. Si A est la matrice de f_F dans la base b_F, alors la matrice $[f]_b$ de f dans la base b est de la forme $\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$. De là, on déduit $\chi_f = \chi_A \chi_C = \chi_{f_F} \chi_C$.

(2) Si $\mu_f = \sum a_i X^i$, alors, pour tout $x \in F$, on a

$$\begin{aligned}
\big(\mu_f(f_F)\big)(x) &= \Big(\sum a_i f_F^i\Big)(x) = \sum a_i (f_F^i)(x) \\
&= \sum a_i f^i(x) \quad \text{car } f_F(x) = f(x) \text{ puisque } x \in F \\
&= \Big(\sum a_i f^i\Big)(x) = \big(\mu_f(f)\big)(x) = 0.
\end{aligned}$$

Ainsi, $\mu_f(f_F) = 0$ donc μ_{f_F} divise μ_f.

□

Corollaire 1.6 *Soit f un endomorphisme de E et F un sous-espace vectoriel de E stable par f.*

1) *Si f est trigonalisable, alors f_F l'est également.*

2) *Si f est diagonalisable, alors f_F l'est également.*

Démonstration. La première propriété est une conséquence directe des théorème 1.1 et proposition1.12, la seconde, de la même proposition et du théorème 1.7.

□

Proposition 1.13 *Soient f un endomorphisme de E et F, G, deux sous-espaces vectoriels de E stables par f tels que $E = F \oplus G$. Alors :*

1) $\chi_f = \chi_{f_F} \chi_{f_G}$ **2)** $\mu_f = \operatorname{ppcm}(\mu_{f_F}, \mu_{f_G})$.

Démonstration. 1) Considérons une base b_F de F et une base b_G de G. Alors $b = b_F \cup b_G$ est une base de E et la matrice de f dans cette base est $[f]_b = \left(\begin{array}{c|c} [f_F]_{b_F} & 0 \\ \hline 0 & [f_G]_{b_G} \end{array}\right)$. De là, on déduit $\chi_f = \chi_{[f_F]_{b_F}} \chi_{[f_G]_{b_G}} = \chi_{f_F} \chi_{f_G}$.

2) On a vu (proposition 1.12) que μ_f est un multiple de μ_{f_F} et de μ_{f_G}, donc un multiple de leur ppcm. Notons P ce ppcm et montrons que $P(f) = 0$ (alors μ divise P donc $\mu = P$ car ces deux polynômes sont unitaires). Soit $x \in E$, que l'on décompose sous la forme $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$. Alors

$$\begin{aligned}
P(f)(x) &= P(f)(x_F) + P(f)(x_G) \\
&= P(f_F)(x_F) + P(f_G)(x_G) \quad \text{car } x_F \in F,\ x_G \in G \implies f(x_F) = f_F(x_F),\ f(x_G) = f_G(x_G) \\
&= 0 + 0 = 0 \quad \text{car } \mu_F \text{ et } \mu_G \text{ divisent } P.
\end{aligned}$$

□

b) Sous-espaces caractéristiques

Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie. Nous avons vu que si f est diagonalisable, alors E est somme directe des sous-espaces propres. Nous allons montrer un résultat plus général lorsque f est trigonalisable. On suppose donc que le polynôme caractéristique de f est scindé sur $\mathbb{K}$:

$$\chi_f = (-1)^n (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_k)^{\alpha_k} \quad \big(\alpha_i = m_f(\lambda_i)\big).$$

Définition *Le sous-espace* $\ker(f-\lambda \mathrm{Id}_E)^{m_f(\lambda)}$*, noté* F_λ*, est appelé* sous-espace caractéristique *de* f *associé à la valeur propre* λ*.*

Proposition 1.14 *Pour toute valeur propre* λ *de* f*, on a*

(i) $E_\lambda \subset F_\lambda$ *;*

(ii) F_λ *est stable par* f*.*

Démonstration. *(i)* Pour tout endomorphisme g et tout entier $k \geq 1$, on a $\ker g \subset \ker g^k$. Puisque $E_\lambda = \ker(f-\lambda \mathrm{Id}_E)$ et $F_\lambda = \ker(f-\lambda \mathrm{Id}_E)^{m_f(\lambda)}$ avec $m_f(\lambda) \geq 1$, nous avons la première propriété.

(ii) C'est une conséquence du corollaire 1.5 (avec $P = (X-\lambda)^{m_f(\lambda)}$).

□

Théorème 1.8 *Soit* E *un* $\mathbb{K}$*-espace vectoriel de dimension finie et* f *un endomorphisme de* E*. On suppose que le polynôme caractéristique de* f *est scindé sur* $\mathbb{K}$ *et on note* $\lambda_1, \ldots, \lambda_k$ *les valeurs propres deux à deux distinctes de* f*. Alors :*

1) E *est somme directe des sous-espaces caractéristiques de* f *:*

$$E = \ker(f-\lambda_1 \mathrm{Id}_E)^{m_f(\lambda_1)} \oplus \cdots \oplus \ker(f-\lambda_k \mathrm{Id}_E)^{m_f(\lambda_k)}.$$

2) *Si* $\mu_f = \prod_{i=1}^{k} (X-\lambda_i)^{\beta_i}$*, alors, pour tout* $i \in [\![1,k]\!]$*, on a*

$$\ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)} = \ker(f-\lambda_i \mathrm{Id}_E)^{\beta_i}.$$

3) *Chaque sous-espace caractéristique* $\ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$ *est de dimension* $m_f(\lambda_i)$*.*

Démonstration. **1)** Cette première partie est un résumé des résultats précédents. En effet, on a $\chi_f(f) = 0$ (théorème de Cayley-Hamilton) et $\chi_f(X) = (-1)^n \prod_{i=1}^{k} (X-\lambda_i)^{m_f(\lambda_i)}$ avec, si i est différent de j, $(X-\lambda_i)^{m_f(\lambda_i)}$ premier avec $(X-\lambda_j)^{m_f(\lambda_j)}$. Le lemme des noyaux permet donc d'écrire

$$E = \ker \chi_f(f) = \bigoplus_{i=1}^{k} \ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}.$$

2) Notons F_{λ_i} le sous-espace caractéristique $\ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$, d_i sa dimension et δ_i la dimension de $\ker(f-\lambda_i \mathrm{Id}_E)^{\beta_i}$. En utilisant le lemme des noyaux avec μ_f, on obtient comme précédemment

$$E = \ker \mu_f(f) = \bigoplus_{i=1}^{k} \ker(f-\lambda_i \mathrm{Id}_E)^{\beta_i} \implies n = \sum_{i=1}^{k} \delta_i.$$

Puisque $\beta_i \leq m_f(\lambda_i)$ (voir corollaire 1.4), on a $\ker(f-\lambda_i \mathrm{Id}_E)^{\beta_i} \subset \ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$ donc $\delta_i \leq d_i$. Mais d'après le point **1)**, on a également $n = \sum_{i=1}^{k} d_i$ donc $\delta_i = d_i$ et $\ker(f-\lambda_i \mathrm{Id}_E)^{\beta_i} = \ker(f-\lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$.

3) Chaque sous-espace caractéristique F_{λ_i} est stable par f (proposition 1.14) : notons f_i l'endomorphisme de F_{λ_i} induit par f. On vient de voir que $F_{\lambda_i} = \ker(f - \lambda_i \mathrm{Id}_E)^{\beta_i}$, donc le polynôme $(X - \lambda_i)^{\beta_i}$ annule f_i, ce qui permet d'affirmer que μ_{f_i} divise $(X - \lambda_i)^{\beta_i}$. Puisque μ_{f_i} et χ_{f_i} ont les mêmes racines, λ_i est la seule valeur propre de f_i donc $\chi_{f_i} = (\lambda_i - X)^{d_i}$. De là, on obtient, en utilisant la proposition 1.13

$$\prod_{i=1}^{k} (\lambda_i - X)^{m_f(\lambda_i)} = \chi_f(X) = \prod_{i=1}^{k} \chi_{f_i}(X) = \prod_{i=1}^{k} (\lambda_i - X)^{d_i}$$

ce qui montre que $d_i = m_f(\lambda_i)$.

$\square$

c) Décomposition de Dunford[3]

Définition *Un endomorphisme f de E est* nilpotent *s'il existe $s \in \mathbb{N}^*$ tel que $f^s = 0$. Le plus petit $s > 0$ vérifiant $f^s = 0$ est appelé* l'indice de nilpotence *de f.*
De même, une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est dite nilpotente s'il existe $s \in \mathbb{N}^*$ tel que $A^s = 0$ et on définit l'indice de nilpotence de A de la même manière.

Remarque : si f est nilpotent, alors toute matrice A représentant f est nilpotente de même indice (car $[f]_b^k = [f^k]_b$ pour tout $k \in \mathbb{N}$).

Proposition 1.15 *Un endomorphisme f de E est nilpotent si, seulement si, $\chi_f = (-X)^n$. En particulier, la seule valeur propre d'un endomorphisme nilpotent est 0.*

Démonstration. $\Longrightarrow$ Soit $\lambda \in \mathbb{C}$ une racine de χ_f. Il existe x non nul dans E tel que $f(x) = \lambda x$, ce qui implique $f^k(x) = \lambda^k x$ pour tout $k \in \mathbb{N}$. Alors, si $s \in \mathbb{N}^*$ vérifie $f^s = 0$, on a $\lambda^s = 0$ dans $\mathbb{C}$, ce qui implique $\lambda = 0$. Ainsi, la seule racine complexe de χ_f est nulle, donc $\chi_f = (-X)^n$ (car χ_f est scindé sur $\mathbb{C}$).

$\Longleftarrow$ Si $\chi_f = (-X)^n$, alors $f^n = 0$ (théorème de Cayley-Hamilton) donc f est nilpotent.

$\square$

Théorème 1.9 (Décompositon de Dunford) *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et f un endomorphisme de E. On suppose que le polynôme caractéristique de f est scindé sur $\mathbb{K}$.*
Alors il existe un unique couple (D, N) d'endomorphismes de E vérifiant :

(i) $f = D + N$;

(ii) D *est diagonalisable ;*

(iii) N *est nilpotent ;*

(iv) $D \circ N = N \circ D$.

En outre, D et N sont des polynômes en f.

Démonstration. Notons $\lambda_1, \ldots, \lambda_k$ les valeurs propres deux à deux distinctes de f, F_{λ_i} le sous-espace caractéristique $\ker(f - \lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$ et π_i la projection sur F_{λ_i} dans la direction $\underset{j \neq i}{\oplus} F_{\lambda_j}$: c'est un polynôme en f (voir lemme des noyaux, théorème 1.6).

• Supposons donné un couple (D, N) satisfaisant aux quatre propriétés du théorème.
Soient μ une valeur propre de D (il en admet au moins une car il est diagonalisable), $F = \ker(D - \mu \mathrm{Id}_E)$ le sous-espace propre associé et t la dimension de F. Puisque $D \circ N = N \circ D$, F est stable par N (proposition 1.11). Puisqu'il est également stable par D, il l'est par $f = D + N$. Notons comme d'habitude N_F, D_F et f_F les endomorphismes de F induits respectivement par N, D et f. Le polynôme caractéristique de N_F divise celui de N (proposition 1.12) donc est de la forme $\chi_{N_F} = (-X)^t$ (proposition 1.15). En particulier, N_F est trigonalisable (théorème 1.1) et il existe donc une base b_F de F telle que $[N_F]_{b_F}$ soit de la forme

$$[N_F]_{b_F} = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

[3]Nelson Dunford, mathématicien américain, 1906-1986.

Puisque $D_F = \mu \mathrm{Id}_F$, on obtient $[f]_{b_F} = [D_F]_{b_F} + [N_F]_{b_F} = \begin{pmatrix} \mu & & * \\ & \ddots & \\ 0 & & \mu \end{pmatrix}$. Par conséquent, $\chi_{f_F} = (\mu - X)^t$ et donc μ est valeur propre de f. De plus, puisque χ_{f_F} divise χ_f et $t = \deg \chi_{f_F}$, $t \leq m_f(\mu)$, ce qui implique $\ker(f - \mu \mathrm{Id}_E)^t \subset \ker(f - \mu \mathrm{Id}_E)^{m_f(\mu)}$. D'autre part, on a $F \subset \ker(f - \mu \mathrm{Id}_E)^t$. En effet, si $x \in F$, alors

$$(f - \mu \mathrm{Id}_E)^t(x) = (-1)^t \chi_{f_F}(f)(x) = (-1)^t \chi_{f_F}(f_F)(x) = 0.$$

Par conséquent, on obtient

$$F = \ker(D - \mu \mathrm{Id}_E) \subset \ker(f - \mu \mathrm{Id}_E)^t \subset \ker(f - \mu \mathrm{Id}_E)^{m_f(\mu)} = F_\mu.$$

Si l'on considère les inclusions ainsi obtenues pour l'ensemble des valeurs propres de D, on obtient

$$E = \bigoplus_{\mu \in \mathrm{Sp}(D)} \ker(D - \mu \mathrm{Id}_E) \subset \bigoplus_{\mu \in \mathrm{Sp}(D)} F_\mu \subset \bigoplus_{\mu \in \mathrm{Sp}(f)} F_\mu = E \quad \text{d'après le théorème 1.8,}$$

ce qui implique $\mathrm{Sp}(D) = \mathrm{Sp}(f)$ et $\ker(D - \mu \mathrm{Id}_E) = F_\mu$ pour tout μ. Par conséquent, pour $x \in E$ s'écrivant sous la forme $x = x_1 + \cdots + x_k$ avec $x_i \in F_{\lambda_i}$ (écriture selon la décomposition $E = \overset{k}{\underset{i=1}{\oplus}} F_{\lambda_i}$), on a $D(x) = \lambda_1 x_1 + \cdots + \lambda_k x_k$, ou encore $D = \lambda_1 \pi_1 + \cdots + \lambda_k \pi_k$ puisque $\pi_i(x) = x_i$.
Ainsi, D est déterminé de façon unique et par conséquent, $N = f - D$ également : ceci montre l'unicité.

• Montrons à présent l'existence de D et N. Posons $D = \lambda_1 \pi_1 + \cdots + \lambda_k \pi_k$ et $N = f - D$. On a bien sûr $f = D + N$; de plus D est diagonalisable. En effet, pour chaque $i \in [\![1, k]\!]$ et tout $x \in F_{\lambda_i}$, on a

$$D(x) = \lambda_1 \pi_1(x) + \cdots + \lambda_k \pi_k(x) = \lambda_i \pi_i(x) = \lambda_i x$$

donc $F_{\lambda_i} \subset \ker(D - \lambda_i \mathrm{Id}_E)$. Puisque $E = \overset{k}{\underset{i=1}{\oplus}} F_{\lambda_i}$, on a $F_{\lambda_i} = \ker(D - \lambda_i \mathrm{Id}_E)$ et D diagonalisable car E est somme directe de ses sous-espaces propres. Il reste donc à démontrer que N est nilpotent et $D \circ N = N \circ D$.

⋆ Soit $x \in E$ que l'on décompose sous la forme $x = x_1 + \cdots + x_k$ avec $x_i \in F_{\lambda_i}$. On a

$$\begin{aligned} f \circ D(x) &= \sum_{i=1}^k f \circ D(x_i) = \sum_{i=1}^k f(\lambda_i x_i) = \sum_{i=1}^k \lambda_i f(x_i) \\ D \circ f(x) &= \sum_{i=1}^k D \circ f(x_i) = \sum_{i=1}^k \lambda_i f(x_i), \end{aligned}$$

la dernière égalité provenant du fait que, F_{λ_i} étant stable par f, $f(x_i)$ est un élément de F_{λ_i} ($x_i \in F_{\lambda_i}$). Ainsi, D commute avec f, donc avec $N = f - D$.

⋆ Notons $s = \max\{m_f(\lambda_1), \ldots, m_f(\lambda_k)\}$ et montrons que $N^s = 0$. Fixons x dans E, que l'on décompose sous la forme $x = x_1 \cdots + x_k$ avec $x_i \in F_{\lambda_i}$. On a

$$N^s(x) = \sum_{i=1}^k N^s(x_i) = \sum_{i=1}^k (f - D)^s(x_i).$$

Mais, pour $i \in [\![1, k]\!]$ et $y \in F_{\lambda_i}$, on a

$$(f - D)(y) = f(y) - D(y) = f(y) - \lambda_i y = (f - \lambda_i \mathrm{Id}_E)(y) \in F_{\lambda_i}$$

car F_{λ_i} est stable par f. Par conséquent, (récurrence sur $l \in \mathbb{N}$) on a $(f - D)^l(x_i) = (f - \lambda_i \mathrm{Id}_E)^l(x_i)$. Ainsi, $(f - D)^s(x_i) = (f - \lambda_i \mathrm{Id}_E)^s(x_i) = 0$ puisque $s \geq m_f(\lambda_i)$ et $x_i \in F_{\lambda_i} = \ker(f - \lambda_i \mathrm{Id}_E)^{m_f(\lambda_i)}$, ceci pour tout $i \in [\![1, k]\!]$. On a donc bien démontré que $N^s(x) = 0$ pour tout $x \in E$.

□

2

Dualité

Définition *Une* forme linéaire *sur E est une application linéaire de E dans $\mathbb{K}$.*
Le dual *de E, noté E^*, est l'ensemble des formes linéaires sur E : $E^* = \mathscr{L}(E, \mathbb{K})$. Le dual de E^* est appelé* bidual *de E et noté E^{**}.*

Rappelons que ce sont des espaces vectoriels sur $\mathbb{K}$.

I _ Base duale

Rappelons le théorème d'algèbre linéaire suivant :

Théorème 2.1 *Soient E et F deux $\mathbb{K}$-espaces vectoriels, $b = (e_1, \ldots, e_n)$ une base de E et $(y_1, \ldots, y_n)$ une famille quelconque de n éléments de F.*

Il existe une unique application linéaire $f : E \to F$ vérifiant $f(e_i) = y_i$.

Fixons une base $(e_1, \ldots, e_n)$ de E. Pour $i \in [\![1, n]\!]$, on utilise ce théorème pour définir la forme linéaire e_i^* par

$$e_i^* : E \to \mathbb{K}, \quad e_i^*(e_k) = \begin{cases} 1 & \text{si } k = i \text{ ;} \\ 0 & \text{sinon.} \end{cases}$$

Nous disposons ainsi de n formes linéaires sur E ; elle vérifient

Théorème 2.2 *$b^* = (e_1^*, \ldots, e_n^*)$ est une base de E^*.*

Définition *La base b^* de E^* ainsi associée à la base b de E est appelée* base duale *de b.*

Démonstration. Nous allons montrer que cette famille est libre et génératrice.

• Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}$ tel que $\sum_{i=1}^{n} \lambda_i e_i^* = 0$. Alors, pour tout $k \in [\![1, n]\!]$, on a

$$0 = \left(\sum_{i=1}^{n} \lambda_i e_i^* \right) (e_k) = \sum_{i=1}^{n} \lambda_i e_i^*(e_k) = \lambda_k$$

donc la famille b^* est libre.

• Soit $l \in E^*$ une forme linéaire sur E ; montrons que $l = \sum_{i=1}^{n} \lambda_i e_i^*$ avec $\lambda_i = l(e_i) \in \mathbb{K}$. Pour $k \in [\![1, n]\!]$, on a

$$\left(\sum_{i=1}^{n} l(e_i) e_i^* \right) (e_k) = \sum_{i=1}^{n} l(e_i) e_i^*(e_k) = l(e_k)$$

donc les deux formes linéaires $\sum_{i=1}^{n} l(e_i) e_i^*$ et l coïncident sur les éléments de la base $(e_1, ..., e_n)$. Elles sont donc égales d'après le théorème 2.1 rappelé précédemment.

□

Corollaire 2.1 *L'espace dual E^* est de dimension finie $n = \dim E$, donc isomorphe à E.*

Un tel isomorphisme est donné par le choix d'une base $(e_1, \ldots, e_n)$ de E : c'est l'isomorphisme envoyant e_i sur e_i^*. Cet isomorphisme dépend du choix de la base : il n'est pas canonique. Par contre, nous verrons (proposition 2.1) que le bidual E^{**} de E (*ie* le dual du dual) est canoniquement isomorphe à E.

Les n formes linéaires e_k^* sont appelées *applications coordonnées*. Ceci est justifié par le fait que si $x = \sum_{i=1}^{n} x_i e_i$ est un élément de E écrit dans la base $(e_1, \ldots, e_n)$, alors, pour $k \in [\![1, n]\!]$, on a

$$e_k^*(x) = \sum_{i=1}^{n} x_i e_k^*(e_i) = x_k.$$

Ainsi, $e_k^*(x)$ est la $k^{\text{ème}}$ coordonnée de x dans la base $(e_1, \ldots, e_n)$. On retiendra donc (en plus du reste !) que, si $(e_1, \ldots, e_n)$ est une base de E, alors pour tout $x \in E$,

$$\boxed{x = \sum_{i=1}^{n} e_i^*(x) e_i.} \tag{2.1}$$

De même, si $l \in E^*$ est une forme linéaire, on a (voir démonstation du théorème 2.2)

$$\boxed{l = \sum_{i=1}^{n} l(e_i) e_i^*.} \tag{2.2}$$

Proposition 2.1 *Soit* $\mathrm{Ev} : E \to E^{**}$ *l'application qui, à x, associe* Ev_x*, la forme linéaire sur E^* définie par*

$$\mathrm{Ev}_x : \left(\begin{array}{ccc} E^* & \to & \mathbb{K} \\ l & \mapsto & l(x) \end{array} \right).$$

Alors Ev *est un isomorphisme de $\mathbb{K}$-espaces vectoriels.*

Démonstration. • Pour $x \in E$ fixé, l'application Ev_x est bien linéaire :

$$\mathrm{Ev}_x(l_1 + \lambda l_2) = (l_1 + \lambda l_2)(x) = l_1(x) + \lambda l_2(x) = \mathrm{Ev}_x(l_1) + \lambda \mathrm{Ev}_x(l_2).$$

Ceci signifie que Ev est bien à valeur dans E^{**}.

• Pour $(x, y) \in E^2$, $\lambda \in \mathbb{K}$ et $l \in E^*$, on a

$$\mathrm{Ev}_{x+\lambda y}(l) = l(x + \lambda y) = l(x) + \lambda l(y) = \mathrm{Ev}_x(l) + \lambda \mathrm{Ev}_y(l) = (\mathrm{Ev}_x + \lambda \mathrm{Ev}_y)(l)$$

donc $\mathrm{Ev}_{x+\lambda y} = \mathrm{Ev}_x + \lambda \mathrm{Ev}_y$: ceci montre que Ev est linéaire.

• Soit $x \in \ker \mathrm{Ev}$. Pour toute forme linéaire l sur E, on a $0 = \mathrm{Ev}_x(l) = l(x)$. Or si $x \neq 0$, il existe $l \in E^*$ tel que $l(x) \neq 0$ (considérer une base $b = (e_1, \ldots, e_n)$ de E telle que $e_1 = x$ et prendre par exemple $l = e_1^*$). Par conséquent, $\ker \mathrm{Ev} = \{0\}$ et Ev est injective.

• Puisque $\dim E^{**} = \dim E^* = \dim E$ (corollaire 2.1), on en déduit que Ev est un isomorphisme.

□

Exercice. Décrire l'isomorphisme réciproque $\mathrm{Ev}^{-1} : E^{**} \to E$.

Proposition 2.2 *Notons $\mathcal{B}$ l'ensemble des bases de E et $\mathcal{B}^*$ l'ensemble des bases de E^*. Alors, l'application de $\mathcal{B}$ vers $\mathcal{B}^*$, qui, à une base, associe sa base duale, est une bijection.*

Démonstration. • Soient $b = (e_1, \dots, e_n)$ et $b' = (f_1, \dots, f_n)$ deux bases de E telles que $e_i^* = f_i^*$ pour tout $i \in [\![1, n]\!]$. Pour chaque $k \in [\![1, n]\!]$, on a

$$\begin{aligned} f_k &= \sum_{i=1}^n e_i^*(f_k)e_i \quad \text{(coordonnées de } f_k \text{ dans la base } b) \\ &= \sum_{i=1}^n f_i^*(f_k)e_i = e_k, \end{aligned}$$

ce qui démontre l'injectivité de l'application.

• Soit $d = (l_1, \dots, l_n)$ une base de E^*. Considérons sa base duale $(l_1^*, \dots, l_n^*)$: c'est une base du bidual E^{**} (théorème 2.2). Puisque $\mathrm{Ev} : E \to E^{**}$ est un isomorphisme (proposition 2.1), il existe un unique élément $u_i \in E$ tel que $\mathrm{Ev}(u_i) = l_i^*$, ceci pour tout $i \in [\![1, n]\!]$. La famille $b = (u_1, \dots, u_n)$ est une base de E car c'est l'image de la base d^* par l'isomorphisme Ev^{-1}. Enfin, pour tout $(i, j) \in [\![1, n]\!]^2$, on a

$$l_j(u_i) = \mathrm{Ev}_{u_i}(l_j) = l_i^*(l_j) = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{sinon} \end{cases}$$

donc $l_j = u_j^*$: ceci montre que d est la base duale de b et donc la surjectivité de l'application.

□

II _ Application transposée

Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et u une application linéaire de E dans F. On définit une application ${}^t u : F^* \to E^*$ par ${}^t u(l) = l \circ u$ si $l \in F^*$:

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[rr, bend right, "{}^t u(l)"'] & F \arrow[r, "l"] & \mathbb{K}. \end{tikzcd}$$

Lemme 1 *${}^t u$ est linéaire.*

Démonstration. Pour $l_1, l_2 \in F^*$ et $\lambda \in \mathbb{K}$, on a

$${}^t u(l_1 + \lambda l_2) = (l_1 + \lambda l_2) \circ u = l_1 \circ u + \lambda(l_2 \circ u) = {}^t u(l_1) + \lambda\, {}^t u(l_2).$$

□

Proposition 2.3 *Soient $e = (e_1, \dots, e_n)$ et $f = (f_1, \dots, f_p)$ des bases de E et F respectivement.*
Si $u \in \mathscr{L}(E, F)$ et $A \in \mathscr{M}_{p,n}(\mathbb{K})$ est sa matrice dans ces bases, alors la matrice de ${}^t u \in \mathscr{L}(F^*, E^*)$ dans les bases duales f^* et e^* est ${}^t A$.

Démonstration. Fixons $(i, j) \in [\![1, n]\!]^2$.

• Le coefficient (i, j) de A est la $i^{\text{ème}}$ coordonnée de $u(e_j)$ dans la base f ; il est donc égal à (voir égalité 2.1)

$$f_i^*\big(u(e_j)\big) = (f_i^* \circ u)(e_j) = \Big({}^t u(f_i^*)\Big)(e_j).$$

• Le coefficient (i, j) de la matrice de ${}^t u$ dans les bases f^* et e^* est la $i^{\text{ème}}$ coordonnée de ${}^t u(f_j^*)$ dans la base e^* ; il est donc égal à (égalité 2.2) $\big({}^t u(f_j^*)\big)(e_i)$, et c'est bien le coefficient (j, i) de A.

□

III _ Orthogonalité

Définition *Soit F un sous-espace vectoriel de E. On appelle orthogonal de F, et on note F°, l'ensemble des formes linéaires sur E qui s'annulent sur F :*

$$F^\circ = \{l \in E^* \,/\, l(x) = 0 \ \forall x \in F\} = \{l \in E^* \,/\, F \subset \ker l\}\,.$$

Lemme 2 *F° est un sous-espace vectoriel de E^*.*

Démonstration. Si l_1 et l_2 sont deux éléments de F° et λ un élément de $\mathbb{K}$, on a, pour tout $x \in F$:

$$(l_1 + \lambda l_2)(x) = l_1(x) + \lambda l_2(x) = 0$$

donc $l_1 + \lambda l_2 \in F^\circ$.

$\square$

Proposition 2.4 *Pour tout sous-espace vectoriel F de E, on a $\dim F + \dim F^\circ = \dim E$.*

Démonstration. Soit $(e_1, \ldots, e_p)$ une base de F que l'on complète en une base $b = (e_1, \ldots, e_n)$ de E. On va montrer que $(e^*_{p+1}, \ldots, e^*_n)$ est une base de F° : alors $\dim F = p$ et $\dim F^\circ = n-p$, ce qui montre la proposition.

• Remarquons que $e^*_{p+1}, \ldots, e^*_n$ sont bien des éléments de F° : si $i \in [\![1,p]\!]$ et $k \in [\![p+1,n]\!]$, alors $e^*_k(e_i) = 0$.

• Soit $l \in F^\circ$; l s'écrit dans la base duale b^* sous la forme $l = \sum_{i=1}^{n} l(e_i)e^*_i$ (égalité 2.2). Mais pour $i \in [\![1,p]\!]$, on a $l(e_i) = 0$ (car $e_i \in F$) donc finalement, $l = \sum_{i=p+1}^{n} l(e_i)e^*_i$. Ceci montre que $(e^*_{p+1}, \ldots, e^*_n)$ engendre F°.

• Mais cette famille est libre (théorème 2.2) ; c'est donc une base de F°.

$\square$

Définition *Soit G un sous-espace vectoriel de E^*. On appelle orthogonal de G, et on note ${}^\circ G$, l'ensemble des éléments x de E qui annulent toutes les formes linéaires de G :*

$${}^\circ G = \{x \in E \,/\, l(x) = 0 \ \forall l \in G\} = \bigcap_{l \in G} \ker l.$$

Il est immédiat que ${}^\circ G$ est un sous-espace vectoriel de E (intersection de sous-espaces vectoriels).

Proposition 2.5 *Pour tout sous-espace vectoriel G de E^*, on a $\dim G + \dim {}^\circ G = \dim E$.*

Démonstration. On suit le même schéma que la démonstration de la proposition 2.4 : on considère une base $d = (l_1, \cdots, l_n)$ de E^* telle que $(l_1, \ldots, l_q)$ soit une base de G, et on montre que $(e_{q+1}, \ldots, e_n)$ est une base de ${}^\circ G$ où $e = (e_1, \ldots, e_n)$ est la base de E dont d est la base duale.

• e_k est dans ${}^\circ G$ pour tout $k \in [\![q+1,n]\!]$ car $l_i(e_k) = e^*_i(e_k) = 0$ si $1 \leq i \leq q < k \leq n$.

• Si $x \in {}^\circ G$ s'écrit dans la base e sous la forme $x = \sum_{i=1}^{n} x_i e_i$ alors, pour tout $k \in [\![1,q]\!]$, on a

$$x_k = e^*_k(x) = l_k(x) = 0$$

donc $x \in \mathrm{Vect}(e_{q+1}, \ldots, e_n)$. Ceci montre que $(e_{q+1}, \ldots, e_n)$ engendre ${}^\circ G$. Comme cette famille est libre, c'est une base de ${}^\circ G$.

$\square$

Corollaire 2.2 *Soient $l_1, \ldots, l_p$ p formes linéaires sur E linéairement indépendantes. Si $F = \ker l_1 \cap \cdots \cap \ker l_p$, alors F est de dimension $n-p$.*

Remarque : ce résultat signifie géométriquement que, dans un espace vectoriel de dimension finie, l'intersection de p hyperplans dont les équations sont linéairement indépendantes est un sous-espace vectoriel de dimension $n-p$.

Démonstration. On considère le sous-espace G engendré par $\{l_1, \dots, l_p\}$ dans E^* : il est de dimension p par hypothèse. Pour conclure, on montre que $F = {}^\circ G$ et on applique le résultat de la proposition 2.5.

- Par définition, ${}^\circ G = \bigcap_{l \in G} \ker l \subset \bigcap_{i=1}^{p} \ker l_i = F$ car $l_i \in G$ pour tout $i \in [\![1, p]\!]$.

- Soient $x \in F$ et $l \in G$. Il existe $\lambda_1, \dots, \lambda_p$ dans $\mathbb{K}$ tels que $l = \sum_{i=1}^{p} \lambda_i l_i$; on obtient alors

$$l(x) = \sum_{i=1}^{p} \lambda_i l_i(x) = 0 \quad \text{puisque } x \in F = \bigcap_{i=1}^{p} \ker l_i,$$

donc $x \in {}^\circ G$.

□

Nous concluons ce chapitre par le résultat attendu suivant :

Proposition 2.6 *Pour tous sous-espaces F de E et G de E^*, on a ${}^\circ(F^\circ) = F$ et $({}^\circ G)^\circ = G$.*

Démonstration. Les propositions 2.4 et 2.5 montrent que $\dim {}^\circ(F^\circ) = \dim F$ et $\dim({}^\circ G)^\circ = \dim G$. Or on a clairement $F \subset {}^\circ(F^\circ)$ et $G \subset ({}^\circ G)^\circ$.

□

3

Formes quadratiques - Espaces vectoriels euclidiens

Dans tout ce chapitre, E est un $\mathbb{R}$-espace vectoriel de dimension finie n.

I _ Formes bilinéaires - Formes quadratiques

a) Définitions

Définition **1)** *Une* forme bilinéaire *sur E est une application $\varphi : E \times E \to \mathbb{R}$ vérifiant :*

(i) *$\forall x \in E$, l'application $y \mapsto \varphi(x,y)$ est linéaire (linéarité à droite) ;*

(ii) *$\forall y \in E$, l'application $x \mapsto \varphi(x,y)$ est linéaire (linéarité à gauche).*

2) *Une forme bilinéaire φ sur E est dite* symétrique *si $\varphi(y,x) = \varphi(x,y)$ pour tout $(x,y) \in E^2$.*

Exemples : • $\varphi : \mathbb{R}^2 \to \mathbb{R}$, $\varphi(x,y) = xy$ est bilinéaire symétrique.
• $\varphi : \mathbb{R}_3[X] \times \mathbb{R}_3[X] \to \mathbb{R}$, $\varphi(P,Q) = \int_{-1}^{1} P(t)Q(t)\mathrm{d}t$ est bilinéaire symétrique.

Proposition 3.1 *L'ensemble des formes bilinéaires (respectivement bilinéaires symétriques) sur E, noté $\mathscr{L}_2(E)$ (resp. $\mathscr{L}_2^s(E)$) est un espace vectoriel sur $\mathbb{R}$.*

Démonstration. On définit la somme et le produit externe par un scalaire comme suit (φ_1 et φ_2 sont deux formes bilinéaires, α un élément de $\mathbb{R}$) :

$$(\varphi_1 + \varphi_2)(x,y) = \varphi_1(x,y) + \varphi_2(x,y),$$
$$(\alpha\,\varphi_1)(x,y) = \alpha\varphi_1(x,y).$$

Il est aisé de vérifier que ces deux opérations font de $\mathscr{L}_2(E)$ un $\mathbb{R}$-espace vectoriel et de $\mathscr{L}_2^s(E)$ un sous-espace vectoriel de $\mathscr{L}_2(E)$.

□

Définition *Une forme* quadratique *sur E est une application $q : E \to \mathbb{R}$ telle qu'il existe une forme bilinéaire φ sur E vérifiant $q(x) = \varphi(x,x)$ pour tout x dans E.*

Conséquence : toute forme bilinéaire φ sur E donne naissance à une forme quadratique q donnée par $q(x) = \varphi(x,x)$. Maintenant, si $\beta : E \times E \to \mathbb{R}$ est définie par $\beta(x,y) = \frac{1}{2}\big(\varphi(x,y) + \varphi(y,x)\big)$, alors β est une forme bilinéaire symétrique et, pour tout $x \in E$, on a $\beta(x,x) = \varphi(x,x) = q(x)$.
Ceci montre que si q est une forme quadratique sur E, il existe une forme bilinéaire symétrique β sur E telle que $q(x) = \beta(x,x)$. On a en fait :

Proposition 3.2 *Soit $q : E \to \mathbb{R}$ une forme quadratique sur E. Il existe une unique forme bilinéaire symétrique β sur E telle que $q(x) = \beta(x,x)$ pour tout x dans E.*

Démonstration. Nous venons de montrer l'existence. L'unicité de β est une conséquence de la formule de polarisation suivante :

$$\forall (x,y)\in E^2,\quad \beta(x,y)=\frac{1}{2}\Big(q(x+y)-q(x)-q(y)\Big)=\frac{1}{4}\Big(q(x+y)-q(x-y)\Big).$$

En effet, en utilisant la bilinéarité et la symétrie de β, on obtient

$$q(x+y)=\beta(x+y,x+y)=\beta(x,x+y)+\beta(y,x+y)=\beta(x,x)+\beta(x,y)+\beta(y,x)+\beta(y,y)=q(x)+q(y)+2\beta(x,y).$$

□

Définition *On appelle* forme polaire *d'une forme quadratique q sur E l'unique forme bilinéaire symétrique β sur E vérifiant $q(x)=\beta(x,x)$ pour tout x dans E.*

Remarques : 1) si on note $\mathcal{Q}(E)$ l'ensemble des formes quadratiques sur E, la proposition 3.2 nous assure que l'application de $\mathscr{L}_2^s(E)$ dans $\mathcal{Q}(E)$, qui, à une forme bilinéaire symétrique β, associe la forme quadratique q définie par $q(x)=\beta(x,x)$, est un isomorphisme d'espaces vectoriels.

2) Avec ces notations, on a, pour tout $(x,y)\in E^2$ et tout $(\lambda,\mu)\in\mathbb{R}^2$

$$q(\lambda x+\mu y)=\lambda^2 q(x)+\mu^2 q(y)+2\lambda\mu\beta(x,y).$$

b) Écriture matricielle

Fixons une base $b=(e_1,\dots,e_n)$ de E.

Définition *On appelle matrice d'une forme bilinéaire φ sur E dans la base b la matrice $A=(a_{i,j})_{1\le i,j\le n}$ de $\mathscr{M}_n(\mathbb{R})$ définie par $a_{i,j}=\varphi(e_i,e_j)$.*

Proposition 3.3 *Soit $b=(e_1,\dots,e_n)$ une base de E et φ une forme bilinéaire sur E, de matrice A dans la base b. Pour tout $x=\sum_{i=1}^n x_ie_i$ et $y=\sum_{i=1}^n y_ie_i$ dans E, on a*

$$\varphi(x,y)={}^tXAY\quad \text{où}\quad X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\quad\text{et}\quad Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

De plus, si $B\in\mathscr{M}_n(\mathbb{R})$ aussi vérifie cette égalité pour tout $(x,y)\in E^2$, alors $B=A$.

Démonstration. • D'une part, par bilinéarité de φ, on a

$$\varphi(x,y)=\varphi\left(\sum_{i=1}^n x_ie_i,\sum_{j=1}^n y_je_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_i\,y_j\,\varphi(e_i,e_j)=\sum_{1\le i,j\le n}x_i\,y_j\,a_{i,j},$$

et d'autre part,

$$({}^tXA)_j=\sum_{k=1}^n({}^tX)_ka_{k,j}=\sum_{k=1}^n x_ka_{k,j}\quad\Longrightarrow\quad {}^tXAY=\sum_{l=1}^n({}^tXA)_ly_l=\sum_{l=1}^n\sum_{k=1}^n x_ka_{k,l}y_l,$$

ce qui donne bien l'égalité $\varphi(x,y)={}^tXAY$.

• Supposons que $B\in\mathscr{M}_n(\mathbb{R})$ vérifie $\varphi(x,y)={}^tXBY$ pour tout $(x,y)\in E^2$. Alors, pour tout couple $(i,j)\in[\![1,n]\!]^2$, on a

$$\begin{aligned}a_{i,j}=\varphi(e_i,e_j)={}^t(E_i)BE_j\quad\text{où}\quad E_i&=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}\leftarrow i\\ &=\sum_{1\le k,l\le n}({}^tE_i)_kb_{k,l}(E_j)_l=b_{i,j}.\end{aligned}$$

□

Corollaire 3.1 *Une forme bilinéaire est symétrique si, et seulement si, sa matrice dans une base quelconque est symétrique.*

Démonstration. Notons A la matrice dans une base $b = (e_1, \ldots, e_n)$ d'une forme bilinéaire φ. Si celle-ci est symétrique, alors A est clairement symétrique car $a_{i,j} = \varphi(e_i, e_j) = \varphi(e_j, e_i) = a_{j,i}$.
Réciproquement, si A est symétrique, alors, d'après la proposition 3.3, on a, pour tout couple (x, y) d'éléments de E (avec les mêmes notations que pour la proposition)

$$\begin{aligned}\varphi(y, x) &= {}^tYAX = {}^t\big({}^tYAX\big)\\ &= {}^tX\,{}^tAY = {}^tXAY = \varphi(x, y).\end{aligned}$$

□

Proposition 3.4 (Changement de bases) *Soit φ une forme bilinéaire sur E, b et b' deux bases de E et P la matrice de passage de b à b'. Si A est la matrice de φ dans b et A' la matrice de φ dans b', on a*

$$\boxed{A' = {}^tP\,A\,P.}$$

Démonstration. Soient x et y deux éléments de E et X, Y (resp. X', Y') les matrices colonnes de leurs coordonnées dans la base b (resp. b'). On a $X = PX'$ et $Y = PY'$, donc on obtient (proposition 3.3) :

$$\varphi(x, y) = {}^tXAY = {}^t(PX')A(PY') = {}^tX'({}^tPAP)Y',$$

ceci pour tout $(x, y) \in E^2$. Le deuxième point de la proposition 3.3 montre que $A' = {}^tPAP$.

□

c) Dualité

Soit φ une forme bilinéaire sur E. Pour $x, y \in E$, on définit deux formes linéaires sur E par :

$$\begin{aligned}\varphi_x : E \to \mathbb{R}, &\quad \varphi_x(z) = \varphi(x, z)\\ \varphi^y : E \to \mathbb{R}, &\quad \varphi^y(z) = \varphi(z, y).\end{aligned}$$

La linéarité de φ_x et φ^y est une conséquence directe de la bilinéarité de φ ; celle-ci implique également la linéarité des deux applications $\varphi_g : E \to E^*$ et $\varphi_d : E \to E^*$ respectivement définies par $\varphi_g(x) = \varphi_x$ et $\varphi_d(y) = \varphi^y$.

Proposition 3.5 *Soient φ une forme bilinéaire sur E, $b = (e_1, \ldots, e_n)$ une base de E et A la matrice de φ dans cette base.*

1) *La matrice de φ_g dans les bases b et b^* (base duale) est égale à ${}^tA$.*

2) *La matrice de φ_d dans les bases b et b^* (base duale) est égale à A.*

Démonstration. 1) Le coefficient (i, j) de la matrice de φ_g dans les bases b et b^* est la $i^{\text{ème}}$ coordonnée de $\varphi_g(e_j) = \varphi_{e_j}$ dans la base $(e_1^*, \ldots, e_n^*)$, donc vaut (voir l'égalité 2.2) $\varphi_{e_j}(e_i) = \varphi(e_j, e_i) = a_{j,i}$.

2) De même, le coefficient (i, j) de la matrice de φ_d vaut $\varphi^{e_j}(e_i) = \varphi(e_i, e_j) = a_{i,j}$.

□

d) Orthogonalité - Forme quadratique non dégénérée

Définition *Soit β une forme bilinéaire symétrique sur E et q la forme quadratique associée.*

1) *Deux éléments x et y de E sont dits* orthogonaux *(pour β ou q) si $\beta(x, y) = 0$.*

2) *Si A est une partie de E, un élément x de E est dit orthogonal à A s'il est orthogonal à tous les éléments de A. On notera $A^\perp$ l'ensemble des éléments de E orthogonaux à A :*

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

Lemme 1 *Pour toute partie A de E :*

(i) $A^\perp$ *est un sous-espace vectoriel de* E *;*

(ii) si B *est une partie de* E *contenant* A*, alors* $B^\perp \subset A^\perp$ *;*

(iii) $A^\perp = (\mathrm{Vect}\, A)^\perp$ *;*

(iv) $\mathrm{Vect}(A) \subset (A^\perp)^\perp$.

Démonstration. *(i)* Soient x et y deux éléments de $A^\perp$ et λ un nombre réel. Pour tout z dans A, on a

$$\beta(x+\lambda y, z) = \beta(x,z) + \lambda\beta(y,z) = 0$$

donc $x+\lambda y$ est orthogonal à tout élément de A : cela signifie que $x+\lambda y$ appartient à $A^\perp$.

(ii) Soit $x \in B^\perp$; tout élément y de A est dans B donc on a $\beta(x,y) = 0$, ce qui montre que $x \in A^\perp$.

(iii) De $A \subset \mathrm{Vect}\, A$ on déduit que $(\mathrm{Vect}\, A)^\perp \subset A^\perp$ en utilisant *(ii)*. Considérons par ailleurs un élément x de $A^\perp$. Pour z élément de $\mathrm{Vect}\, A$, il existe $y_1, \ldots, y_k$ dans A et $\lambda_1, \ldots, \lambda_k$ dans $\mathbb{R}$ tels que $z = \lambda_1 y_1 + \cdots + \lambda_k y_k$. De là, on obtient

$$\beta(x,z) = \beta\left(x, \sum \lambda_i y_i\right) = \sum \lambda_i \beta(x, y_i) = 0 \quad \text{car } x \in A^\perp \text{ et } y_i \in A\ \forall i.$$

Ainsi, x est orthogonal à tous les éléments de $\mathrm{Vect}\, A$, donc appartient à $(\mathrm{Vect}\, A)^\perp$.

(iv) Un élément x de A est orthogonal à tout élément de $A^\perp$, donc appartient à $(A^\perp)^\perp = A^{\perp\perp}$. On a donc $A \subset A^{\perp\perp}$, et puisque $A^{\perp\perp}$ est un sous-espace vectoriel de E, on obtient $\mathrm{Vect}(A) \subset A^{\perp\perp}$.

□

Définition *Soit q une forme quadratique sur E.*

1) *L'orthogonal de E est appelé* radical *de q ; on le note* $\mathrm{rad}\, q$ *:* $\mathrm{rad}\, q = E^\perp$.

2) *q est dite* non dégénérée *si son radical est réduit à $\{0\}$; elle est dite dégénérée dans le cas contraire.*

3) *Le rang de q (ou de sa forme polaire) est l'entier* $\mathrm{rg}(q) = \dim E - \dim \mathrm{rad}\, q$.

Remarque : ainsi, q est non dégénérée si, et seulement si, $\mathrm{rg}(q) = \dim E$.

Lemme 2 *Soit q une forme quadratique et β sa forme polaire. Alors :* $\mathrm{rad}\, q = \ker \beta_g = \ker \beta_d$.

Démonstration. Rappelons que $\beta_g : E \to E^*$ est définie par $\beta_g(x) = \beta_x$ avec $\beta_x(y) = \beta(x,y)$. Par conséquent, $x \in \ker \beta_g$ si, et seulement si, la forme linéaire β_x est nulle, c'est-à-dire $\beta_x(y) = \beta(x,y) = 0$ pour tout y dans E : ceci signifie que $x \in \mathrm{rad}\, q$. On procède de même pour β_d.

□

On déduit immédiatement de ce lemme, des définitions et de la proposition 3.5 :

Corollaire 3.2
1) *Les propriétés suivantes sont équivalents :*

(i) q est non dégénérée, *(ii) β_g est un isomorphisme,* *(iii) β_d est un isomorphisme.*

2) *Si A est la matrice de q dans une base de E, alors* $\mathrm{rg}(q) = \mathrm{rg} A$.

Théorème 3.1 *Soient q une forme quadratique sur E et F un sous-espace vectoriel de E. On a :*

$$\dim F + \dim F^\perp = \dim E + \dim(F \cap \mathrm{rad}\, q).$$

Démonstration. Notons $\beta_g^F : F \to E^*$ la restriction de β_g à F.

• Le lemme 2 montre que $\ker \beta_g^F = F \cap \operatorname{rad} q$. Le théorème du rang nous permet donc d'affirmer que

$$\dim F = \dim\bigl(\operatorname{Im} \beta_g^F\bigr) + \dim(F \cap \operatorname{rad} q).$$

• D'autre part, puisque $\beta_g(x) = \beta_x$, on a

$$\begin{aligned} F^\perp &= \bigl\{y \in E \,/\, \beta(x,y) = 0 \ \forall x \in F\bigr\} = \bigl\{y \in E \,/\, \beta_x(y) = 0 \ \forall x \in F\bigr\} \\ &= \bigcap_{x \in F} \ker \beta_x = \bigcap_{l \in \operatorname{Im} \beta_g^F} \ker l = {}^\circ\bigl(\operatorname{Im} \beta_g^F\bigr) \end{aligned}$$

donc $\dim F^\perp = \dim E - \dim\bigl(\operatorname{Im} \beta_g^F\bigr)$ (proposition 2.5), ce qui nous donne l'égalité demandée.

□

Corollaire 3.3 *Si q est une forme bilinéaire non dégénérée et F un sous-espace vectoriel de E, alors :*

(i) $\dim F + \dim F^\perp = \dim E$;

(ii) $F^{\perp\perp} = F$.

Démonstration. Le point *(i)* est un cas particulier de théorème 3.1 (cas où $\operatorname{rad} q = \{0\}$). Ce premier point montre que $F^{\perp\perp}$ et F ont même dimension, donc sont égaux puisque $F \subset F^{\perp\perp}$ (lemme 1).

□

II _ Décomposition en somme de carrés - Signature

a) Algorithme de Gauss[1]

Théorème 3.2 *Soit q une forme quadratique non nulle sur un $\mathbb{R}$-espace vectoriel E de dimension finie n. Il existe des formes linéaires linéairement indépendantes $l_1, \dots, l_k$ et des nombres réels non nuls $\alpha_1, \dots, \alpha_k$ tels que*

$$q(x) = \sum_{i=1}^{k} \alpha_i l_i(x)^2 \quad \textit{pour tout } x \in E.$$

De plus, on a $\operatorname{rg}(q) = k$ et $\operatorname{rad} q = \bigcap_{i=1}^{k} \ker l_i$.

Démonstration. • Supposons dans un premier temps que q s'écrive sous la forme $q(x) = \sum_{i=1}^{k} \alpha_i l_i(x)^2$. L'application $\beta : E \times E \to \mathbb{R}$ définie par

$$\beta(x,y) = \sum_{i=1}^{k} \alpha_i l_i(x) l_i(y)$$

est bilinéaire symétrique et vérifie $\beta(x,x) = q(x)$ pour tout $x \in E$; c'est donc la forme polaire de q. Soit $F = \bigcap_{i=1}^{k} \ker l_i$. Pour tout $x \in F$ et tout $y \in E$, on a $\beta(x,y) = 0$ car $l_1(x) = \cdots = l_k(x) = 0$; par conséquent, $F \subset \operatorname{rad} q$. Supposons de plus que $(l_1, \dots, l_k)$ soit libre et complétons cette famille en une base $(l_1, \cdots, l_n)$ de E^*. D'après la proposition 2.2, il existe une base $(u_1, \dots, u_n)$ de E telle que $u_i^* = l_i$ pour tout $i \in [\![1,n]\!]$. Si $x \in \operatorname{rad} q$, alors, pour $i \in [\![1,k]\!]$, on a

$$0 = \beta(x, u_i) = \sum_{j=1}^{k} \alpha_j l_j(x) l_j(u_i) = \sum_{j=1}^{k} \alpha_j l_j(x) u_j^*(u_i) = \alpha_i l_i(x)$$

[1] Carl Friedrich Gauss, mathématicien, astronome et physicien allemand, 1777-1855.

donc $l_i(x) = 0$ si $\alpha_i \neq 0$. Ainsi, dans les conditions du théorème, on a $\operatorname{rad} q = F$. Enfin, d'après le corollaire 2.2, $\dim \operatorname{rad} q = n - k$ donc $\operatorname{rg}(q) = n - \dim \operatorname{rad} q = k$.

• Montrons à présent l'existence d'une telle décomposition pour q. Soit $(e_1, \dots, e_n)$ une base de E et $x = \sum_{i=1}^{n} x_i e_i$ un élément de E. Par bilinéarité de la forme polaire β de q, on a

$$\begin{aligned} q(x) = \beta(x, x) = \beta\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} x_j e_j\right) &= \sum_{1 \leq i,j \leq n} x_i\, x_j\, a_{i,j} \quad \text{où} \quad a_{i,j} = \beta(e_i, e_j) = a_{j,i} \\ &= \sum_{i=1}^{n} a_{i,i}\, x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{i,j}\, x_i\, x_j. \end{aligned}$$

Nous allons démontrer la première partie du théorème par récurrence sur n, la dimension de E. Si $n = 1$, le calcul ci-dessus montre que pour tout $x \in E$, $q(x) = a_{11} x_1^2$; on a donc le résultat en prenant $l_1 = e_1^*$. Nous supposons à présent que le théorème est démontré pour tout espace vectoriel de dimension $n - 1$.
Considérons une forme quadratique q sur un espace vectoriel E de dimension n.

<u>1^{er} cas : il existe $i \in [\![1, n]\!]$ tel que $a_{i,i} \neq 0$.</u> Quitte à effectuer une permutation sur les vecteurs de base, on peut supposer que $i = 1$. Soit $y = x - x_1 e_1 = x_2 e_2 + \cdots + x_n e_n$. Reprenant le calcul précédent, on obtient :

$$\begin{aligned} q(x) = q(x_1 e_1 + y) &= \beta(x_1 e_1 + y, x_1 e_1 + y) = a_{1,1}\, x_1^2 + 2x_1 \beta(e_1, y) + q(y) \\ &= \frac{1}{a_{1,1}} \Big(\underbrace{a_{1,1} x_1 + \beta(e_1, y)}_{l_1(x)}\Big)^2 + \underbrace{q(y) - \frac{\beta(e_1, y)^2}{a_{1,1}}}_{q'(y)}. \end{aligned}$$

L'application $l_1 : E \to \mathbb{R}$ est donnée par

$$l_1(x) = a_{1,1} x_1 + \beta(e_1, y) = x_1 \beta(e_1, e_1) + \beta(e_1, y) = \beta(e_1, x_1 e_1 + y) = \beta(e_1, x)$$

donc est une forme linéaire sur E. Notons F le sous-espace vectoriel de E engendré par $\{e_2, \dots, e_n\}$ et $q' : F \to \mathbb{R}$ l'application définie par

$$q'(u) = q(u) - \frac{\beta(e_1, u)^2}{a_{1,1}}.$$

C'est une forme quadratique sur F, de forme polaire β' donnée par

$$\beta'(u, v) = \beta(u, v) - \frac{1}{a_{1,1}} \beta(e_1, u) \beta(e_1, v).$$

Puisque F est de dimension $n-1$, il existe, par hypothèse de récurrence, $k-1$ formes linéaires sur F linéairement indépendantes $l'_2, \dots, l'_k$ et $k - 1$ nombres réels non nuls $\alpha_2, \dots, \alpha_k$ tels que, pour tout $u \in F$,

$$q'(u) = \sum_{i=2}^{k} \alpha_i\, l'_i(u)^2.$$

Pour $i \in [\![2, k]\!]$, introduisons la forme linéaire sur E $l_i = l'_i \circ \pi$ où $\pi : E \to F$ est la projection sur F suivant $\mathbb{R} e_1$ $\big(l_i(x) = l'_i(x - x_1 e_1)\big)$. Pour tout $x \in E$, on a, en posant $\alpha_1 = \frac{1}{a_{1,1}} \neq 0$,

$$q(x) = \alpha_1 l_1(x)^2 + \sum_{i=2}^{k} \alpha_i\, l_i(x)^2.$$

Pour conclure la démonstration dans ce premier cas, il reste à démontrer que les formes linéaires $l_1, \dots, l_k$ sont linéairement indépendantes. Si $\lambda_1 l_1 + \cdots + \lambda_k l_k = 0$ (avec $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$), alors

$$\lambda_1 l_1(e_1) + \cdots + \lambda_k l_k(e_1) = 0.$$

Mais $l_1(e_1) = \beta(e_1, e_1) = a_{1,1} \neq 0$ et, pour $i \geq 2$, $l_i(e_1) = 0$. Par conséquent, $\lambda_1 = 0$ et $\lambda_2 l_2 + \cdots + \lambda_k l_k = 0$. Donc $\lambda_2 l'_2 + \cdots + \lambda_k l'_k = 0$, ce qui implique $\lambda_2 = \cdots = \lambda_k = 0$ puisque $(l'_2, \dots, l'_k)$ est libre.

2$^{\text{ème}}$ cas : $a_{i,i} = 0$ pour tout $i \in [\![1, n]\!]$. Il existe alors $(i, j) \in [\![1, n]\!]$, $i < j$, tel que $a_{i,j} \neq 0$. Comme précédemment, on peut supposer que $a_{1,2} \neq 0$. Posant cette fois-ci $y = x - x_1e_1 - x_2e_2 = x_3e_3 + \cdots + x_ne_n$, on a

$$\begin{aligned}
q(x) &= q(x_1e_1 + x_2e_2 + y) = \beta(x_1e_1 + x_2e_2 + y, x_1e_1 + x_2e_2 + y)\\
&= a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + 2x_1\beta(e_1, y) + a_{2,2}x_2^2 + 2x_2\beta(e_2, y) + q(y)\\
&= 2a_{1,2}\left[\left(x_1 + \frac{\beta(e_2, y)}{a_{1,2}}\right)\left(x_2 + \frac{\beta(e_1, y)}{a_{1,2}}\right) - \frac{\beta(e_1, y)\beta(e_2, y)}{a_{1,2}^2}\right] + q(y)\\
&= \frac{a_{1,2}}{2}\left(x_1 + x_2 + \frac{\beta(e_1, y) + \beta(e_2, y)}{a_{1,2}}\right)^2 - \frac{a_{1,2}}{2}\left(x_1 - x_2 + \frac{\beta(e_2, y) - \beta(e_1, y)}{a_{1,2}}\right)^2 +\\
&\qquad \underbrace{q(y) - 2\frac{\beta(e_1, y)\beta(e_2, y)}{a_{1,2}}}_{q'(y)}\\
&= \frac{a_{1,2}}{2}\left(x_1 + x_2 + \frac{\beta(e_1 + e_2, y)}{a_{1,2}}\right)^2 - \frac{a_{1,2}}{2}\left(x_1 - x_2 + \frac{\beta(e_2 - e_1, y)}{a_{1,2}}\right)^2 + q'(y)\\
&= \frac{1}{2a_{1,2}}\Big(\underbrace{a_{1,2}x_1 + a_{1,2}x_2 + \beta(e_1 + e_2, y)}_{l_1(x)}\Big)^2 - \frac{1}{2a_{1,2}}\Big(\underbrace{a_{1,2}x_1 - a_{1,2}x_2 + \beta(e_2 - e_1, y)}_{l_2(x)}\Big)^2 + q'(y).
\end{aligned}$$

L'application $l_1 : E \to \mathbb{R}$ est donnée par

$$\begin{aligned}
l_1(x) &= \beta(e_1, e_2)x_1 + \beta(e_1, e_2)x_2 + \beta(e_1 + e_2, y) = \beta(x_1e_1, e_2) + \beta(e_1, x_2e_2) + \beta(e_1 + e_2, y)\\
&= \beta(x_1e_1, e_1) + \beta(e_2, x_2e_2) + \beta(x_1e_1, e_2) + \beta(e_1, x_2e_2) + \beta(e_1 + e_2, y) \quad \text{car } \beta(e_i, e_i) = a_{i,i} = 0\\
&= \beta(x_1e_1, e_1 + e_2) + \beta(e_1 + e_2, x_2e_2) + \beta(e_1 + e_2, y)\\
&= \beta(e_1 + e_2, x_1e_1 + x_2e_2 + y) = \beta(e_1 + e_2, x),
\end{aligned}$$

et de même, $l_2 : E \to \mathbb{R}$ est donnée par $l_2(x) = \beta(e_2 - e_1, x)$. Ce sont donc deux formes linéaires sur E. De plus, (l_1, l_2) est libre. En effet, ces deux formes linéaires vérifient

$$\begin{cases}
l_1(e_1 + e_2) = \beta(e_1 + e_2, e_1 + e_2) = \beta(e_1, e_1) + 2\beta(e_1, e_2) + \beta(e_2, e_2) = a_{1,1} + 2a_{1,2} + a_{2,2} = 2a_{1,2} \neq 0\\
l_2(e_1 + e_2) = \beta(e_2 - e_1, e_1 + e_2) = \beta(e_2, e_2) - \beta(e_1, e_1) = a_{2,2} - a_{1,1} = 0\\
l_2(e_2 - e_1) = \beta(e_2 - e_1, e_2 - e_1) = a_{2,2} - 2a_{1,2} + a_{1,1} = -2a_{1,2} \neq 0,
\end{cases}$$

donc l_1 et l_2 sont non nulles et non proportionnelles. Suivant le schéma du premier cas, notons F le sous-espace vectoriel de E engendré par $\{e_3, \ldots, e_n\}$ et $q' : F \to \mathbb{R}$ l'application définie par

$$q'(u) = q(u) - 2\frac{\beta(e_1, u)\beta(e_2, u)}{a_{1,2}}.$$

C'est une forme quadratique sur F, de forme polaire β' donnée par

$$\beta'(u, v) = \beta(u, v) - \frac{1}{a_{1,2}}\Big(\beta(e_1, u)\beta(e_2, v) + \beta(e_1, v)\beta(e_2, u)\Big).$$

Puisque F est de dimension $n-2$, il existe, par hypothèse de récurrence, $k-2$ formes linéaires sur F linéairement indépendantes $l'_3, \ldots, l'_k$ et $k-2$ nombres réels non nuls $\alpha_3, \ldots, \alpha_k$ tels que, pour tout $u \in F$,

$$q'(u) = \sum_{i=3}^{k} \alpha_i\, l'_i(u)^2.$$

Pour $i \in [\![3, k]\!]$, considérons la forme linéaire sur E $l_i = l'_i \circ \pi'$ où $\pi' : E \to F$ est la projection sur F suivant $\mathbb{R}e_1 \oplus \mathbb{R}e_2$ $\big(l_i(x) = l'_i(x - x_1e_1 - x_2e_2)\big)$. Pour tout $x \in E$, on a, en posant $\alpha_1 = \frac{1}{2a_{1,2}} \neq 0$,

$$q(x) = \alpha_1 l_1(x)^2 - \alpha_1 l_2(x)^2 + \sum_{i=3}^{k} \alpha_i\, l_i(x)^2.$$

Pour conclure ce second cas, il reste à démontrer que les formes linéaires $l_1, \dots, l_k$ sont linéairement indépendantes. Si $\lambda_1 l_1 + \cdots + \lambda_k l_k = 0$ (avec $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$), alors

$$\lambda_1 l_1(e_1 + e_2) + \cdots + \lambda_k l_k(e_1 + e_2) = 0.$$

Mais $l_1(e_1 + e_2) \neq 0$ et pour $i \geq 2$, $l_i(e_1 + e_2) = 0$. Par conséquent, $\lambda_1 = 0$ et $\lambda_2 l_2 + \cdots + \lambda_k l_k = 0$. De même, $l_2(e_1 - e_2) \neq 0$ et $l_i(e_1 - e_2) = 0$ pour $i \geq 3$ donc $\lambda_2 = 0$. Donc $\lambda_3 l'_3 + \cdots + \lambda_k l'_k = 0$, ce qui implique $\lambda_3 = \cdots = \lambda_k = 0$ puisque $(l'_3, \dots, l'_k)$ est libre.

□

b) Base orthogonale

On considère une forme quadratique q sur un espace vectoriel réel de dimension finie n, et on note β sa forme polaire.

Définition *Une famille $(e_1, \dots, e_k)$ de E est dite q-*orthogonale *si $\beta(e_i, e_j) = 0$ pour $i \neq j$.*

Proposition 3.6 *Une base b de E est q-orthogonale si, et seulement si, la matrice de q dans b est diagonale.*

Démonstration. Si $b = (e_1, \dots, e_n)$, le coefficient (i, j) de cette matrice vaut par définition $\beta(e_i, e_j)$.

□

Théorème 3.3 *Soit q une forme quadratique sur E. Soient $l_1, \dots, l_k$ k formes linéaires sur E linéairement indépendantes et $(\alpha_1, \dots, \alpha_k) \in (\mathbb{R}^*)^k$ tels que $q(x) = \sum_{i=1}^{k} \alpha_i l_i(x)^2$ pour tout $x \in E$ (existent d'après le théorème 3.2).*
Il existe une base q-orthogonale $(e_1, \dots, e_n)$ de E vérifiant :

- *(i) pour tout $i \in [\![1, k]\!]$, $q(e_i) = \alpha_i$;*
- *(ii) $(e_{k+1}, \dots, e_n)$ est une base du radical de q.*

Démonstration. Complétons $(l_1, \dots, l_k)$ en une base $(l_1, \dots, l_n)$ de E^* et considérons la base $(e_1, \dots, e_n)$ de E telle que $e_i^* = l_i$ pour tout $i \in [\![1, n]\!]$ (voir proposition 2.2). Pour $(p, q) \in [\![1, n]\!]^2$, on a

$$\beta(e_p, e_q) = \sum_{i=1}^{k} \alpha_i l_i(e_p) l_i(e_q) = \sum_{i=1}^{k} \alpha_i e_i^*(e_p) e_i^*(e_q) = \begin{cases} \alpha_i & \text{si } 1 \leq i = p = q \leq k, \\ 0 & \text{sinon.} \end{cases}$$

Par conséquent :

1) $e_{k+1}, \dots, e_n$ sont dans $\operatorname{rad} q$. Or ces vecteurs forment une famille libre et $\dim \operatorname{rad} q = n - \operatorname{rg}(q) = n - k$; la famille $(e_{k+1}, \dots, e_n)$ est donc une base de $\operatorname{rad} q$.

2) $(e_1, \dots, e_n)$ est q-orthogonale.

3) pour $i \in [\![1, k]\!]$, $q(e_i) = \beta(e_i, e_i) = \alpha_i$.

□

c) Signature d'une forme quadratique

Théorème 3.4 (Loi d'inertie de Sylvester[2]) *Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E. Il existe un couple $(s, t) \in \mathbb{N}^2$ tel que, pour toute base $(e_1, \dots, e_n)$ q-orthogonale de E, on ait :*

$$s = \operatorname{Card}\{e_i \,/\, q(e_i) > 0\} \quad \text{et} \quad t = \operatorname{Card}\{e_i \,/\, q(e_i) < 0\}.$$

[2] James Joseph Sylvester, mathématicien anglais, 1814-1897.

Démonstration. Soient $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_n)$ deux bases q-orthogonales et notons :

$$s = \mathrm{Card}\{e_i \,/\, q(e_i) > 0\}, \quad t = \mathrm{Card}\{e_i \,/\, q(e_i) < 0\},$$
$$s' = \mathrm{Card}\{f_i \,/\, q(f_i) > 0\}, \quad t' = \mathrm{Card}\{f_i \,/\, q(f_i) < 0\}.$$

Les matrices de q dans ces deux bases sont diagonales avec pour termes diagonaux respestivement $q(e_1), \ldots, q(e_n)$ et $q(f_1), \ldots, q(f_n)$. Par conséquent, on a $\mathrm{rg}(q) = s + t = s' + t'$ (corollaire 3.2).

Notons $I = \{i \in [\![1, n]\!] \,/\, q(e_i) > 0\}$, $J = \{i \in [\![1, n]\!] \,/\, q(f_i) \leq 0\}$, $F = \mathrm{Vect}\{e_i \,/\, i \in I\}$ et $G = \mathrm{Vect}\{f_i \,/\, i \in J\}$. Pour tout $x \in F \setminus \{0\}$, on a, si $x = \sum_{i \in I} \lambda_i e_i$,

$$q(x) = \sum_{i \in I} \lambda_i^2 q(e_i) > 0$$

et de même, pour tout $x \in G$, $q(x) \leq 0$. Par conséquent, $F \cap G = \{0\}$ et $\dim(F + G) = \dim F + \dim G \leq n$. Mais $\dim F = \mathrm{Card}(I) = s$ et $\dim G = n - s'$ donc on obtient $s + n - s' \leq n$, soit $s \leq s'$. En échangeant les rôles des bases (e_i) et (f_i), on obtient $s' \leq s$. Ainsi, $s = s'$ et donc $t = t'$.

□

Définition *Le couple (s, t) est appelé la* signature *de q.*

Remarque : si q est de signature (s, t), alors $\mathrm{rg}(q) = s + t$.

III _ Produit scalaire - Espaces vectoriels euclidiens

a) Définition

Définition *Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E.*

1) *q est dite* positive *si $q(x) \geq 0$ pour tout $x \in E$.*

2) *q est dite* définie *si, pour tout élément x non nul de E, $q(x) \neq 0$.*

Proposition 3.7 *Une forme quadratique définie est non dégénérée.*

Démonstration. Si q est dégénérée, il existe un élément x non nul dans E qui est orthogonal à tout élément de E. Il est en particulier orthogonal à lui-même, c'est-à-dire que $q(x) = 0$. Par conséquent, q n'est pas définie.

□

Remarque : la réciproque est fausse, comme le montre la forme quadratique q définie sur $\mathbb{R}^2$ par $q\big((x, y)\big) = x^2 - y^2$. En effet, elle est de rang 2 (sa signature vaut $(1, 1)$) donc non dégénérée, mais n'est pas définie car $q\big((1, 1)\big) = 0$.

Proposition 3.8 *Si q est une forme quadratique définie, toute famille q-orthogonale de vecteurs non nuls est libre.*

Démonstration. Soit $(e_1, \ldots, e_k)$ une famille q-orthogonale de vecteurs non nuls de E et supposons donné $(\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$ tel que $\sum_{i=1}^{k} \lambda_i e_i = 0$. Si β est la forme polaire de q, on a pour $j \in [\![1, k]\!]$:

$$0 = \beta\left(\sum_{i=1}^{k} \lambda_i e_i, e_j\right) = \sum_{i=1}^{k} \lambda_i \beta(e_i, e_j) = \lambda_j \beta(e_j, e_j) = \lambda_j q(e_j)$$

donc $\lambda_j = 0$ puisque $q(e_j) \neq 0$ (q définie).

□

Théorème 3.5 (Cauchy[3]-Schwarz[4]) *Soit q une forme quadratique positive sur un $\mathbb{R}$-espace vectoriel E et β sa forme polaire. On a, pour tout $(x,y) \in E$:*

$$\beta(x,y)^2 \leq q(x)q(y).$$

De plus, si q est définie, l'égalité est vérifiée si, et seulement si, (x,y) est liée.

Démonstration. • Fixons x et y non nuls dans E (si x ou y est nul, l'inégalité est claire) et considérons l'application f de $\mathbb{R}$ dans $\mathbb{R}$ définie par $f(\lambda) = q(x+\lambda y)$. Puisque q est positive, f ne prend que des valeurs positives et, par bilinéarité, on a $f(\lambda) = q(x) + 2\lambda\beta(x,y) + \lambda^2 q(y)$.

Si $q(y) = 0$, alors $q(x) + 2\lambda\beta(x,y)$ est positif pour tout $\lambda \in \mathbb{R}$ donc $\beta(x,y) = 0$; l'inégalité est donc vérifiée dans ce cas. Si $q(y) \neq 0$, f est un pôlynôme de degré 2 toujours positif, donc son discriminant est négatif :

$$\Delta = \beta(x,y)^2 - q(x)q(y) \leq 0,$$

ce qui donne l'inégalité demandée.

• Si $x = \lambda y$, alors $q(x) = \lambda^2 q(y)$ et $\beta(x,y) = \lambda\beta(y,y) = \lambda q(y)$ donc $\beta(x,y)^2 = \lambda^2 q(y)^2 = q(x)q(y)$.

• Supposons à présent que $\beta(x,y)^2 = q(x)q(y)$. Le discriminant de f est alors nul donc il existe $\lambda_0 \in \mathbb{R}$ tel que $f(\lambda_0) = 0$, c'est-à-dire $q(x+\lambda_0 y) = 0$. Si q est définie, alors $x + \lambda_0 y = 0$ donc (x,y) est liée.

□

Théorème 3.6 (Minkowski[5]) *Soit q une forme quadratique positive sur un $\mathbb{R}$-espace vectoriel E.*
On a, pour tout $(x,y) \in E^2$:

$$\sqrt{q(x+y)} \leq \sqrt{q(x)} + \sqrt{q(y)}.$$

*Si q est définie positive, alors $\sqrt{q(x+y)} = \sqrt{q(x)} + \sqrt{q(y)}$ si, et seulement si, x et y sont positivement liés (*ie *$\exists\ \lambda \geq 0$ tel que $x = \lambda y$ ou $y = \lambda x$).*

Démonstration. Si x ou y est nul, le résultat est clair. Considérons donc deux éléments x et y non nuls de E. Notons β la forme polaire de q. On obtient, en utilisant la bilinéarité de β et l'inégalité de Cauchy-Schwarz :

$$q(x+y) = q(x) + q(y) + 2\beta(x,y) \leq q(x) + q(y) + 2\sqrt{q(x)q(y)} = \left(\sqrt{q(x)} + \sqrt{q(y)}\right)^2,$$

ce qui donne l'inégalité recherchée. Ensuite, s'il y a égalité, alors $\beta(x,y) = \sqrt{q(x)q(y)}$ donc (x,y) est liée si q est de plus définie (théorème 3.5) : soit $\lambda \in \mathbb{R}^*$ tel que $x = \lambda y$. On a

$$\beta(x,y) = \lambda\beta(y,y) = \lambda q(y) \quad \text{et} \quad \sqrt{q(x)q(y)} = \sqrt{\lambda^2 q(y)^2} = |\lambda| q(y)$$

donc $\lambda = |\lambda|$ est positif. Réciproquement, si $x = \lambda y$ avec λ positif, alors $\beta(x,y) = \lambda q(y) = \sqrt{q(x)q(y)}$ donc $q(x+y) = \left(\sqrt{q(x)} + \sqrt{q(y)}\right)^2$.

□

Définition
1) *Un* produit scalaire *sur un $\mathbb{R}$-espace vectoriel E est une forme bilinéaire symétrique telle que la forme quadratique associée est définie positive.*

2) *Un* espace vectoriel euclidien *est un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.*

Remarque : Une forme quadratique est associée à un produit scalaire si, et seulement si, sa signature est égale à $(n,0)$ (où $n = \dim E$).
Remarque : Le fait que la forme quadratique associée à un produit scalaire est définie positive et le théorème 3.6 montrent que l'application $x \mapsto \sqrt{q(x)}$ est une norme sur E. Dorénavant, nous noterons $||x||$ la norme d'un vecteur x de E et $\langle x\,|y\rangle$ le produit scalaire de deux vecteurs x et y de E.

[3] Augustin Louis Cauchy, mathématicien français, 1789-1857.
[4] Hermann Amandus Schwarz, mathématicien allemand,1843-1921.
[5] Hermann Minkowski, mathématicien allemand, 1864-1909.

Proposition 3.9 *Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E. Alors F et $F^\perp$ sont supplémentaires : $E = F \oplus F^\perp$.*

Démonstration. Si $x \in F \cap F^\perp$, alors $||x||^2 = \langle x \,|x \rangle = 0$ donc $x = 0$: $F \cap F^\perp = \{0\}$. D'autre part, puisqu'un produit scalaire est non dégénéré, on déduit du théorème 3.1 que $\dim F + \dim F^\perp = \dim E$. Ces deux points montrent que F et $F^\perp$ sont supplémentaires dans E.

□

Cette proposition nous permet de donner la définition suivante :

Définition *Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E.*

1) On appelle projection orthogonale *sur F la projection sur F dans la direction $F^\perp$.*

2) On appelle symétrie orthogonale *par rapport à F la symétrie par rapport à F dans la direction $F^\perp$.*

Théorème 3.7 (Pythagore[6]) *Soient E un espace vectoriel euclidien et $(x, y) \in E^2$. Alors :*

$$x \text{ est orthogonal à } y \quad \text{si, et seulement si,} \quad ||x+y||^2 = ||x||^2 + ||y||^2.$$

Démonstration. Il suffit de remarquer que l'on a $||x+y||^2 = ||x||^2 + ||y||^2 + 2\langle x \,|y \rangle$ par bilinéarité.

□

b) Bases orthonormées

Définition *Une base orthonormée d'un espace vectoriel euclidien est une base orthogonale dont tous les vecteurs sont de norme 1.*

Proposition 3.10 *Tout espace vectoriel euclidien admet des bases orthonormées.*

Démonstration. Notons q la forme quadratique associée au produit scalaire. Nous avons déjà vu qu'il existe des bases q-orthogonales (théorème 3.3). Si $(e_1, \dots, e_n)$ est une telle base, alors $q(e_i) > 0$ pour tout $i \in [\![1, n]\!]$ (car q est définie positive) donc, si $f_i = \frac{e_i}{\sqrt{q(e_i)}}$, $(f_1, \dots, f_n)$ est une base orthonormée de E.

□

Théorème 3.8 (Procédé d'orthonormalisation de Gram[7]-Schmidt[8])
Soit $(e_1, ..., e_n)$ une base d'un espace vectoriel euclidien E. Il existe une base orthonormée $(f_1, ..., f_n)$ de E telle que, pour tout $k \in [\![1, n]\!]$, on ait $\mathrm{Vect}\,(e_1, ..., e_k) = \mathrm{Vect}\,(f_1, ..., f_k)$.

Démonstration. • La première étape est simple : il suffit de normaliser e_1, c'est-à-dire de poser $f_1 = \dfrac{e_1}{\|e_1\|}$. Les deux vecteurs e_1 et f_1 étant colinéaires, on a bien $\mathrm{Vect}(e_1) = \mathrm{Vect}(f_1)$.

• La deuxième étape consiste à projeter orthogonalement e_2 sur l'orthogonal de f_1 dans le plan $\mathrm{Vect}(f_1, e_2)$, c'est-à-dire de poser $e'_2 = e_2 - \langle e_2 \,|f_1 \rangle\, f_1$. En effet, on a par bilinéarité du produit scalaire

$$\langle e'_2 \,|f_1 \rangle = \langle e_2 \,|f_1 \rangle - \langle e_2 \,|f_1 \rangle \langle f_1 \,|f_1 \rangle = \langle e_2 \,|f_1 \rangle - \langle e_2 \,|f_1 \rangle = 0$$

car $\langle f_1 \,|f_1 \rangle = \|f_1\|^2 = 1$. Puisque la famille (f_1, e_2) est libre, le vecteur e'_2 n'est pas nul et on peut, pour finir, le normaliser en posant $f_2 = \dfrac{e'_2}{\|e'_2\|}$: la famille (f_1, f_2) est orthonormale par construction. La définition de f_2 et de e'_2 montre que $\mathrm{Vect}(f_1, f_2) = \mathrm{Vect}(f_1, e_2) = \mathrm{Vect}(e_1, e_2)$.

• On définit ensuite e'_3 comme étant le projeté orthogonal de e_3 sur l'orthogonal de $\mathrm{Vect}(f_1, f_2)$ dans $\mathrm{Vect}(f_1, f_2, e_3)$:

$$e'_3 = e_3 - \langle e_3 \,|f_1 \rangle\, f_1 - \langle e_3 \,|f_2 \rangle\, f_2.$$

[6] Pythagore de Samos, philosophe et mathématicien grec, vers 580-495 av. J.-C.
[7] Jørgen Pedersen Gram, mathématicien danois, 1850-1916.
[8] Erhard Schmidt, mathématicien allemand, 1876-1959.

On vérifie que e'_3 est orthogonal à f_1 et à f_2 :

$$\langle e'_3 \,|f_1 \rangle = \langle e_3 \,|f_1 \rangle - \langle e_3 \,|f_1 \rangle \langle f_1 \,|f_1 \rangle - \langle e_3 \,|f_2 \rangle \langle f_2 \,|f_1 \rangle = 0 \quad \text{puisque} \quad \langle f_1 \,|f_1 \rangle = 1 \quad \text{et} \quad \langle f_2 \,|f_1 \rangle = 0,$$

et de même $\langle e'_3 \,|f_2 \rangle = 0$. On conclut en posant $f_3 = \dfrac{e'_3}{\|e'_3\|}$ (e'_3 n'est pas nul car la famille (f_1, f_2, e_3) est libre).

• La démonstration générale consiste à poursuivre la construction des vecteurs f_k par récurrence en utilisant la formule suivante à l'étape k :

$$e'_k = e_k - \sum_{i=1}^{k-1} \langle e_k \,|f_i \rangle f_i,$$

puis en normalisant e'_k : $f_k = \dfrac{e'_k}{\|e'_k\|}$.

□

Proposition 3.11 (Calcul dans une base orthonormée) *Soient $(e_1, \ldots, e_n)$ une base orthonormée d'un espace vectoriel euclidien E et $x = \sum\limits_{i=1}^{n} x_i e_i$, $y = \sum\limits_{i=1}^{n} y_i e_i$, deux éléments de E. Alors :*

(i) $\forall\, i \in [\![1, n]\!],\ x_i = \langle x \,|e_i \rangle$; *(ii)* $\langle x \,|y \rangle = \sum\limits_{i=1}^{n} x_i y_i$; *(iii)* $||x||^2 = \sum\limits_{i=1}^{n} x_i^2$.

Démonstration. • Pour $j \in [\![1, n]\!]$, on a $\langle x \,|e_j \rangle = \left\langle \sum\limits_{i=1}^{n} x_i e_i \,|e_j \right\rangle = \sum\limits_{i=1}^{n} x_i \langle e_i \,|e_j \rangle = x_j ||e_j||^2 = x_j$.

• $\langle x \,|y \rangle = \left\langle \sum\limits_{i=1}^{n} x_i e_i \,\middle|\, \sum\limits_{j=1}^{n} y_j e_j \right\rangle = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} x_i y_j \langle e_i \,|e_j \rangle = \sum\limits_{i=1}^{n} x_i y_i$ car $\langle e_i \,|e_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$

• La dernière égalité est un cas particulier de la deuxième ($y = x$).

□

Proposition 3.12 (Changement de bases orthonormées) *Soient b et b' deux bases orthonormées d'un espace vectoriel euclidien et P la matrice de passage de b à b'. Alors :*

$${}^tPP = \mathrm{I}_n.$$

Démonstration. Il s'agit simplement de la formule de changement de bases (proposition 3.4) dans le cas particulier où la matrice de la forme bilinéaire (le produit scalaire ici) est I_n.

□

Définition *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite* orthogonale *si ${}^tAA = \mathrm{I}_n$. On notera $\mathrm{O}(n)$ l'ensemble des matrices orthogonales de taille n.*

Proposition 3.13 *Si $A \in \mathrm{O}(n)$, alors $A \in \mathrm{GL}_n(\mathbb{R})$ (avec $A^{-1} = {}^tA$) et* $\det A \in \{-1, 1\}$.

Démonstration. On a

$$1 = \det \mathrm{I}_n = \det\big({}^tAA\big) = \big(\det {}^tA\big)\big(\det A\big) = \big(\det A\big)\big(\det A\big) = \big(\det A\big)^2$$

donc $\det A \in \{-1, 1\}$. En particulier $\det A \neq 0$ donc A est inversible.

□

IV _ Adjoint d'un endomorphisme

Théorème 3.9 *Soient E un espace vectoriel euclidien et f un endomorphisme de E. Il existe un unique endomorphisme f^* de E vérifiant :*

$$\forall (x,y)\in E^2,\quad \langle f(x)\,|\,y\rangle = \langle x\,|\,f^*(y)\rangle\,.$$

Démonstration. Fixons une base orthonormée $(e_1,\ldots,e_n)$ de E.

• Si f^* existe, on a, pour $x\in E$ et $i\in[\![1,n]\!]$, $\langle f(e_i)\,|\,x\rangle = \langle e_i\,|\,f^*(x)\rangle$ donc, en utilisant la proposition 3.11, on obtient

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

Ceci montre l'unicité.

• Soit à présent $f^* : E\to E$ définie par $f^*(x)=\sum\limits_{i=1}^n \langle f(e_i)\,|\,x\rangle\, e_i$. Par linéarité à droite du produit scalaire, f^* est linéaire. Pour $x=\sum\limits_{i=1}^n x_ie_i\in E$ et $y\in E$, on a

$$\langle f(x)\,|\,y\rangle = \left\langle \sum_{i=1}^n x_i f(e_i)\middle|\, y\right\rangle = \sum_{i=1}^n x_i\,\langle f(e_i)\,|\,y\rangle\,,$$

$$\langle x\,|\,f^*(y)\rangle = \left\langle \sum_{i=1}^n x_ie_i\,\middle|\,\sum_{j=1}^n \langle f(e_j)\,|\,y\rangle\, e_j\right\rangle = \sum_{i=1}^n\sum_{j=1}^n x_i\,\langle f(e_j)\,|\,y\rangle\,\langle e_i\,|\,e_j\rangle = \sum_{i=1}^n x_i\,\langle f(e_i)\,|\,y\rangle\,,$$

ce qui montre que $\langle f(x)\,|\,y\rangle = \langle x\,|\,f^*(y)\rangle$.

$\square$

Définition *L'endomorphisme f^* ainsi défini est appelé* adjoint *de f.*

Proposition 3.14 *Pour tous endomorphismes f et g d'un espace vectoriel euclidien E on a :*

$$(i)\quad \forall\lambda\in\mathbb{R},\ (f+\lambda g)^* = f^*+\lambda g^* \qquad (ii)\quad (f\circ g)^* = g^*\circ f^* \qquad (iii)\quad f\in \mathrm{GL}(E) \implies \left(f^{-1}\right)^* = (f^*)^{-1}$$

$$(iv)\quad \left(f^*\right)^* = f \qquad (v)\quad \forall\,\mathbf{b}\ \text{b.o.n. de } E,\ \left[f^*\right]_{\mathbf{b}} = {}^t[f]_{\mathbf{b}}$$

Démonstration. *(i)* Pour tout $(x,y)\in E^2$, on a

$$\begin{aligned}\langle (f+\lambda g)(x)\,|\,y\rangle &= \langle f(x)+\lambda g(x)\,|\,y\rangle = \langle f(x)\,|\,y\rangle + \lambda\,\langle g(x)\,|\,y\rangle\\ &= \langle x\,|\,f^*(y)\rangle + \lambda\,\langle x\,|\,g^*(y)\rangle = \langle x\,|\,f^*(y)+\lambda g^*(y)\rangle\\ &= \langle x\,|\,(f^*+\lambda g^*)(y)\rangle\end{aligned}$$

donc $(f+\lambda g)^* = f^*+\lambda g^*$ (unicité de l'adjoint).

(ii) Pour $(x,y)\in E^2$, on a

$$\langle (f\circ g)(x)\,|\,y\rangle = \left\langle f\big(g(x)\big)\,|\,y\right\rangle = \langle g(x)\,|\,f^*(y)\rangle = \left\langle x\,|\,g^*\big(f^*(y)\big)\right\rangle = \langle x\,|\,(g^*\circ f^*)(y)\rangle$$

donc, toujours par unicité, $(f\circ g)^* = g^*\circ f^*$.

(iii) C'est une conséquence de la propriété (ii) :

$$\mathrm{Id} = \mathrm{Id}^* = (f\circ f^{-1})^* = \left(f^{-1}\right)^*\circ f^* \implies f^* \text{ inversible et } \left(f^*\right)^{-1} = \left(f^{-1}\right)^*.$$

(iv) Cette quatrième identité découle directement de la symétrie du produit scalaire (et de l'unicité de l'adjoint) :

$$\forall (x,y)\in E^2,\ \langle f^*(x)\,|\,y\rangle = \langle y\,|\,f^*(x)\rangle = \langle f(y)\,|\,x\rangle = \langle x\,|\,f(y)\rangle\,.$$

(v) Si $[f]_{\mathbf{b}} = (a_{i,j})_{1\le i,j\le n}$, alors $a_{i,j} = \langle f(e_j)\,|\,e_i\rangle$ (la $i^{\text{ème}}$ coordonnée de $f(e_j)$ dans la base orthonormée $\mathbf{b}=(e_1,\ldots,e_n)$; voir proposition 3.11) donc $a_{i,j} = \langle e_j\,|\,f^*(e_i)\rangle$. D'autre part, si $[f^*]_{\mathbf{b}} = (b_{i,j})_{1\le i,j\le n}$, alors $b_{i,j} = \langle f^*(e_j)\,|\,e_i\rangle = \langle e_i\,|\,f^*(e_j)\rangle = a_{j,i}$. On a donc bien $[f^*]_{\mathbf{b}} = {}^t[f]_{\mathbf{b}}$.

$\square$

Proposition 3.15 *Soit E un espace vectoriel euclidien et f un endomorphisme de E. Alors :*

$$\ker f^* = \big(\operatorname{Im} f\big)^{\perp} \qquad \textit{et} \qquad \operatorname{Im} f^* = \big(\ker f\big)^{\perp}.$$

Démonstration. • Soit $x \in \ker f^*$; pour tout $y \in E$, on a $\langle x \,|\, f(y)\rangle = \langle f^*(x) \,|\, y\rangle = \langle 0 \,|\, y\rangle = 0$ donc x est orthogonal à tout élément de $\operatorname{Im} f$. Ceci montre que $\ker f^* \subset \big(\operatorname{Im} f\big)^{\perp}$.

• Soit $x \in \big(\operatorname{Im} f\big)^{\perp}$; pour tout $y \in E$, on a $\langle f^*(x) \,|\, y\rangle = \langle x \,|\, f(y)\rangle = 0$ donc $f^*(x) \in E^{\perp} = \{0\}$. Par conséquent, $\big(\operatorname{Im} f\big)^{\perp} = \ker f^*$.

• On obtient la deuxième égalité en appliquant la première à f^* :

$$\ker f = \ker(f^*)^* = \big(\operatorname{Im} f^*\big)^{\perp} \implies \big(\ker f\big)^{\perp} = \operatorname{Im} f^*.$$

$\square$

Proposition 3.16 *Soient f un endomorphisme d'un espace vectoriel euclidien E et F un sous-espace de E stable par f. Alors $F^{\perp}$ est stable par f^*.*

Démonstration. Soit x un élément de $F^{\perp}$. Pour tout y dans F, on a

$$\langle f^*(x) \,|\, y\rangle = \langle x \,|\, f(y)\rangle = 0 \quad \text{car } f(y) \in F.$$

Ainsi, $f^*(x)$ est orthogonal à tout élément de F : $f^*(x) \in F^{\perp}$.

$\square$

V _ Endomorphismes orthogonaux

a) Définition

Théorème 3.10 *Soient E un espace vectoriel euclidien et f un endomorphisme de E. Les sept assertions suivantes sont équivalentes :*

(1) $f \circ f^* = \mathrm{Id}_E = f^* \circ f$ (ie f *est inversible et* $f^{-1} = f^*$) ;

(2) f préserve le produit scalaire : $\forall\,(x,y) \in E^2$, $\langle f(x) \,|\, f(y)\rangle = \langle x \,|\, y\rangle$;

(3) f préserve la norme : $\forall\, x \in E$, $||f(x)|| = ||x||$;

(4) l'image par f de toute base orthonormée de E est une base orthonormée de E ;

(5) il existe une base orthonormée de E dont l'image par f est une base orthonormée de E ;

(6) la matrice de f dans toute base orthonormée de E est orthogonale ;

(7) il existe une base orthonormée de E dans laquelle la matrice de f est orthogonale.

Démonstration. $\underline{(1) \Longrightarrow (2)}$ Pour tout $(x,y) \in E^2$, on a $\langle f(x) \,|\, f(y)\rangle = \big\langle x \,\big|\, f^*\big(f(y)\big)\big\rangle = \langle x \,|\, y\rangle$.

$\underline{(2) \Longrightarrow (3)}$ Pour tout $x \in E$, on a $||f(x)||^2 = \langle f(x) \,|\, f(x)\rangle = \langle x \,|\, x\rangle = ||x||^2$.

$\underline{(3) \Longrightarrow (2)}$ Pour tout $(x,y) \in E^2$, on a

$$\begin{aligned}
\langle f(x) \,|\, f(y)\rangle &= \frac{1}{2}\Big[||f(x)+f(y)||^2 - ||f(x)||^2 - ||f(y)||^2\Big] = \frac{1}{2}\Big[||f(x+y)||^2 - ||f(x)||^2 - ||f(y)||^2\Big] \\
&= \frac{1}{2}\Big[||x+y||^2 - ||x||^2 - ||y||^2\Big] \\
&= \langle x \,|\, y\rangle .
\end{aligned}$$

<u>(2) $\Longrightarrow$ (4)</u> Si $b = (e_1, \ldots, e_n)$ est une base orthonormée de E, alors, pour tout $(i,j) \in [\![1,n]\!]^2$, on a

$$\langle f(e_i) \,|\, f(e_j) \rangle = \langle e_i \,|\, e_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

donc $\big(f(e_1), \ldots, f(e_n)\big)$ est une base orthonormée de E (c'est une base d'après la proposition 3.8).

<u>(4) $\Longrightarrow$ (5)</u> Si c'est vrai pour toutes les bases orthonormées, ça l'est pour au moins une !

<u>(5) $\Longrightarrow$ (7)</u> Soit b une base orthonormée dont l'image par f est une base orthonormée. La matrice de f dans les bases b et $f(b)$ est alors la matrice de passage d'une base orthonormée à une autre, donc est orthogonale d'après la proposition 3.12.

<u>(7) $\Longrightarrow$ (6)</u> Soit b une base orthonormée telle que $[f]_b$ soit orthogonale : ${}^t[f]_b \cdot [f]_b = \mathrm{I}_n$. Soient b' une base orthonormée et P la matrice de passage de b à b' : ${}^tPP = \mathrm{I}_n$ d'après la proposition 3.12. Puisque $[f]_{b'} = P^{-1}[f]_bP = {}^tP[f]_bP$, on a

$${}^t[f]_{b'} \cdot [f]_{b'} = {}^t\Big({}^tP[f]_bP\Big) \cdot \Big({}^tP[f]_bP\Big) = {}^tP\,{}^t[f]_bP\,{}^tP[f]_bP = {}^tP\,{}^t[f]_b\mathrm{I}_n[f]_bP = {}^tP\,{}^t[f]_b[f]_bP = {}^tPP = \mathrm{I}_n,$$

c'est-à-dire que $[f]_{b'}$ est orthogonale.

<u>(6) $\Longrightarrow$ (7)</u> Il n'y a rien à dire.

<u>(7) $\Longrightarrow$ (1)</u> Soit b une base orthonormée de E telle que $[f]_b$ soit orthogonale. Puisque la matrice de f^* dans la base b est ${}^t[f]_b$ (proposition 3.14), on a

$$[f^* \circ f]_b = [f^*]_b \cdot [f]_b = {}^t[f]_b \cdot [f]_b = \mathrm{I}_n = [f]_b \cdot {}^t[f]_b = [f \circ f^*]_b$$

donc $f \circ f^* = \mathrm{Id}_E = f^* \circ f$.

□

Définition *Un endomorphisme vérifiant les propriétés équivalentes précédentes est dit* orthogonal.
On note $\mathrm{O}(E)$ *l'ensemble des endomorphismes orthogonaux de* E *: cet ensemble est appelé* groupe orthogonal.

Remarque : On utilise également les termes *d'isométrie vectorielle* pour désigner un endomorphisme orthogonal.

Exemple. La symétrie orthogonale par rapport à un sous-espace vectoriel F de E est un endomorphisme orthogonal de E. En effet, si $x \in E$ se décompose en $x = u + v$ avec $u \in F$ et $v \in F^\perp$, alors $s(x) = u - v$ et on a, en utilisant le théorème de Pythagore :

$$||x||^2 = ||u||^2 + ||v||^2 = ||u||^2 + ||-v||^2 = ||s(x)||^2.$$

Définition *Une* réflexion *d'un espace vectoriel euclidien est une symétrie orthogonale par rapport à un hyperplan.*

Proposition 3.17 *Soient f et g deux endomorphismes orthogonaux d'un espace vectoriel euclidien E. Alors :*

- *(i)* $f \circ g$ *est orthogonal ;*
- *(ii)* $f^{-1} = f^*$ *est orthogonal ;*
- *(iii)* $\det f \in \{-1, 1\}$ *;*
- *(iv)* $\mathrm{Sp}(f) \subset \{-1, 1\}$.

Démonstration. *(i) et (ii)* Si f et g préservent la norme, il en est de même de $f \circ g$ et f^{-1}.

(iii) La matrice A de f dans une base orthonormée est orthogonale donc $\det f = \det A \in \{-1, 1\}$ (proposition 3.13).

(iv) Si, pour x élément non nul de E, on a $f(x) = \lambda x$ ($\lambda \in \mathbb{R}$), alors

$$||x|| = ||f(x)|| = ||\lambda x|| = |\lambda| \cdot ||x|| \implies |\lambda| = 1.$$

□

b) Orientation

Notons $\mathscr{B}$ l'ensemble des bases orthonormées de E. Pour $\mathbf{b} \in \mathscr{B}$, on note

$$\mathscr{B}_{\mathbf{b}}^{+} = \{\mathbf{e} \in \mathscr{B} \,/\, \det_{\mathbf{b}} \mathbf{e} = 1\} \quad \text{et} \quad \mathscr{B}_{\mathbf{b}}^{-} = \{\mathbf{e} \in \mathscr{B} \,/\, \det_{\mathbf{b}} \mathbf{e} = -1\}.$$

Puisque $\det_{\mathbf{b}} \mathbf{e} = \det P$ où P est la matrice de passage de $\mathbf{b}$ à $\mathbf{e}$, on a $\det_{\mathbf{b}} \mathbf{e} = \pm 1$ (P est orthogonale) donc $\left(\mathscr{B}_{\mathbf{b}}^{+}, \mathscr{B}_{\mathbf{b}}^{-}\right)$ forme une partition de $\mathscr{B}$.

Lemme 3 *Soient $\mathbf{b}_1$ et $\mathbf{b}_2$ deux bases orthonormées de E et $\varepsilon = \det_{\mathbf{b}_1} \mathbf{b}_2$. Alors les partitions associées à $\mathbf{b}_1$ et $\mathbf{b}_2$ coïncident. Plus précisément :*

1) $\varepsilon = 1 \Longrightarrow \mathscr{B}_{\mathbf{b_1}}^{+} = \mathscr{B}_{\mathbf{b_2}}^{+}$ *et* $\mathscr{B}_{\mathbf{b_1}}^{-} = \mathscr{B}_{\mathbf{b_2}}^{-}$;

2) $\varepsilon = -1 \Longrightarrow \mathscr{B}_{\mathbf{b_1}}^{+} = \mathscr{B}_{\mathbf{b_2}}^{-}$ *et* $\mathscr{B}_{\mathbf{b_1}}^{-} = \mathscr{B}_{\mathbf{b_2}}^{+}$.

Démonstration. C'est une conséquence du fait que pour toute base orthonormée $\mathbf{b}$ de E, on a

$$\det_{\mathbf{b}_1} \mathbf{b} = \left(\det_{\mathbf{b}_1} \mathbf{b_2}\right)\left(\det_{\mathbf{b}_2} \mathbf{b}\right) = \varepsilon \det_{\mathbf{b}_2} \mathbf{b}.$$

□

Définition *Soit E un espace vectoriel euclidien.* Orienter E, *c'est choisir une des deux parties de la partition précédente. Les bases appartenant à cette partie sont dites* directes, *les autres* indirectes.

Proposition 3.18 *Soient $\mathbf{b}$ une base orthonormée directe d'un espace vectoriel euclidien orienté E et f un endomorphisme orthogonal de E. Alors :*

$$\textit{la base orthonormée } f(\mathbf{b}) \textit{ est directe si, et seulement si, } \det f = 1.$$

Démonstration. On a $\det f = \det[f]_{\mathbf{b}} = \det_{\mathbf{b}} f(\mathbf{b}) = \begin{cases} 1 & \text{si } f(\mathbf{b}) \text{ est directe} \\ -1 & \text{sinon.} \end{cases}$

□

Définition *On notera* $\mathrm{SO}(E)$ *l'ensemble des endomorphismes orthogonaux préservant l'orientation : c'est le* groupe spécial orthogonal.

c) Cas de la dimension 2

Proposition 3.19 *Soient E un espace vectoriel euclidien orienté de dimension 2 et f un endomorphisme orthogonal de E.*

1) Si $\det f = 1$, il existe $\theta \in \mathbb{R}$ tel que, pour toute base orthonormée directe b de E, on ait $[f]_b = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

*2) Si $\det f = -1$, alors f est une réflexion (*ie *une symétrie orthogonale par rapport à une droite).*

Définition *Une* rotation *d'un espace vectoriel euclidien de dimension 2 est un endomorphisme orthogonal de E de déterminant 1. Lorsque E est orienté, la rotation est dite d'angle θ si sa matrice dans une base orthonormée directe (et donc dans toutes) est de la forme* $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Démonstration. Soit $b = (e_1, e_2)$ une base orthonormée de E et notons $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ la matrice de f dans cette base. Puisque $f(e_1) = ae_1 + be_2$ et $||f(e_1)|| = ||e_1|| = 1$, on a $a^2 + b^2 = 1$. Il existe donc $\theta \in \mathbb{R}$ tel que $(a, b) = (\cos\theta, \sin\theta)$.
Ensuite, puisque e_2 est orthogonal à e_1, $f(e_2)$ est orthogonal à $f(e_1)$. Mais $f(e_1)^{\perp}$ est de dimension 2-1=1, donc engendré par $be_1 - ae_2$ (ce vecteur est non nul et orthogonal à $f(e_1) = ae_1 + be_2$). Il existe donc $\lambda \in \mathbb{R}$ tel que $f(e_2) = \lambda(be_1 - ae_2)$. De $||f(e_2)|| = ||e_2|| = 1$, on déduit que $\lambda^2(a^2 + b^2) = 1$, soit $\lambda^2 = 1$. On obtient donc $[f]_b = \begin{pmatrix} \cos\theta & \lambda\sin\theta \\ \sin\theta & -\lambda\cos\theta \end{pmatrix}$ avec $\det f = \det[f]_b = -\lambda$.

<u>1^{er} cas : $\det f = -1$.</u> Alors $[f]_b = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{smallmatrix}\right)$ donc

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

Ainsi f est diagonalisable et $E = \ker(f - \mathrm{Id}_E) \oplus \ker(f + \mathrm{Id}_E)$: f est la symétrie par rapport à $\ker(f - \mathrm{Id}_E)$ dans la direction $\ker(f + \mathrm{Id}_E)$. Nous concluons ce cas en démontrant que ces deux sous-espaces sont orthogonaux. Si $x \in \ker(f - \mathrm{Id}_E)$ et $y \in \ker(f + \mathrm{Id}_E)$, alors

$$\langle x \,|y\,\rangle = \langle f(x) \,|\, f(y)\,\rangle = \langle x \,|{-y}\,\rangle = -\langle x \,|y\,\rangle \qquad \text{donc} \qquad \langle x \,|y\,\rangle = 0.$$

<u>$2^{\text{ème}}$ cas : $\det f = 1$.</u> Alors $[f]_b = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Supposons que b soit directe et considérons une seconde base orthonormée directe b'. La matrice de passage P de b à b' est alors une matrice spéciale orthogonale. En appliquant ce que l'on vient de faire pour $[f]_b$ à P, on déduit l'existence de $\alpha \in \mathbb{R}$ tel que $P = \left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right)$. On a $[f]_{b'} = P^{-1}[f]_b P$. Mais

$$\begin{aligned} [f]_b P &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\alpha - \sin\theta\sin\alpha & -\cos\theta\sin\alpha - \sin\theta\cos\alpha \\ \sin\theta\cos\alpha + \cos\theta\sin\alpha & -\sin\theta\sin\alpha + \cos\theta\cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\alpha) & -\sin(\theta+\alpha) \\ \sin(\theta+\alpha) & \cos(\theta+\alpha) \end{pmatrix} = P[f]_b \end{aligned}$$

donc $[f]_{b'} = P^{-1}P[f]_b = [f]_b$.

$\square$

d) Classification en dimension quelconque

Théorème 3.11 *Soient E un espace vectoriel euclidien et f un endomorphisme orthogonal de E. Il existe une base orthonormée de E dans laquelle la matrice de f est de la forme*

$$\begin{pmatrix} \boxed{\mathrm{I}_p} & & & & 0 \\ & \boxed{-\mathrm{I}_q} & & & \\ & & \boxed{R_{\theta_1}} & & \\ & 0 & & \ddots & \\ & & & & \boxed{R_{\theta_k}} \end{pmatrix} \qquad \text{où} \quad R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Remarque : dans l'énoncé ci-dessus, p, q ou k peuvent éventuellement être nuls.

Lemme 4 *Tout endomorphisme d'un $\mathbb{R}$-espace vectoriel E admet un sous-espace stable de dimension 1 ou 2.*

Démonstration. Considérons la décomposition du polynôme minimal μ_f de f en facteurs irréductibles dans $\mathbb{R}[X]$: $\mu_f = P_1 \cdots P_s$ avec P_i de degré 1 ou 2 (les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 dont le discriminant est strictement négatif). Puisque $0 = \mu_f(f) = P_1(f) \circ \cdots \circ P_s(f)$, il existe $i \in [\![1, s]\!]$ tel que $P_i(f)$ n'est pas bijectif, donc pas injectif ($\dim E < \infty$) : soit $x \in E \setminus \{0\}$ tel que $P_i(f)(x) = 0$.

• Si P_i est de degré 1, alors $P_i = X - \lambda$ avec $\lambda \in \mathbb{R}$ et $0 = P_i(f)(x) = f(x) - \lambda x$: x est un vecteur propre pour f. Par conséquent, $\mathrm{Vect}(x)$ est un sous-espace de dimension 1 de E, stable par f.

• Si P_i est de degré 2, alors $P_i = X^2 + aX + b$ avec $(a, b) \in \mathbb{R}^2$ et $0 = P_i(f)(x) = f^2(x) + af(x) + bx$ donc $f^2(x) = -af(x) - bx \in \mathrm{Vect}\big(x, f(x)\big)$. Ainsi, $\mathrm{Vect}\big(x, f(x)\big)$ est un sous-espace stable par f, et est de dimension 1 ou 2.

$\square$

Démonstration du théorème 3.11. On procède par récurrence sur $n = \dim E$.
En dimension 1, les seuls endomorphismes orthogonaux sont $-\mathrm{Id}$ et Id ; leur matrice dans n'importe quelle base est (± 1). Supposons établi le résultat pout tout endomorphisme orthogonal d'un espace vectoriel euclidien

de dimension inférieure ou égale à $n-1$ et considérons un endomorphisme orthogonal f d'un espace vectoriel euclidien E de dimension n.

• Supposons dans un premier temps que f admette un vecteur propre x : $f(x) = \pm x$ d'après la proposition 3.17. Notons H l'hyperplan $\mathrm{Vect}(x)^{\perp}$. Puisque $\mathrm{Vect}(x)$ est stable par f, H est stable par $f^* = f^{-1}$ (proposition 3.16), donc est stable par f[9] : notons f_H l'endomorphisme de H induit par f. C'est un endomorphisme orthogonal de l'espace vectoriel euclidien H (car il préserve la norme : voir théorème 3.10). Par hypothèse de récurrence ($\dim H = n-1$), il existe une base orthonormée b_1 de H dans laquelle la matrice de f_H est de la forme voulue. Alors, si $b = \left\{\frac{x}{||x||}\right\} \cup b_1$, b est une base orthonormée de E dans laquelle la matrice de f est

$$\left(\begin{array}{c|c} \pm 1 & 0 \\ \hline 0 & [f_H]_{b_1} \end{array}\right).$$

Quitte à échanger deux vecteurs dans b, on obtient la matrice voulue.

• Si f n'admet pas de vecteur propre, alors il n'admet pas de sous-espace stable de dimension 1. Mais d'après le lemme 4, il existe un sous-espace F de dimension 2 stable par f. Comme précédemment, $F^{\perp}$ est stable par f : notons respectivement f_F et $f_{F^{\perp}}$ les endomorphismes de F et $F^{\perp}$ induits par f.

- ∗ D'après la proposition 3.19, il existe une base orthonormée b_1 de F et $\theta \in \mathbb{R}$ tel que $[f_F]_{b_1} = R_\theta$ (f_F n'est pas une symétrie car f n'admet pas de vecteur propre).
- ∗ Par hypothèse de récurrence, il existe une base orthonormée b_2 de $F^{\perp}$ telle que $[f_{F^{\perp}}]_{b_2}$ soit du type cherché.

Alors, $b = b_2 \cup b_1$ est une base orthonormée de E dans laquelle la matrice de f est de la forme demandée.

□

Cas de la dimension 3. Soient E un espace vectoriel euclidien de dimension 3 et f un endomorphisme orthogonal de E. Le théorème 3.11 nous permet d'affirmer qu'il existe une base orthonormée (e_1, e_2, e_3) de E dans laquelle la matrice de f est de la forme suivante (éventuellement θ est égal à 0 ou à π)

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Supposons E orienté par (e_1, e_2, e_3) et la droite $\mathbb{R}e_1$ orientée par e_1. Le plan $e_1^{\perp}$ est alors naturellement[10] orienté par (e_2, e_3). Nous obtenons les différents endomorphismes orthogonaux suivants :

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ f est l'application identité.

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ f est la réflexion par rapport au plan $\mathrm{Vect}(e_1, e_2)$.

$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ f est l'homothétie vectorielle de rapport -1.

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$ On dit que f est la *rotation* d'axe orienté par e_1 et d'angle θ. Dans le cas où $\theta = \pi$, la matrice est $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ et on parle dans ce cas de demi-tour d'axe $\mathrm{Vect}(e_1)$.

$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$ On dit que f est l'*anti-rotation* d'axe orienté par e_1 et d'angle θ. C'est la composée commutative de la rotation d'axe $\mathbb{R}e_1$ et d'angle θ avec la réflexion de plan $\mathrm{Vect}(e_2, e_3)$.

---

[9] f étant un isomorphisme, on a $f^{-1}(H) \subset H \Leftrightarrow f^{-1}(H) = H \Leftrightarrow f(H) = H$.

[10] La droite $\mathbb{R}e_1$ étant orientée par e_1, une orientation *naturelle* du plan $e_1^{\perp}$ est le choix d'une base (u, v) (que l'on dira *directe*) de celui-ci telle que (e_1, u, v) soit une base directe de E.

VI _ Endomorphismes symétriques

a) Théorème de réduction

Définition *Un endomorphisme f d'un espace vectoriel euclidien est dit* symétrique *s'il coïncide avec son adjoint.*

Remarque : Par définition de f^*, ceci équivaut à : $\forall (x,y) \in E^2,\ \langle f(x)\,|\,y\rangle = \langle x\,|\,f(y)\rangle$.

Remarque : Des propriétés de l'adjoint (voir proposition 3.14), on déduit que l'ensemble $\mathscr{S}(E)$ des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathscr{L}(E)$, stable par composition.

Exemples. • Toute symétrie orthogonale s : $s^* = s^{-1} = s$.

• Les homothéties : si $h = \lambda \mathrm{Id}_E$, alors $\langle h(x)\,|\,y\rangle = \langle \lambda x\,|\,y\rangle = \lambda \langle x\,|\,y\rangle = \langle x\,|\,\lambda y\rangle = \langle x\,|\,h(y)\rangle$ pour tout $(x,y) \in E^2$.

• Les projecteurs orthogonaux. En effet, en F un sous-espace vectoriel de E et π la projection orthogonale sur F. Pour $(x,y) \in E^2$, on a, si $x = u + v$ et $y = u' + v'$ avec $(u,u') \in F$ et $(v,v') \in F^\perp$:

$$\langle \pi(x)\,|\,y\rangle = \langle u\,|\,u'+v'\rangle = \langle u\,|\,u'\rangle + \langle u\,|\,v'\rangle = \langle u\,|\,u'\rangle\,,$$
$$\langle x\,|\,\pi(y)\rangle = \langle u+v\,|\,u'\rangle = \langle u\,|\,u'\rangle\,.$$

Proposition 3.20 *Soient b une base orthonormée de E et $f \in \mathscr{L}(E)$. Alors :*

$$f \text{ est symétrique} \iff [f]_b \text{ est une matrice symétrique.}$$

Démonstration. Cela provient du fait que, b étant orthonormée, $[f^*]_b = {}^t[f]_b$ (proposition 3.14).

□

Proposition 3.21
1) Si F est un sous-espace stable par un endomorphisme symétrique f, alors $F^\perp$ est également stable par f.

2) Si f est un endomorphisme symétrique, alors $(\ker f)^\perp = \operatorname{Im} f$; en particulier, on a $E = \ker f \overset{\perp}{\oplus} \operatorname{Im} f$.

Démonstration. 1) D'après la proposition 3.16, si F est stable par f, alors $F^\perp$ est stable par f^*, donc par f si $f = f^*$.

2) D'après la proposition 3.15, on a $(\ker f)^\perp = \operatorname{Im} f^* = \operatorname{Im} f$ si f est symétrique.

□

Théorème 3.12 *Soient E un espace vectoriel euclidien et f un endomorphisme symétrique de E. Alors E admet une base orthonormée formée de vecteurs propres pour f (*ie *f est diagonalisable dans une base orthonormée).*

Lemme 5 *Si $A \in \mathscr{M}_n(\mathbb{R})$ est une matrice symétrique, alors son polynôme caractéristique χ_A est scindé sur $\mathbb{R}$.*

Démonstration. Puisque χ_A est scindé sur $\mathbb{C}$, il suffit de montrer que si $\lambda \in \mathbb{C}$ est valeur propre de A, alors $\lambda \in \mathbb{R}$. Pour un tel λ, il existe X non nul dans $\mathscr{M}_{n,1}(\mathbb{C})$ tel que $AX = \lambda X$. Alors, $A\overline{X} = \overline{A}\,\overline{X} = \overline{\lambda}\,\overline{X}$ car A est réelle ($\overline{A}$ et $\overline{X}$ désignent les matrices dont les coefficients sont les conjugués des coefficients de A et X). De là, on obtient :

$$\begin{aligned} {}^t\!\left(A\overline{X}\right) &= {}^t\overline{X}\,{}^tA = {}^t\overline{X}A \quad \text{car } A \text{ est symétrique} \\ &= {}^t\!\left(\overline{\lambda}\,\overline{X}\right) = \overline{\lambda}\,{}^t\overline{X}, \end{aligned}$$

ce qui implique $\overline{\lambda}\,{}^t\overline{X}X = {}^t\overline{X}AX = \lambda\,{}^t\overline{X}X$. Mais si $X = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, alors ${}^t\overline{X}X = \sum_{i=1}^{n} |z_i|^2$ qui est non nul puisque $X \neq 0$. Par conséquent, $\overline{\lambda} = \lambda$, c'est-à-dire $\lambda \in \mathbb{R}$.

□

Démonstration du théorème 3.12. On procède par récurrence sur $n = \dim E$, le cas de la dimension 1 étant clair (n'importe quel vecteur unitaire convient). On suppose donc que tout endomorphisme symétrique d'un espace vectoriel euclidien de dimension n est diagonalisable dans une base orthonormée et on considère un endomorphisme symétrique f d'un espace vectoriel euclidien de dimension $n+1$. D'après le lemme 5, χ_f est scindé sur $\mathbb{R}$: soit x un vecteur propre pour f et $H = \big(\mathrm{Vect}\, x\big)^{\perp}$. Puisque $\mathrm{Vect}\, x$ est stable par f, il en est de même pour H (proposition 3.21). L'endomorphisme f_H de H induit par f est symétrique (car f l'est), et $\dim H = n-1$ (car $x \neq 0$). Par hypothèse de récurrence, il existe donc une base orthonormée b_1 de H dans laquelle $[f_H]_{b_1}$ est diagonale. Si $b = \left\{\frac{x}{||x||}\right\} \cup b_1$, b est une base orthonormée de E dans laquelle la matrice de f est diagonale.

□

b) Application aux formes quadratiques

Soient E un espace vectoriel euclidien et f un endomorphisme symétrique de E. Alors $\varphi_f : E \times E \to \mathbb{R}$ définie par $\varphi_f(x,y) = \langle f(x)\,|y\,\rangle$ est une forme bilinéaire symétrique sur E :

- ∗ bilinéaire car f est linéaire et $\langle \cdot\,|\cdot\,\rangle$ est bilinéaire ;
- ∗ symétrique car f est symétrique $\big(\varphi_f(y,x) = \langle f(y)\,|x\,\rangle = \langle y\,|f(x)\,\rangle = \varphi_f(x,y)\big)$.

Notons q_f la forme quadratique associée à φ_f : pour tout $x \in E$, $q_f(x) = \langle f(x)\,|x\,\rangle$.

Lemme 6 *Si b est une base orthonormée de E, alors les matrices de f et de q_f dans b coïncident.*

Démonstration. Soit $b = (e_1, \dots, e_n)$ une base orthonormée de E.

- Le coefficient $a_{i,j}$ de $[f]_b$ est la $i^{\text{ème}}$ coordonnée de $f(e_j)$ dans b : $a_{i,j} = \langle f(e_j)\,|e_i\,\rangle$ d'après la proposition 3.11.
- Le coefficient $b_{i,j}$ de $[q_f]_b$ vaut par définition $\varphi_f(e_i,e_j) = \varphi_f(e_i,e_j) = \varphi_f(e_j,e_i) = \langle f(e_j)\,|e_i\,\rangle = a_{i,j}$.

□

Proposition 3.22 *L'application $\Phi : \mathscr{S}(E) \to \mathscr{Q}(E)$, qui, à f, associe q_f, est un isomorphisme d'espaces vectoriels.*

Démonstration. • Φ est linéaire :

$$q_{f+\lambda g}(x) = \langle (f+\lambda g)(x)\,|x\,\rangle = \langle f(x) + \lambda g(x)\,|x\,\rangle = \langle f(x)\,|x\,\rangle + \lambda\,\langle g(x)\,|x\,\rangle = q_f(x) + \lambda q_g(x).$$

• Soit q une forme quadratique sur E et A sa matrice dans une base orthonormée b de E : A est une matrice symétrique (corollaire 3.1). Si f est l'endomorphisme de E admettant A pour matrice dans la base b, alors :

- → f est symétrique car A l'est (proposition 3.20) ;
- → d'après le lemme 6, on a $[q_f]_b = [f]_b$ donc q_f et q ont même matrice dans la base b : $q = q_f$.

Ceci montre que Φ est surjectif.

• Si, pour $f \in \mathscr{S}(E)$, $q_f = 0$, alors sa forme polaire φ_f, qui, à (x,y) associe $\langle f(x)\,|y\,\rangle$, est nulle. Ceci signifie que pour tout $x \in E$, $f(x)$ est orthogonal à tout élément de E (pour le produit scalaire !), donc que $f(x) = 0$. Ceci montre que Φ est injectif.

□

Théorème 3.13 *Soient E un espace vectoriel euclidien et q une forme quadratique sur E.*
Il existe une base orthonormée de E qui est q-orthogonale.

Démonstration. Soit f l'endomorphisme symétrique de E tel que $q = q_f$. D'après le théorème 3.12, il existe une base orthonormée $b = (e_1, \dots, e_n)$ de E formée de vecteurs propres de f. Alors, pour tout $(i,j) \in [\![1,n]\!]^2$, on a

$$\varphi_f(e_i,e_j) = \langle f(e_i)\,|e_j\,\rangle = \langle \lambda_i e_i\,|e_j\,\rangle = \lambda_i\,\langle e_i\,|e_j\,\rangle = \begin{cases} \lambda_i & \text{si } i = j\ ; \\ 0 & \text{si } i \neq j. \end{cases}$$

Ainsi, puisque φ_f est la forme polaire de $q = q_f$, $(e_1, \dots, e_n)$ est une base q-orthogonale.

□

4

Coniques du plan affine euclidien

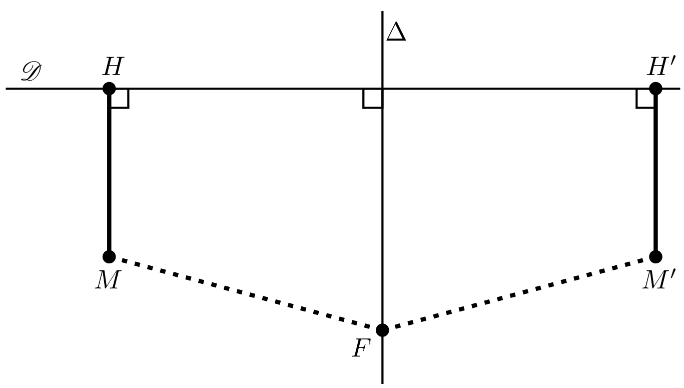
Dans tout ce chapitre, $\mathscr{E}$ est un espace affine euclidien de dimension 2 et E désigne sa direction.

I _ Point de vue géométrique

a) Définition par foyer et directrice

Définition *Soient $\mathscr{D}$ une droite de $\mathscr{E}$, F un point de $\mathscr{E}$ n'appartenant pas à $\mathscr{D}$ et e un nombre réel strictement positif. On appelle* conique *de* directrice *$\mathscr{D}$, de* foyer *F et d'*excentricité *e l'ensemble des points M du plan vérifiant* $\dfrac{MF}{d(M,\mathscr{D})} = e$.

Proposition 4.1 *Soit Γ une conique de foyer F et de directrice $\mathscr{D}$. Notons Δ la droite perpendiculaire à $\mathscr{D}$ passant par F. Alors Δ est axe de symétrie de Γ.*



Démonstration. Notons σ_Δ la réflexion d'axe Δ. Soit M un point de $\mathscr{E}$, H son projeté orthogonal sur $\mathscr{D}$ et H' le projeté orthogonal de $M' = \sigma_\Delta(M)$ sur $\mathscr{D}$. On a

$$d(M,\mathscr{D}) = MH, \quad d(M',\mathscr{D}) = M'H', \quad \sigma_\Delta(F) = F \quad \text{et} \quad \sigma_\Delta(H) = H'.$$

Puisque σ_Δ est une isométrie, on obtient donc

$$\frac{MF}{d(M,\mathscr{D})} = \frac{M'F}{d(M',\mathscr{D})}$$

et donc $M \in \Gamma$ si, et seulement si, $M' \in \Gamma$.

$\square$

Définition *La droite Δ est appelée* axe focal *de la conique Γ.*

Une question légitime est de savoir si Γ est non vide. Pour y répondre, on va démontrer que $\Gamma \cap \Delta$ est non vide :

Proposition 4.2 *Soit Γ une conique d'excentricité e et d'axe focal Δ.*

- *Si $e=1$, alors $\Gamma\cap\Delta$ est réduit à un point.*
- *Si $e\neq 1$, $\Gamma\cap\Delta$ est constitué de deux points distincts S et S'.*

Démonstration. Notons respectivement $\mathscr{D}$ et F la directrice et le foyer de Γ et K le projeté orthogonal de F sur $\mathscr{D}$. On a, si M est un point de $\mathscr{E}$ et H son projeté orthogonal sur $\mathscr{D}$:

$$\begin{aligned} M\in\Gamma\cap\Delta &\iff M\in\Delta \quad \text{et} \quad MF^2=e^2\,MH^2 \\ &\iff M\in\Delta \quad \text{et} \quad \left\langle \overrightarrow{MF}-e\overrightarrow{MH} \,\middle|\, \overrightarrow{MF}+e\overrightarrow{MH}\right\rangle = 0. \end{aligned}$$

Or si $M\in\Delta$, M, F et H sont alignés, $H=K$ et les vecteurs $\overrightarrow{MF}-e\overrightarrow{MK}$ et $\overrightarrow{MF}+e\overrightarrow{MK}$ sont colinéaires ; ils sont alors orthogonaux si, et seulement si, l'un des deux est nul. On déduit :

$$M\in\Gamma\cap\Delta \quad\iff\quad \overrightarrow{MF}=e\overrightarrow{MK} \quad \text{ou} \quad \overrightarrow{MF}=-e\overrightarrow{MK}$$

<u>1^er cas : $e=1$</u>. Dans ce cas, $\overrightarrow{MF}=e\overrightarrow{MK}$ est exclu car $F\notin\mathscr{D}$, donc $M\in\Gamma\cap\Delta$ si, et seulement si, $\overrightarrow{MF}=-\overrightarrow{MK}$, c'est-à-dire M est le milieu de $[F,K]$.

<u>2^ème cas : $e\neq 1$</u>. Dans ce cas, $M\in\Gamma\cap\Delta$ si, et seulement, $M=S$ ou S', définis comme étant les barycentres suivants :

$$S=\frac{1}{1-e}(F-eK) \qquad \text{et} \qquad S'=\frac{1}{1+e}(F+eK).$$

Notons que $S\neq S'$ car $e\neq 0$.

□

Remarque : le point S' appartient au segment $[F,K]$.

Dans le cas où $e=1$, l'unique point de $\Delta\cap\Gamma$ est appelé *sommet* de Γ.
Dans le cas où $e\neq 1$, les deux points S et S' de $\Delta\cap\Gamma$ sont également appelés *sommets* de Γ. On notera Δ' la médiatrice de $[S,S']$: c'est une droite perpendiculaire à $(SS')=(FK)=\Delta$ donc parallèle à la directrice $\mathscr{D}$. De plus, le milieu C de $[S,S']$ est le point d'intersection de Δ et Δ'.

Proposition 4.3 *Δ' est axe de symétrie de Γ.*

Corollaire 4.1 *Le point C est centre de symétrie de Γ.*

Remarque : Lorsque $e\neq 1$, on parle également de coniques à centre.

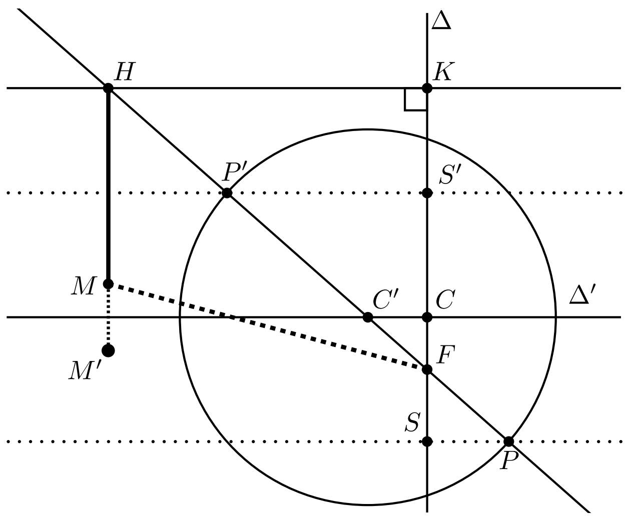
Démonstration. La symétrie de centre C coïncide avec la composée $\sigma_\Delta\circ\sigma_{\Delta'}$.

□

Démonstration de la proposition 4.3. Soit M un point de Γ, H son projeté orthogonal sur la directrice $\mathscr{D}$ et M' son image par la réflexion d'axe Δ'. On note p la projection sur la droite (FH) parallèlement à Δ'. Si $P=p(S)$ et $P'=p(S')$, alors on a

$$P=\frac{1}{1-e}(F-eH) \qquad \text{et} \qquad P'=\frac{1}{1+e}(F+eH)$$

car p est affine, $p(F)=F$ et $p(K)=H$.



Soit $\mathcal{C} = \{N \in \mathscr{E} \;/\; \frac{NF}{NH} = e\}$. Puisque, pour $N \in \mathscr{E}$, on a

$$\begin{aligned} NF = eNH &\iff NF^2 - e^2 NH^2 = 0 \\ &\iff \left\langle \overrightarrow{NF} - e\overrightarrow{NH} \,\middle|\, \overrightarrow{NF} + e\overrightarrow{NH} \right\rangle = 0 \\ &\iff \left\langle \frac{1}{1-e}\overrightarrow{NP} \,\middle|\, \frac{1}{1+e}\overrightarrow{NP'} \right\rangle = 0 \\ &\iff \left\langle \overrightarrow{NP} \,\middle|\, \overrightarrow{NP'} \right\rangle = 0, \end{aligned}$$

$\mathcal{C}$ est le cercle de diamètre $[P, P']$. Si C' est son centre, on a

$$C' = \frac{1}{2}(P + P') = \frac{1}{2}\big(p(S) + p(S')\big) = p\Big(\frac{1}{2}(S + S')\Big) = p(C)$$

donc C' est sur la parallèle à Δ' passant par C, c'est-à-dire sur Δ'. En particulier, Δ' est axe de symétrie de $\mathcal{C}$.

Maintenant, puisque $M \in \Gamma$, on a $MF = e\, d(M, \mathscr{D}) = e\, MH$ donc M est sur $\mathcal{C}$. Par conséquent, $M' = \sigma_{\Delta'}(M)$ est également sur $\mathcal{C}$ donc $M'F = e\, M'H = e\, d(M', \mathscr{D})$ car M et M' ont même projeté orthogonal sur $\mathscr{D}$. Ceci montre que M' est sur Γ.

□

Proposition 4.4 *Soit Γ une conique d'excentricité $e \neq 1$, de foyer F, de directrice $\mathscr{D}$ et de centre C. Soit F' et $\mathscr{D}'$ les symétriques de F et $\mathscr{D}$ par rapport à C. Alors, la conique de directrice $\mathscr{D}'$, de foyer F' et d'excentricité e coïncide avec Γ :*

$$\Gamma = \left\{ M \in \mathscr{E} \;/\; \frac{MF}{d(M, \mathscr{D})} = e \right\} = \left\{ M \in \mathscr{E} \;/\; \frac{MF'}{d(M, \mathscr{D}')} = e \right\}.$$

Démonstration. Si Γ' est la conique de directrice $\mathscr{D}'$, de foyer F' et d'excentricité e, alors, pour $M \in \mathscr{E}$ et $M' = s_C(M)$, on a

$$\begin{aligned} M \in \Gamma &\iff \frac{MF}{d(M, \mathscr{D})} = e \\ &\iff \frac{M'F'}{d(M', \mathscr{D}')} = e \quad \text{car } s_C \text{ est une isométrie} \\ &\iff M' \in \Gamma', \end{aligned}$$

donc $s_C(\Gamma) = \Gamma'$, ce qui permet de conclure que $\Gamma = \Gamma'$ d'après le corollaire 4.1.

□

b) Recherche d'une équation

Soit Γ une conique de directrice $\mathscr{D}$, de foyer F et d'excentricité e. On note comme précédemment K le projeté orthogonal de F sur $\mathscr{D}$.

Premier cas : $e = 1$. Soit $\vec{\imath} = \frac{\overrightarrow{KF}}{KF}$ et $\vec{\jmath} \in \overrightarrow{\mathscr{D}}$ de norme 1 : $(\vec{\imath}, \vec{\jmath})$ est une base orthonormée de E. Si S est le sommet de Γ et $p = FK$, alors, puisque S est le milieu de $[F, K]$ (voir la démonstration de la proposition 4.2), les coordonnées de F et K dans le repère $(S; \vec{\imath}, \vec{\jmath})$ sont respectivement $(\frac{p}{2}, 0)$ et $(-\frac{p}{2}, 0)$. De là, si un point M a pour coordonnées (x, y) dans ce repère, on a $MF^2 = (x - \frac{p}{2})^2 + y^2$ et $d(M, \mathscr{D})^2 = \left\langle \overrightarrow{MK} \,|\vec{\imath}\right\rangle^2 = (x + \frac{p}{2})^2$ donc

$$\begin{aligned} M \in \Gamma &\iff MF^2 = d(M, \mathscr{D})^2 \\ &\iff \left(x - \frac{p}{2}\right)^2 + y^2 = \left(x + \frac{p}{2}\right)^2 \\ &\iff y^2 = 2px. \end{aligned}$$

Définition *Dans ce cas, Γ est appelée* parabole *de sommet S, de paramètre p et d'équation réduite $y^2 = 2px$.*

Second cas : $e \neq 1$. Rappelons que le centre C de Γ est le milieu de $[S, S']$ où

$$S = \frac{1}{1-e}(F - eK) \quad \text{et} \quad S' = \frac{1}{1+e}(F + eK).$$

Soit $\vec{\imath} = \frac{\overrightarrow{CF}}{CF}$ et $\vec{\jmath} \in \overrightarrow{\mathscr{D}}$ de norme 1. On note $a = CS = CS'$ et $c = CF$. Puisque

$$\overrightarrow{CF} - e\,\overrightarrow{CK} = (1-e)\overrightarrow{CS}, \quad \overrightarrow{CF} + e\,\overrightarrow{CK} = (1+e)\overrightarrow{CS'} \quad \text{et} \quad \overrightarrow{CS} = -\overrightarrow{CS'},$$

on a $\overrightarrow{CF} = e\,\overrightarrow{CS'}$ et $e\,\overrightarrow{CK} = \overrightarrow{CS'}$. Par conséquent, $c = ea$ et on obtient

$$\overrightarrow{CK} = \frac{1}{e}\,\overrightarrow{CS'} = \frac{1}{e^2}\,\overrightarrow{CF} = \frac{c}{e^2}\,\vec{\imath} = \frac{a}{e}\,\vec{\imath}.$$

En particulier, $CK = \frac{c}{e^2} = \frac{a}{e}$ et $\mathscr{D}$ a pour équation dans le repère $(C; \vec{\imath}, \vec{\jmath})$ $x = \frac{a}{e} = \frac{c}{e^2}$. Nous pouvons à présent donner une équation de Γ dans le repère $(C; \vec{\imath}, \vec{\jmath})$. Pour $M \in \mathscr{E}$ de coordonnées (x, y), on a $MF^2 = (x - c)^2 + y^2$ et $d(M, \mathscr{D})^2 = \left\langle \overrightarrow{MK} \,|\vec{\imath}\right\rangle^2 = \left(x - \frac{a}{e}\right)^2$ donc

$$\begin{aligned} M \in \Gamma &\iff (x-c)^2 + y^2 = e^2\left(x - \frac{a}{e}\right)^2 = (ex - a)^2 \\ &\iff (1 - e^2)x^2 + y^2 = a^2(1 - e^2) \quad (\text{car} \ \ c = ea) \\ &\iff \frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1. \end{aligned}$$

Si $e < 1$, on pose $b^2 = a^2(1 - e^2)$ et si $e > 1$, on pose $b^2 = a^2(e^2 - 1)$. Dans le premier cas ($e < 1$), l'équation de Γ est

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

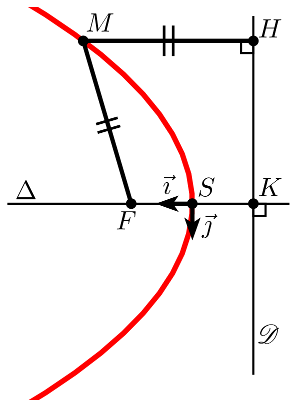
et dans le second cas ($e > 1$), l'équation est

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Définition *Une* ellipse *est une conique d'excentricité $e < 1$. Une* hyperbole *est une conique d'excentricité $e > 1$. Dans les deux cas, le réel a donné par les équations ci-dessus est appelé* demi-grand axe *de la conique.*

On peut récapituler les résultats obtenus précédemment.

Parabole ($e = 1$)



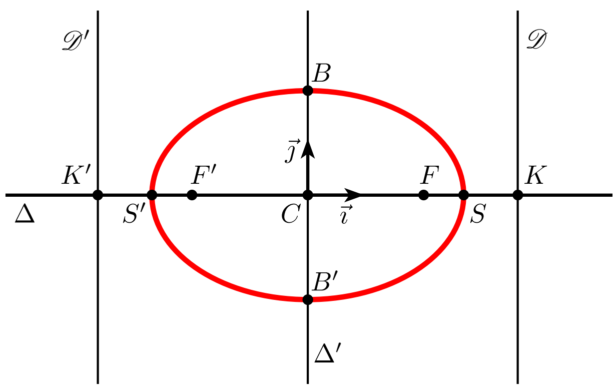
Repère privilégié : $(S; \vec{\imath}, \vec{\jmath})$

Équation de Γ : $y^2 = 2px$

Équation de la directrice $\mathscr{D}$: $x = -\frac{p}{2}$

$p = FK$

Ellipse ($e < 1$)



Repère privilégié : $(C; \vec{\imath}, \vec{\jmath})$

Équation de Γ : $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

$a = CS = CS'$

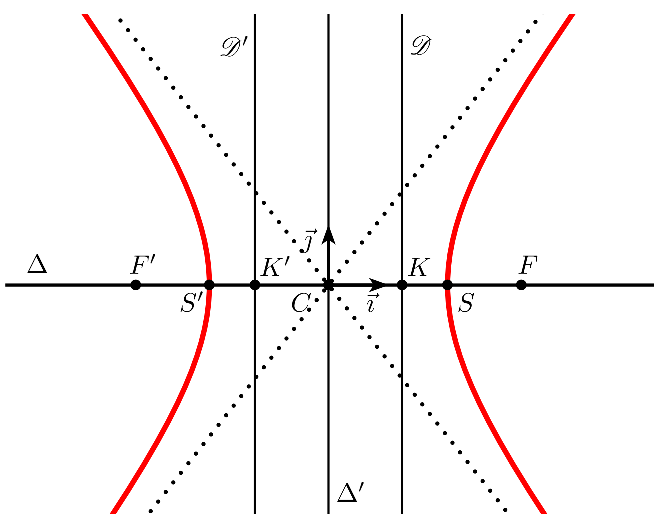
$c = CF = CF' = ea$

$b = a\sqrt{1-e^2} = \sqrt{a^2 - c^2} = CB = CB'$

Équation de la directrice $\mathscr{D}$: $x = \frac{a}{e}$

Équation de la directrice $\mathscr{D}'$: $x = -\frac{a}{e}$

Hyperbole ($e > 1$)



Repère privilégié : $(C; \vec{\imath}, \vec{\jmath})$

Équation de Γ : $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

$a = CS = CS'$

$c = CF = CF' = ea$

$b = a\sqrt{e^2 - 1} = \sqrt{c^2 - a^2}$

Équation de la directrice $\mathscr{D}$: $x = \frac{a}{e}$

Équation de la directrice $\mathscr{D}'$: $x = -\frac{a}{e}$

Équation des asymptotes : $\frac{x}{a} \pm \frac{y}{b} = 0$

Complément : asymptote d'une hyperbole. Soient D_1 et D_2 les deux droites d'équations respectives $\frac{x}{a} - \frac{y}{b} = 0$ et $\frac{x}{a} + \frac{y}{b} = 0$ dans le repère $(C; \vec{\imath}, \vec{\jmath})$. Pour $M \in \mathscr{E}$ de coordonnées (x, y), on considère les points $N_1 = N_1(x)$ et $N_2 = N_2(x)$ de coordonnées respectives $(x, \frac{b}{a}x)$ et $(x, -\frac{b}{a}x)$: $N_1 \in D_1$ et $N_2 \in D_2$. On a

$$MN_1^2 = \left(y - \frac{b}{a}x\right)^2 \quad \text{et} \quad MN_2^2 = \left(y + \frac{b}{a}x\right)^2$$

donc

$$MN_1^2 \cdot MN_2^2 = \left(y - \frac{b}{a}x\right)^2 \cdot \left(y + \frac{b}{a}x\right)^2 = \left(y^2 - \frac{b^2}{a^2}x^2\right)^2.$$

Par conséquent, si M est sur l'hyperbole d'équation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, on obtient

$$MN_1^2 \cdot MN_2^2 = \left(y^2 - b^2\left(1 + \frac{y^2}{b^2}\right)\right)^2 = b^4.$$

Ainsi, lorsque $CM^2 = x^2 + y^2 \underset{M \in \Gamma}{\longrightarrow} +\infty$ et que x et y gardent un signe constant, MN_1 ou MN_2 tend vers $+\infty$ et donc MN_2 ou MN_1 tend vers 0. Ceci montre que D_1 et D_2 sont deux asymptotes de Γ.

c) Définition bifocale des coniques à centre

Théorème 4.1 *Soit Γ une conique à centre, de foyers F et F', d'excentricité e et de grand axe $2a$.*

*1) Si Γ est une ellipse (*ie $e < 1$*), alors $\Gamma = \{M \in \mathscr{E} \,/\, MF + MF' = 2a\}$.*

*2) Si Γ est une hyperbole (*ie $e > 1$*), alors $\Gamma = \{M \in \mathscr{E} \,/\, |MF - MF'| = 2a\}$.*

Démonstration. • Notons $\mathscr{D}$ et $\mathscr{D}'$ les directrices associées respectivement à F et F'. Soit M un point de Γ et H, H' ses projetés orthogonaux respectifs sur $\mathscr{D}$ et $\mathscr{D}'$.

Si $e < 1$, alors $M \in [H, H']$ donc $MH + MH' = HH' = 2\,\frac{a}{e}$. On obtient donc

$$MF + MF' = e\big(d(M, \mathscr{D}) + d(M, \mathscr{D}')\big) = e(MH + MH') = 2a.$$

Si $e > 1$, alors $M \notin [H, H']$ donc $|MH - MH'| = HH' = 2\,\frac{a}{e}$. On obtient donc

$$|MF - MF'| = e\big|d(M, \mathscr{D}) - d(M, \mathscr{D}')\big| = e|MH - MH'| = 2a.$$

• Réciproquement, on note comme précédemment C le milieu de $[F, F']$ et $c = CF$. On travaille dans un repère orthonormé $(C; \vec{\imath}, \vec{\jmath})$ où $\vec{\imath} = \frac{\overrightarrow{CF}}{c}$. Ainsi, F et F' ont respectivement pour coordonnées $(c, 0)$ et $(-c, 0)$, et si M a pour coordonnées (x, y), alors $MF^2 = (x - c)^2 + y^2$ et $MF'^2 = (x + c)^2 + y^2$. De là, on obtient :

$$\begin{aligned}
\left[(MF + MF')^2 - 4a^2\right]\left[(MF - MF')^2 - 4a^2\right] &= (MF + MF' - 2a)(MF + MF' + 2a)\times \\
&\qquad\qquad (MF - MF' - 2a)(MF - MF' + 2a) \\
&= \left[(MF - 2a)^2 - MF'^2\right]\left[(MF + 2a)^2 - MF'^2\right] \\
&= \left[MF^2 - MF'^2 + 4a^2 - 4aMF\right]\left[MF^2 - MF'^2 + 4a^2 + 4aMF\right] \\
&= \left[-4cx + 4a^2 - 4aMF\right]\left[-4cx + 4a^2 + 4aMF\right] \\
&= 16\left[-cx + a^2 - aMF\right]\left[-cx + a^2 + aMF\right] \\
&= 16\left[(-cx + a^2)^2 - a^2MF^2\right] \\
&= 16\left[c^2x^2 - 2ca^2x + a^4 - a^2\left((x - c)^2 + y^2\right)\right] \\
&= 16\left((c^2 - a^2)x^2 - a^2y^2 + a^4 - a^2c^2\right) \\
&= 16a^2(c^2 - a^2)\left(\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} - 1\right).
\end{aligned}$$

On a donc $MF + MF' = 2a$ ou $|MF - MF'| = 2a$ si, et seulement si, $\dfrac{x^2}{a^2} + \dfrac{y^2}{a^2 - c^2} = 1$.

Pour conclure, rappelons les inégalités triangulaires

$$\left|MF - MF'\right| \leq FF' = 2c \leq MF + MF'.$$

Ainsi, si $MF + MF' = 2a$, alors $a \geq c$ et $\frac{x^2}{a^2} + \frac{y^2}{a^2-c^2} = 1$ est bien l'équation de Γ (ellipse dans ce cas). Si $|MF - MF'| = 2a$, alors $a \leq c$ et Γ est l'hyperbole d'équation $\frac{x^2}{a^2} - \frac{y^2}{c^2-a^2} = 1$. Dans les deux cas, M appartient à Γ.

□

Remarque : si $a = c$, nous sommes dans le cas d'égalité dans les inégalités triangulaires. En conséquence, si $FF' = 2a$, on a $MF + MF' = 2a$ si, et seulement si, $M \in [F, F']$, et $|MF - MF'] = 2a$ si, et seulement si, $M \in (FF') \setminus [F, F']$.

II _ Équation d'une conique

Rappelons que l'équation d'une droite dans un repère est de la forme $ax + by + c = 0$: c'est un polynôme de degré 1 en les deux variables x et y. Cette équation peut s'écrire sans se référer à la base $(\vec{\imath}, \vec{\jmath})$ sous la forme $l\left(\overrightarrow{\Omega M}\right) + c = 0$ où l est une forme linéaire sur E. Nous allons ici considérer des équations de degré 2 :

$$ax^2 + by^2 + cxy + dx + ey + f = 0 \qquad \text{avec } (a, b, c, d, e, f) \in \mathbb{R}^6.$$

Une telle équation peut se mettre sous la forme

$$q\left(\overrightarrow{\Omega M}\right) + l\left(\overrightarrow{\Omega M}\right) + f = 0 \qquad \text{où} \begin{cases} q : E \to \mathbb{R} & \text{est une forme quadratique,} \\ l : E \to \mathbb{R} & \text{est une forme linéaire.} \end{cases}$$

Nous avons vu précédemment qu'une conique admet une équation de cette forme. Nous allons voir ici que, réciproquement, l'ensemble des points satisfaisant une telle équation est une conique.
Commençons par examiner en quel(s) terme(s) cette équation dépend de l'origine. Soit donc Ω' un point de $\mathscr{E}$. On a, si φ est la forme polaire de q,

$$\begin{aligned} q\left(\overrightarrow{\Omega M}\right) + l\left(\overrightarrow{\Omega M}\right) + f &= q\left(\overrightarrow{\Omega\Omega'} + \overrightarrow{\Omega' M}\right) + l\left(\overrightarrow{\Omega\Omega'} + \overrightarrow{\Omega' M}\right) + f \\ &= q\left(\overrightarrow{\Omega\Omega'}\right) + q\left(\overrightarrow{\Omega' M}\right) + 2\varphi\left(\overrightarrow{\Omega\Omega'}, \overrightarrow{\Omega' M}\right) + l\left(\overrightarrow{\Omega\Omega'}\right) + l\left(\overrightarrow{\Omega' M}\right) + f \\ &= q\left(\overrightarrow{\Omega' M}\right) + l'\left(\overrightarrow{\Omega' M}\right) + f' \end{aligned}$$

où $l' : E \to \mathbb{R}$ est donnée par $l'(\vec{u}) = l(\vec{u}) + 2\varphi\left(\overrightarrow{\Omega\Omega'}, \vec{u}\right)$ et $f' = q\left(\overrightarrow{\Omega\Omega'}\right) + l\left(\overrightarrow{\Omega\Omega'}\right) + f$. Ainsi, l'équation est de la même forme, avec la même forme quadratique, si l'on change d'origine. D'où la définition suivante.

Définition *On appelle* équation de conique *dans $\mathscr{E}$ toute équation de la forme $f(M) = 0$ où*

$$f : \mathscr{E} \to \mathbb{R} \quad \textit{est donnée par} \quad f(M) = q\left(\overrightarrow{\Omega M}\right) + l\left(\overrightarrow{\Omega M}\right) + k$$

avec q forme quadratique non nulle sur E, l forme linéaire sur E, $\Omega \in \mathscr{E}$ et $k \in \mathbb{R}$. La forme quadratique q sera appelée forme quadratique de l'équation.

On a vu au chapitre précédent (théorème 3.13) qu'il existe une base orthonormée $(\overrightarrow{e_1}, \overrightarrow{e_2})$ de E qui est q-orthogonale. Il existe donc $(\lambda, \mu) \in \mathbb{R}^2$ tel que si un élément $\vec{u}$ de E s'écrit $\vec{u} = x\overrightarrow{e_1} + y\overrightarrow{e_2}$, alors

$$q(\vec{u}) = \lambda x^2 + \mu y^2.$$

Remarquons que, quitte à changer l'équation $f(M) = 0$ en $-f(M) = 0$ et quitte à échanger $\overrightarrow{e_1}$ et $\overrightarrow{e_2}$, on peut supposer que $\lambda > 0$ (le but est d'étudier les coniques, pas leurs équations).

Étude dans le cas où q est non dégénérée. Puisque $\lambda > 0$, la signature de q est dans ce cas égale à $(2,0)$ ou $(1,1)$. Notons φ la forme polaire de q. L'application $\varphi_g : E \to E^*$, $x \mapsto \varphi(x,\cdot)$ est alors un isomorphisme (corollaire 3.2). Si $\overrightarrow{u_0} = \varphi_g^{-1}\left(\frac{1}{2}l\right)$, alors $l(\vec{u}) = 2\varphi(\overrightarrow{u_0}, \vec{u})$ pour tout $\vec{u} \in E$. De là :

$$\begin{aligned} q\left(\overrightarrow{\Omega M}\right) + l\left(\overrightarrow{\Omega M}\right) &= \varphi\left(\overrightarrow{\Omega M}, \overrightarrow{\Omega M}\right) + 2\varphi\left(\overrightarrow{u_0}, \overrightarrow{\Omega M}\right) \\ &= \varphi\left(\overrightarrow{\Omega M}, \overrightarrow{\Omega M}\right) + 2\varphi\left(\overrightarrow{u_0}, \overrightarrow{\Omega M}\right) + \varphi\left(\overrightarrow{u_0}, \overrightarrow{u_0}\right) - \varphi\left(\overrightarrow{u_0}, \overrightarrow{u_0}\right) \\ &= \varphi\left(\overrightarrow{\Omega M} + \overrightarrow{u_0}, \overrightarrow{\Omega M} + \overrightarrow{u_0}\right) - \varphi\left(\overrightarrow{u_0}, \overrightarrow{u_0}\right) \\ &= q\left(\overrightarrow{\Omega M} + \overrightarrow{u_0}\right) - q\left(\overrightarrow{u_0}\right) \end{aligned}$$

donc $f(M) = q\left(\overrightarrow{\Omega M} + \overrightarrow{u_0}\right) + k'$ si $k' = k - q\left(\overrightarrow{u_0}\right)$. Par conséquent, si $O = \Omega - \overrightarrow{u_0}$, on obtient

$$f(M) = q\left(\overrightarrow{OM}\right) + k'.$$

Puisque $q\left(-\overrightarrow{OM}\right) = q\left(\overrightarrow{OM}\right)$, le point O est centre de symétrie de la conique d'équation $f(M) = 0$. De plus, l'équation s'écrit dans le repère $(O; \overrightarrow{e_1}, \overrightarrow{e_2})$ $f(M) = \lambda x^2 + \mu y^2 + k'$ (avec $M = O + x\overrightarrow{e_1} + y\overrightarrow{e_2}$).

On peut résumer cette première étude comme suit (on a supposé $\lambda > 0$).

- Si q est de signature $(2,0)$ (*ie* $\mu > 0$), alors la conique
  - ∗ est vide si $k' > 0$;
  - ∗ réduite au point O si $k' = 0$;
  - ∗ admet pour équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ dans le repère orthonormé $(O; \overrightarrow{e_1}, \overrightarrow{e_2})$ (où $a = \sqrt{\frac{-k'}{\lambda}}$ et $b = \sqrt{\frac{-k'}{\mu}}$). Dans ce cas, la conique est une ellipse.
- Si q est de signature $(1,1)$ (*ie* $\mu < 0$), alors la conique
  - ∗ coïncide avec l'union de deux droites sécantes si $k' = 0$ ($\lambda x^2 = -\mu y^2$ avec $\mu < 0$) ;
  - ∗ admet pour équation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \pm 1$ dans le repère orthonormé $(O; \overrightarrow{e_1}, \overrightarrow{e_2})$ (où $a = \sqrt{\frac{|k'|}{\lambda}}$ et $b = \sqrt{\frac{|k'|}{-\mu}}$). Dans ce cas, la conique est une hyperbole.

Étude dans le cas où q est dégénérée. Puisque l'on a supposé $\lambda > 0$, on a $\mu = 0$ (q est de signature $(1,0)$). Dans le repère $(\Omega; \overrightarrow{e_1}, \overrightarrow{e_2})$, on a ($M = \Omega + x\overrightarrow{e_1} + y\overrightarrow{e_2}$)

$$f(M) = \lambda x^2 + ax + by + k \quad \text{avec} \quad (a,b) \in \mathbb{R}^2.$$

Si $b = 0$, alors $f(M) = \lambda(x + \frac{a}{2\lambda})^2 + k - \frac{a^2}{4\lambda}$ et donc, si $\alpha = \frac{a^2}{4\lambda} - k$, la conique

- ∗ est vide si $\alpha < 0$;
- ∗ réduite à une droite si $\alpha = 0$ ($x + \frac{a}{2\lambda} = 0$) ;
- ∗ coïncide avec l'union de deux droites si $\alpha > 0$ $\left(x + \frac{a}{2\lambda} = \pm\sqrt{\frac{\alpha}{\lambda}}\right)$.

Si $b \neq 0$, on a, toujours avec $\alpha = \frac{a^2}{4\lambda} - k$:

$$f(M) = \lambda\left(x + \frac{a}{2\lambda}\right)^2 + by - \alpha.$$

Si $S = \Omega - \frac{a}{2\lambda}\overrightarrow{e_1} + \frac{\alpha}{b}\overrightarrow{e_2}$, alors l'équation dans le repère $(S; \overrightarrow{e_1}, \overrightarrow{e_2})$ est $x^2 = 2py$ où $p = -\frac{b}{2\lambda}$; la conique est dans ce cas une parabole.

Index