Algèbre 2,
Cours de deuxième année de l'Université de Bordeaux

Jean-Jacques Ruch [1]

1. Institut de Mathématiques Bordeaux, UMR 5251 du CNRS, Université de Bordeaux, 351 cours de la Libération, F33405 Talence Cedex, France.

Table des Matières

Chapitre 1

DÉTERMINANT

Montrer qu'une application linéaire est inversible n'est à priori pas une chose évidente. Le déterminant permettra, dans certains cas, de montrer si c'est le cas ou non. Il permettra aussi, toujours dans certains cas, de résoudre des systèmes ou bien d'obtenir l'inverse d'une matrice. Enfin il servira, comme on le verra dans la prochaine leçon, à la diagonalisation et la trigonalisation des endomorphismes d'un espace vectoriel.

Dans tout ce chapitre K désigne un corps. Rappelons qu'un corps est un espace vectoriel sur lui-même de dimension 1.

I. Préliminaire sur les permutations (sans démonstrations)

Définition 1. *Une* permutation σ *de* $\{1, 2, \ldots, n\}$ *est une bijection de* $\{1, 2, \ldots, n\}$ *dans lui-même.*
Une transposition τ *est une permutation qui laisse invariant tous les éléments sauf deux qu'elle échange : il existe deux éléments distincts* i *et* j *tels que :*

$$\tau(i) = j \text{ et } \tau(j) = i \text{ et } \quad \forall k \neq i \text{ et } k \neq j,\ \tau(k) = k.$$

Proposition 2. *L'ensemble des permutations de* $\{1, 2, \ldots, n\}$*, muni de la composition des applications est un groupe appelé* groupe symétrique *et est noté* S_n*. L'élément neutre est l'identité, noté* Id.

Cela signifie simplement que l'on a les propriétés suivantes :
- $S_n \neq \emptyset$;
- $\exists e \in S_n \mid \forall \sigma \in S_n,\ \sigma \circ e = e \circ \sigma = \sigma$; ici $e = \mathrm{Id}$;
- $\forall \sigma_1 \in S_n,\ \exists \sigma_2 \in S_n \mid \sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1 = \mathrm{Id}$; on note alors σ_2 par σ_1^{-1}.

Remarque 3. *On peut remarquer que ce groupe symétrique n'est pas un groupe commutatif (si* $n > 2$*) car en général, pour* σ_1, σ_2 *dans* S_n*, on a* $\sigma_1 \circ \sigma_2 \neq \sigma_2 \circ \sigma_1$.
On note souvent $\sigma_1\sigma_2$ *au lieu de* $\sigma_1 \circ \sigma_2$.

Exemple 4. *Les éléments de* S_3 *sont*

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad (1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$(2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad (1\ 2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \text{ et } (1\ 3\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On remarque qu'il y a $6 = 3!$ *éléments. On peut montrer de façon générale que* S_n *a* $n!$ *éléments.*

Proposition 5. *Toute permutation de* S_n *peut s'écrire comme une composée de transpositions de* S_n.

Démonstration. Plutôt que de démontrer l'existence de la décomposition en générale, montrons le principe de la démonstration sur l'exemple

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 5 & 1 \end{pmatrix} = (1\ 2\ 4\ 5).$$

Le principe est de multiplier σ à gauche (composer à gauche) par des transpositions de façon à obtenir chaque fois un élément fixe de plus. Ici $\sigma(1) = 2$, on compose donc par $(1\ 2)$, et on obtient $(1\ 2)\sigma = (2\ 4\ 5)$. On continue (en oubliant 1) ; comme 2 donne 4, on compose par $(2\ 4)$. On trouve $(2\ 4)(1\ 2)\sigma = (4\ 5)$. Une transposition étant sa propre inverse on en déduit que $\sigma = (1\ 2)(2\ 4)(4\ 5)$. □

Remarque 6. *Cette décomposition n'est pas unique, mais la parité du nombre de transpositions qui interviennent dans une décomposition ne dépend pas de la décomposition.*

Exemple 7. *Dans S_3 on a*

$$(1\ 2\ 3) = (1\ 2)(2\ 3) = (1\ 2)(1\ 3)(1\ 3)(2\ 3);$$

les deux décompositions contiennent un nombre pair de transpositions.

Définition 8. *On appelle* signature *d'une permutation σ, le nombre $\varepsilon(\sigma) = (-1)^k$ où k est le nombre de transpositions qui interviennent dans une décomposition de σ en produit de transpositions.*

Remarque 9. *La signature d'une permutation est bien définie, car si σ se décompose de deux manières en produit de k et k' transpositions, alors $k - k'$ est un multiple de 2 et donc $(-1)^k = (-1)^{k'}$.*
En particulier on a pour une transposition τ, $\varepsilon(\tau) = -1$ et pour l'identité $\varepsilon(\mathrm{Id}) = 1$.

Proposition 10. *L'application $\varepsilon : S_n \to \{-1, 1\}$ est un morphisme de groupe. Cela signifie que*

$$\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2).$$

Remarque 11. *En particulier on a*

$$\varepsilon(\sigma)\varepsilon(\sigma^{-1}) = \varepsilon(\sigma\sigma^{-1}) = \varepsilon(\mathrm{Id}) = 1,$$

et donc

$$\varepsilon(\sigma^{-1}) = \frac{1}{\varepsilon(\sigma)} = \varepsilon(\sigma) \text{ car } \varepsilon(\sigma) = \pm 1.$$

Exemple 12. *Dans S_3 on a*

$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	*Id*	$\varepsilon(\sigma_1) = 1$
$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	$(1\ 2)$	$\varepsilon(\sigma_2) = -1$
$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	$(1\ 3)$	$\varepsilon(\sigma_3) = -1$
$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	$(2\ 3)$	$\varepsilon(\sigma_4) = -1$
$\sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$(1\ 3)(2\ 3)$	$\varepsilon(\sigma_5) = 1$
$\sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$(1\ 2)(2\ 3)$	$\varepsilon(\sigma_6) = 1$

II. Définition et propriété des déterminants

II.1. Déterminant d'une matrice.

Définition 13. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n à coefficients dans un corps K. On appelle* déterminant de la matrice M *et on note* $\det(M)$ *le scalaire*

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(j)}$$

Remarque 14. *On note en général le déterminant de la matrice M :*

$$\det(M) = \begin{vmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ldots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{vmatrix}.$$

Proposition 15. *On a aussi*

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j}.$$

Démonstration. La preuve repose essentiellement sur deux changements de variable : un premier dans le produit et un second dans la somme. Soit ρ et σ dans S_n, alors comme ρ est une bijection de $\{1,\ldots,n\}$, en posant $j=\rho(k)$ on a

$$\prod_{j=1}^{n} m_{j\sigma(j)} = \prod_{k=1}^{n} m_{\rho(k)\sigma(\rho(k))}.$$

En particulier avec $\rho = \sigma^{-1}$, on a $\sigma(\rho(k)) = k$ et donc

$$\prod_{j=1}^{n} m_{j\sigma(j)} = \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

En revenant à la définition du déterminant on obtient, par sommation, que

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

On se rappelle ensuite que $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ de sorte que

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma^{-1}) \prod_{k=1}^{n} m_{\sigma^{-1}(k)k}.$$

Comme $\sigma \to \sigma^{-1}$ est une bijection de S_n, en posant $\rho = \sigma^{-1}$ on obtient

$$\det(M) = \sum_{\rho\in S_n} \varepsilon(\rho) \prod_{k=1}^{n} m_{\rho(k)k},$$

et on renomme la variable de sommation σ, au lieu de ρ, pour conclure. $\square$

Corollaire 16. *Une matrice carrée et sa transposée ont des déterminants égaux.*

Démonstration. Soit $M = (m_{ij})$ une matrice carrée. On note ${}^tM = (m^*_{ij}) = (m_{ji})$ sa transposée. On a :

$$\det({}^tM) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m^*_{j\sigma(j)} = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j} = \det(M),$$

d'après la proposition précédente. □

Remarque 17. *Cela signifie que toutes les propriétés du déterminant qui seront établies par rapport aux colonnes d'une matrice seront aussi valables pour les lignes.*

Proposition 18.
a) *Si $M = (m)$ est une matrice d'ordre 1 alors $\det(M) = m$.*
b) *Si on note I_n la matrice identité d'ordre n alors $\det(\mathrm{I}_n) = 1$.*
c) *Pour tout $\lambda \in K$ et toute matrice carrée M d'ordre n, on a $\det(\lambda M) = \lambda^n \det(M)$.*

Démonstration. Le premier point est évident.
Pour le second, on remarque que si on note a_{ij} les coefficients de la matrice identité, alors $a_{j\sigma(j)} = 0$ si $\sigma(j) \neq j$, et donc $\det(\mathrm{I}_n) = \prod_{j=1}^{n} a_{jj} = 1$.
Enfin pour le troisième point si $M = (m_{ij})$ alors $\lambda M = (\lambda m_{ij})$ et donc

$$\det(\lambda M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} \lambda m_{j\sigma(j)} = \lambda^n \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(j)} = \lambda^n \det(A).$$

□

Exemple 19. *On a $S_2 = \{\mathrm{Id},\ (1\ 2)\}$ et donc :*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Exemple 20. *On a $S_3 = \{\mathrm{Id},\ (1\ 2),\ (1\ 3),\ (2\ 3),\ (1\ 2)(1\ 3),\ (1\ 2)(2\ 3)\}$ et donc*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33}.$$

II.2. Déterminant d'une famille de vecteurs.

Définition 21. *Soient $(x_1, \ldots, x_n)$ des vecteurs d'un K espace vectoriel E de dimension n. On appelle* déterminant *de $x_1, \ldots, x_n$ dans une base $\mathcal{B}$ et on note $\det_{\mathcal{B}}(x_1, \ldots, x_n)$ (on omet l'indice $\mathcal{B}$ lorsqu'il n'y a pas ambiguité) le déterminant de la matrice dont les colonnes sont constituées des vecteurs $x_1, \ldots, x_n$.*

Proposition 22. *Soient $(x_1, \ldots, x_n)$ des vecteurs d'un K espace vectoriel E de dimension n.*
a) *Soit $\rho \in S_n$, alors*

$$\det(x_{\rho(1)}, \ldots, x_{\rho(n)}) = \varepsilon(\rho)\det(x_1, \ldots, x_n).$$

En particulier si on échange deux colonnes d'une matrice, le déterminant est changé en son opposé.
b) *Si $x_i = x_j$ pour $i \neq j$ alors on a $\det(x_1, \ldots, x_n) = 0$.*
c) *Soient y_k un vecteur de E et λ et μ deux scalaires, alors :*

$$\det(x_1, \ldots, \lambda x_k + \mu y_k, \ldots, x_n) = \lambda\det(x_1 \ldots, x_k, \ldots, x_n) + \mu\det(x_1 \ldots, y_k, \ldots, x_n).$$

Le déterminant d'une matrice dépend donc linéairement de chacun de ses vecteurs colonnes. En particulier il est nul si une colonne est nulle.
d) *Le déterminant d'une matrice ne change pas lorsqu'on ajoute à l'un de ses vecteurs-colonnes une combinaison linéaire des autres vecteurs-colonnes ; il est nul si l'un des vecteurs-colonnes est combinaison des autres.*

DÉMONSTRATION. On note $M=(m_{ij})$ la matrice dont les colonnes sont les vecteurs $x_1,\dots,x_n$.
a) On a

$$\begin{aligned}
\det(x_{\rho(1)},\dots,x_{\rho(n)}) &= \sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{j=1}^n m_{j\sigma(\rho(j))}\\
&= \varepsilon(\rho)\sum_{\sigma\in S_n}\varepsilon(\rho)\varepsilon(\sigma)\prod_{j=1}^n m_{j\sigma\rho(j)} \quad \text{car} \quad \varepsilon(\rho)^2=1\\
&= \varepsilon(\rho)\sum_{\sigma\rho\in S_n}\varepsilon(\sigma\rho)\prod_{j=1}^n m_{j\sigma\rho(j)} \quad \text{car} \quad \varepsilon(\sigma\rho)=\varepsilon(\sigma)\varepsilon(\rho)\\
&= \varepsilon(\rho)\sum_{\tau\in S_n}\varepsilon(\tau)\prod_{j=1}^n m_{j\tau(j)} \quad \text{en posant } \tau=\sigma\rho\\
&= \varepsilon(\rho)\det(x_1,\dots,x_n)
\end{aligned}$$

En particulier si on note τ la transposition $(i\ j)$ alors $\varepsilon(\tau)=-1$ et par conséquent on a

$$\det(x_1,\dots,x_j,\dots,x_i,\dots,x_n)=\det(x_1,\dots,x_{\tau(i)},\dots,x_{\tau(j)},\dots,x_n)=-\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n).$$

b) On a d'après la propriété précédente

$$\begin{aligned}
\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n) &= -\det(x_1,\dots,x_j,\dots,x_i,\dots,x_n)\\
&= -\det(x_1,\dots,x_i,\dots,x_j,\dots,x_n) \quad \text{car } x_i=x_j
\end{aligned}$$

et donc ce déterminant est nul.
c) Soit

$$x_k=\begin{pmatrix}a_{1k}\\ \vdots\\ a_{nk}\end{pmatrix} \text{ et } y_k=\begin{pmatrix}b_{1k}\\ \vdots\\ b_{nk}\end{pmatrix},$$

alors en utilisant la deuxième formule du déterminant

$$\begin{aligned}
\det(x_1,\dots,\lambda x_k+\mu y_k,\dots,x_n) &= \sum_{\sigma\in S_n}\varepsilon(\sigma)(\lambda a_{\sigma(k)k}+\mu b_{\sigma(k)k})\prod_{j=1,j\neq k}^n m_{\sigma(j)j}\\
&= \lambda\sum_{\sigma\in S_n}\varepsilon(\sigma)a_{\sigma(k)k}\prod_{j=1,j\neq k}^n m_{\sigma(j)j}+\mu\sum_{\sigma\in S_n}\varepsilon(\sigma)b_{\sigma(k)k}\prod_{j=1,j\neq k}^n m_{\sigma(j)j}\\
&= \lambda\det(x_1\dots,x_k,\dots,x_n)+\mu\det(x_1\dots,y_k,\dots,x_n)
\end{aligned}$$

d) Par la linéarité du déterminant par rapport à chaque variable (propriété précédente) on a

$$\det(x_1,\dots,x_k+\sum_{l=1,l\neq k}^n\lambda_l x_l,\dots,x_n)=\det(x_1,\dots,x_k,\dots,x_n)+\sum_{l=1,l\neq k}^n\lambda_l\det(x_1,\dots,x_l,\dots,x_n).$$

Or tous les déterminants de la somme sont nuls d'après la propriété **b)**, car ils ont deux colonnes identiques. Ceci montre le résultat voulu. En particulier, si un vecteur est égal à une combinaison linéaire des autres vecteurs, on ne change pas le déterminant si on ajoute au vecteur l'opposé de la combinaison linéaire. Donc on est ramené à calculer le déterminant d'une matrice dont une colonne est nulle et par conséquent ce déterminant est nul. □

II.3. Propriétés.

Le prochain résultat ètablit le caractère multiplicatif du dèterminant.

Théorème 23.

Soient M et N deux matrices carrées d'ordre n, alors

$$\det(MN) = \det(M)\det(N).$$

DÉMONSTRATION. Notons par $M_1, \dots, M_n$ les colonnes de M de sorte que celles du produit MN sont $\sum_{i=1}^n n_{i,1}M_i, \dots, \sum_{i=1}^n n_{i,n}M_i$. En utilisant la propriété **c)** de Proposition 22 on obtient

$$\begin{aligned}
\det(MN) &= \det(\sum_{i_1=1}^n n_{i_1,1}M_{i_1}, \dots, \sum_{i_n=1}^n n_{i_n,n}M_{i_n}) \\
&= \sum_{1\leqslant i_1,\dots,i_n\leqslant n} n_{i_1,1} \dots n_{i_n,n} \det(M_{i_1}, \dots, M_{i_n}) \\
&= \sum_{1\leqslant i_1,\dots,i_n\leqslant n} n_{i_1,1} \dots n_{i_n,n} \det(M_{i_1}, \dots, M_{i_n}),
\end{aligned}$$

et grâce à la propriété **b)** de Proposition 22 on peut restreindre la somme prècèdente aux indices $i_1, \dots, i_n$ distincts deux à deux. Notons

$$E := \{(i_1, \dots, i_n)/ \quad 1 \leqslant i_1, \dots, i_n \leqslant n, \quad \text{pour tout } 1 \leq j \neq k \leq n, \quad i_j \neq i_k\}.$$

Cet ensemble E est en bijection avec S_n, il suffit d'associer à un n-uplet $(i_1, \dots, i_n)$ de E la permutation σ de S_n telle que pour tout j entre 1 et n, $\sigma(j) = i_j$. Ainsi

$$\begin{aligned}
\det(MN) &= \sum_{\sigma\in S_n} n_{\sigma(1),1} \dots n_{\sigma(n),n} \det(M_{\sigma(1)}, \dots, M_{\sigma(n)}) \\
&= \sum_{\sigma\in S_n} n_{\sigma(1),1} \dots n_{\sigma(n),n}\, \varepsilon(\sigma)\det(M_1, \dots, M_n) \\
&= \Big(\sum_{\sigma\in S_n} \varepsilon(\sigma)\, n_{\sigma(1),1} \dots n_{\sigma(n),n}\Big)\det M \\
&= \det(N)\det(M),
\end{aligned}$$

grâce à la propriété **a)** de Proposition 22 et à Proposition 15. □

Remarque 24. *Il n'est pas inutile de répéter ici que le déterminant n'est pas linéaire ! On a en général*

$$\det(M + N) \neq \det(M) + \det(N).$$

On peut prendre par exemple le cas des matrices carrées :

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \text{ and } N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Corollaire 25. *Soient M et N deux matrices carrées d'ordre n, alors*

$$\det(MN) = \det(NM)$$

même si $MN \neq NM$.

Démonstration. D'après le théorème précédent on a :

$$\det(MN) = \det(M)\det(N) = \det(N)\det(M) = \det(NM).$$

□

Théorème 26.

Soit M une matrice carrée d'ordre n, alors

$$M \text{ est inversible} \iff \det(M) \neq 0,$$

et dans ce cas

$$\det(M^{-1}) = (\det(M))^{-1}.$$

Démonstration. On a

$$M \text{ est inversible} \iff MM^{-1} = \mathrm{I}_n \Rightarrow \det(M)\det(M^{-1}) = \det(MM^{-1}) = \det(\mathrm{I}_n) = 1,$$

et donc on en déduit que $\det(M) \neq 0$.

Pour la réciproque on va démontrer la contraposée. Si la matrice M n'est pas inversible, en particulier, comme c'est une matrice carrée, cela implique que son image n'est pas $\mathbb{R}^n$ tout entier, or cette image est l'espace vectoriel engendré par les vecteurs $Me_1, \ldots, Me_n$, où $e_1, \ldots, e_n$ sont les vecteurs de la base canonique. On se convainc rapidement que pour tout i, Me_i est la i ème colonne de M. Ainsi les vecteurs-colonnes de M sont liés, donc l'un des vecteurs est combinaison linéaire des autres et donc $\det(M) = 0$. Ainsi on en déduit $\det(M) \neq 0 \Rightarrow M$ est inversible. □

Nous allons maintenant caractériser les déterminants de matrices qui sont semblables. Rappelons que deux matrices carrées d'ordre n, M et N, sont *semblables* s'il existe une matrice inversible P telle que $N = P^{-1}MP$. Pour que deux matrices soient semblables, il faut et il suffit qu'elles soient les matrices du même endomorphisme d'un espace vectoriel E dans deux bases.

Proposition 27. *Des matrices semblables ont même déterminant.*

Démonstration. On a

$$\det(N) = \det(P^{-1}MP) = \det(P^{-1})\det(M)\det(P) = \frac{1}{\det(P)}\det(M)\det(P) = \det(M).$$

□

Nous pouvons ainsi définir le déterminant d'un endomorphisme.

Théorème 28.

Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. Alors le scalaire $\det_{\mathcal{B}}(f(e_1), \ldots, f(e_n))$ *ne dépend pas de la base* $\mathcal{B} = (e_1, \ldots, e_n)$ *choisie. On l'appelle* déterminant de l'endomorphisme f *et on le note* $\det(f)$.

Démonstration. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ deux bases de E. On note M et N les matrices de l'endomorphisme f de E dans ces deux bases et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors on a $N = P^{-1}MP$ et d'après la proposition précédente $\det(M) = \det(N)$, d'où le résultat. □

Théorème 29.

> *Soient f et g deux endomorphismes d'un K-espace vectoriel E. Alors les propriétés suivantes sont vérifiées :*
> *a)* $\det(f \circ g) = \det(f)\det(g)$ *;*
> *b)* f *est bijectif si et seulement si* $\det(f) \neq 0$ *et alors on a* $\det(f^{-1}) = \dfrac{1}{\det(f)}$.

DÉMONSTRATION. Ces résultats s'obtiennent facilement lorsqu'on passe à l'écriture matricielle. □

II.4. Formes multilinéaires.

Nous allons étudier les applications qui vérifient les propriétés de Proposition 22. Nous allons voir qu'elles caractérisent univoquement le déterminant, à la normalisation

Définition 30. *Soit E un $K-$espace vectoriel. Soit l'application suivante :*

$$F : \overbrace{E \times E \times \cdots \times E}^{n\ fois} \to K$$

On dit que F est une forme n-linéaire *ou une* forme multilinéaire *sur E, si pour tout $i = 1, \ldots, n$, pour tout $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in E$, l'application $x \mapsto F(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$ est linéaire de E dans K. On note $\mathcal{L}^n(E)$ l'ensemble des formes n-linéaires sur E.*

La définition dit juste qu'une forme multilinéaire est linéaire par rapport à chacune de ses variables. Une forme linéaire est simplement une application linéaire de E dans K. Si $n = 2$ on dit que F est une *forme bilinéaire* et si $n = 3$ on dit que F est une *forme trilinéaire*.

Exemple 31. *Soit $F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ une forme bilinéaire sur $\mathbb{R}^2$. Posons $e_1 = (1,0)$ et $e_2 = (0,1)$ et $x_1 = x_{11}e_1 + x_{21}e_2$, $x_2 = x_{12}e_1 + x_{22}e_2$. On a*

$$F(x_1, x_2) = F(x_{11}e_1 + x_{21}e_2, x_{12}e_1 + x_{22}e_2) = ax_{11}x_{12} + bx_{11}x_{22} + cx_{21}x_{12} + dx_{21}x_{22}$$

avec $a = F(e_1, e_1)$, $b = F(e_1, e_2)$, $c = F(e_2, e_1)$ et $d = F(e_2, e_2)$. Inversement, on vérifie facilement que toute application de cette forme est bilinéaire.

Proposition 32. *Soit E un $K-$espace vectoriel. L'ensemble des formes n-linéaires sur E, $\mathcal{L}^n(E)$, muni de l'addition des fonctions et de la multiplication par un scalaire est un K-espace vectoriel.*

DÉMONSTRATION. On montre que c'est un sous-espace vectoriel de l'espace vectoriel des fonctions définies sur E et à valeurs dans K. □

Définition 33. *Soit E un $K-$espace vectoriel. Une forme n-linéaire F sur E est dite* alternée *si pour tout $(x_1, \ldots, x_n) \in E^n$ tel qu'il existe $1 \leq i, j \leq n$, $i \neq j$, $x_i = x_j$, alors $F(x_1, \ldots, x_n) = 0$.*

Définition 34. *On note $\mathcal{A}^n(E)$ le sous-espace vectoriel de $\mathcal{L}^n(E)$ constitué des formes n-linéaires alternées sur E.*

Définition 35. *Soit E un $K-$espace vectoriel. Une forme n-linéaire F sur E est dite*

— symétrique *si pour tout $(x_1, \ldots, x_n) \in E^n$, pour tout $1 \leq i, j \leq n$, $i \neq j$,*

$$F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n),$$

— antisymétrique *si pour tout $(x_1, \ldots, x_n) \in E^n$, pour tout $1 \leq i, j \leq n$, $i \neq j$,*

$$F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n). \tag{1}$$

Proposition 36. *Soit E un $K-$espace vectoriel. Si K est un corps de caractéristique différente de 2 alors une forme n-linéaire F sur E est alternée si et seulement si elle est antisymétrique.*

DÉMONSTRATION. Si F est antisymétrique, alors pour tout $(x_1, \ldots, x_n) \in E^n$ et pour tout $1 \leq i, j \leq n,\ i \neq j$, on a (1). En particulier quand $x_i = x_j$, on obtient

$$2F(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) = 0.$$

Comme K n'est pas de caractéristique 2 on en déduit que $F(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) = 0$.

Pour la réciproque il n'y a pas besoin de supposer que la caractéristique de K est différente de 2. Si F est alternée on a

$$\begin{aligned} 0 &= F(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\ &= F(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) + F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \\ &\quad + F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + F(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\ &= F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n), \end{aligned}$$

d'où le résultat. □

Exemple 37. *En reprenant l'exemple précédent, F est alternée si et seulement si pour tout (x_{11}, x_{12}) on a*

$$0 = F(x_{11}e_1 + x_{21}e_2, x_{11}e_1 + x_{21}e_2) = ax_{11}^2 + (b + c)x_{11}x_{21} + dx_{21}^2$$

c'est-à-dire seulement si $a = b + c = d = 0$. Donc la forme f s'écrit

$$F((x_{11}, x_{21}), (x_{12}, x_{22})) = b(x_{11}x_{22} - x_{21}x_{12}).$$

Proposition 38. *Soit F une forme n-linéaire alternée définie sur un K-espace vectoriel E. Alors on a*

a) Pour toute permutation σ de S_n :

$$\forall (x_1, \ldots, x_n) \in E^n, \quad F(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma) F(x_1, \ldots, x_n);$$

b) Le scalaire $F(x_1, \ldots, x_n)$ ne change pas lorsqu'on ajoute à l'un des x_i une combinaison linéaire des autres ;

c) Le scalaire $F(x_1, \ldots, x_n)$ est nul si les vecteurs x_i sont liés. En particulier toute forme n-linéaire alternée est nulle si $n > \dim(E)$

DÉMONSTRATION. On a vu dans la proposition précédente que si F est une forme multilinéaire et si τ est une transposition alors $F(x_{\tau(1)}, \ldots, x_{\tau(n)}) = -F(x_1, \ldots, x_n)$. Comme toute permutation σ se décompose en produit de transpositions, $\sigma = \tau_k \circ \cdots \circ \tau_1$, le point a) s'obtient en effectuant une récurrence sur k.

Pour b) on a

$$\begin{aligned} F(x_1, \ldots, x_{i-1}, x_i + \sum_{j \neq i} \lambda_j x_j, x_{i+1}, \ldots, x_n) &= F(x_1, \ldots, x_n) + \sum_{j \neq i} \lambda_j F(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\ &= F(x_1, \ldots, x_n) \end{aligned}$$

car F est alternée, d'où le résultat. Enfin, si les x_i sont liés, l'un d'eux est combinaison linéaire des autres.

D'après b), on peut ajouter à ce vecteur l'opposé de cette combinaison linéaire sans changer la valeur de F, qui est donc nulle. Ceci montre c). □

Théorème 39.

> *Soit E un K espace vectoriel de dimension n et $\mathcal{B} := (e_1, \ldots, e_n)$ une base de E. Alors l'application* $\det : E^n \to K$, *qui à $(x_1, \ldots, x_n)$ dans E^n associe* $\det_{\mathcal{B}}(x_1, \ldots, x_n)$ *est l'unique forme n-linéaire alternée sur E^n telle que* $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$.

Remarque 40. *L'espace $\mathcal{A}^n(E)$ est donc de dimension* 1.

DÉMONSTRATION. Soit L une forme n-linéaire alternée sur E^n telle que $L(e_1, \ldots, e_n) = 1$. Décomposons $(x_1, \ldots, x_n)$ selon la base $(e_1, \ldots, e_n)$ en écrivant pour tout $i = 1, \ldots, n$, $x_i = \sum_{j=1}^n \alpha_{i,j} e_j$. Alors en utilisant la multilinéarité et le caractère alternée comme dans la preuve du Théorème 23 on obtient

$$\begin{aligned} L(x_1, \ldots, x_n) &= \Big(\sum_{\sigma \in S_n} \varepsilon(\sigma)\, \alpha_{1,\sigma(1)}\, \alpha_{n,\sigma(n)} \Big) L(e_1, \ldots, e_n) \\ &= \det_{\mathcal{B}}(x_1, \ldots, x_n) L(e_1, \ldots, e_n), \end{aligned}$$

ce qui termine la preuve. □

III. Calcul de déterminants

III.1. Déterminants de matrices particulières.

Nous allons voir que pour certaines matrices le calcul du déterminant se fait assez facilement. Considérons tout d'abord des matrices dont l'écriture par blocs est

$$M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix},$$

où M' est une matrice carrée d'ordre p, M'' est une matrice carrée d'ordre q, N est une matrice à p lignes et q colonnes et le dernier bloc (noté 0 dans la matrice) est la matrice nulle à q lignes et p colonnes, avec $p + q = n$ l'ordre de M.

Proposition 41. *Le déterminant de la matrice* $M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix}$ *est* $\det(M')\det(M'')$.

DÉMONSTRATION. Notons $M = (m_{ij})_{1 \leq i,j \leq n}$. On a

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) m_{\sigma(1)1} \ldots m_{\sigma(n)n}.$$

Pour qu'un terme soit non nul il faut que

$$\sigma(1), \ldots, \sigma(p) \in \{1, \ldots, p\},$$

et donc

$$\sigma(p+1), \ldots, \sigma(n) \in \{p+1, \ldots, n\}.$$

Il existe donc des permutations $\sigma' \in S_p$ et $\sigma'' \in S_q$ telle que

$$\begin{aligned} &\sigma(j) = \sigma'(j) \quad \text{si} \quad 1 \leq j \leq p, \\ &\sigma(p+k) = p + \sigma''(k) \quad \text{si} \quad 1 \leq k \leq q. \end{aligned}$$

On a donc

$$\det(M) = \sum_{\sigma' \in S_p, \sigma'' \in S_q} \varepsilon(\sigma) m'_{\sigma'(1)_1} \ldots m'_{\sigma'(p)p} m''_{\sigma''(p+1)(p+1)} \ldots m''_{\sigma''(n)n}.$$

En décomposant σ' et σ'' en produit de transpositions on remarque que $\varepsilon(\sigma) = \varepsilon(\sigma')\varepsilon(\sigma'')$, et donc

$$\det(M) = \left(\sum_{\sigma' \in S_p} \varepsilon(\sigma') m'_{\sigma'(1)_1} \ldots m'_{\sigma'(p)p} \right) \left(\sum_{\sigma'' \in S_q} \varepsilon(\sigma'') m''_{\sigma''(1)_1} \ldots m''_{\sigma''(p)p} \right) = \det(M')\det(M'')$$

ce qui montre la proposition. □

Exemple 42. *En combinant les propriétés que l'on a vues, on obtient assez facilement que*

$$\begin{vmatrix} 11 & 0 & 12 & 0 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 7 & 5 & 8 & 6 \end{vmatrix} = -\begin{vmatrix} 7 & 5 & 8 & 6 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 11 & 0 & 12 & 0 \end{vmatrix} = \begin{vmatrix} 6 & 5 & 8 & 7 \\ 2 & 1 & 4 & 3 \\ 0 & 0 & 10 & 9 \\ 0 & 0 & 12 & 11 \end{vmatrix}$$
$$= \begin{vmatrix} 6 & 5 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} 10 & 9 \\ 12 & 11 \end{vmatrix} = (6.1 - 5.2)(10.11 - 9.12) = -8.$$

Une simple généralisation de ce résultat donne :

Corollaire 43. *Soit M une matrice de la forme*

$$\begin{pmatrix} M_1 & * & * & * \\ 0 & M_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & M_r \end{pmatrix}$$

où, pour i compris entre 1 et r, M_i est une matrice carrée d'ordre k_i, avec $k_1 + \cdots + k_r = n$. Alors on a

$$\det(M) = \det(M_1)\dots\det(M_r)$$

On en déduit alors le déterminant d'une matrice triangulaire supérieure. Rappelons qu'une matrice carrée d'ordre n est dite *triangulaire supérieure (inférieure)* si tous ses coefficients en dessous (respectivement au-dessus) de sa diagonale sont nuls.

Corollaire 44. *Soit M une matrice triangulaire supérieure, c'est-à-dire*

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m_{nn} \end{pmatrix},$$

alors $\det(M) = m_{11}\dots m_{nn}$

Remarque 45. *Le résultat est identique pour une matrice diagonale (matrice triangulaire supérieure dont les coefficients au-dessus de la diagonale sont tous nuls). D'autre part il reste encore vrai pour les matrices triangulaires inférieures (transposées de matrices triangulaires supérieures).*

III.2. Méthodes de calcul.

Dans la pratique, il est extrêmement rare que l'on calcule un déterminant par la formule directe, qui est bien trop compliquée. On a recours à des astuces. Nous allons donner une formule qui permet effectivement de calculer le déterminant d'une matrice carrée d'ordre n.

Définition 46. *Soit $M = (m_{ij})_{1\le i,j\le n}$ une matrice carrée d'ordre n. On note M_{ij} la matrice d'ordre $n-1$, obtenue à partir de M en supprimant la i ième ligne et la j ième colonne (tous les autres coefficients restent dans le même ordre). On appelle* mineur *de m_{ij} le scalaire* $\det(M_{ij})$ *et* cofacteur *de m_{ij} le scalaire* $\Delta_{ij} = (-1)^{i+j}\det(M_{ij})$.

Théorème 47.

Avec les notations de la définition précédente on a :
a) développement par rapport à la colonne j, $j = 1, \dots, n$,

$$\det(M) = (-1)^{1+j}\, m_{1j} \det(M_{1j}) + \cdots + (-1)^{n+j}\, m_{nj} \det(M_{nj}) = m_{1j}\, \Delta_{1j} + \cdots + m_{nj}\, \Delta_{nj};$$

b) développement par rapport à la ligne i, $i = 1, \dots, n$,

$$\det(M) = (-1)^{i+1}\, m_{i1} \det(M_{i1}) + \cdots + (-1)^{i+n}\, m_{in} \det(M_{in}) = m_{i1}\, \Delta_{i1} + \cdots + m_{in}\, \Delta_{in}.$$

DÉMONSTRATION. Démontrons la première formule (la seconde se démontre de la même façon ou s'obtient en considérant la transposée de M). Pour une colonne de M, on peut écrire

$$\begin{pmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{nj} \end{pmatrix} = \begin{pmatrix} m_{1j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ m_{1j} \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{nj} \end{pmatrix}.$$

Comme le déterminant est $n-$linéaire, en particulier il est linéaire en la j ième colonne. Si on note N_{ij} la matrice obtenue à partir de M en remplaçant tous les éléments de la j ième colonne, sauf celui dans la i ième ligne, par 0, on obtient

$$\det(M) = \det(N_{1j}) + \cdots + \det(N_{nj}).$$

Si l'on fait passer la j ième colonne de N_{ij} à la première place, sans changer l'ordre des autres, on effectue $j-1$ transpositions de colonnes, donc le déterminant est multiplié par $(-1)^{j-1}$; de même, si l'on fait passer la i ième ligne de N_{ij} à la première place (sans changer l'ordre des autres) le déterminant est multiplié par $(-1)^{i-1}$. On a donc

$$\det(N_{ij}) = (-1)^{(i-1)+(j-1)} \begin{vmatrix} m_{ij} & m_{i1} & \cdots & m_{i\,j-1} & m_{i\,j+1} & \cdots & m_{in} \\ 0 & & & & & & \\ \vdots & & M_{ij} & & & & \\ 0 & & & & & & \end{vmatrix} = (-1)^{i+j}\, m_{ij} \det(M_{ij})$$

d'après le corollaire 20. On obtient donc le a). □

Exemple 48. *Il est avantageux de développer selon une ligne ou une colonne où il y a beaucoup de zéros :*

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 0 & 5 & 6 \\ 0 & 0 & 3 & 0 \\ 1 & 6 & 4 & 2 \end{vmatrix} &= (-1)^{3+1}.0. \begin{vmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 6 & 4 & 2 \end{vmatrix} + (-1)^{3+2}.0. \begin{vmatrix} 1 & 3 & 4 \\ 4 & 5 & 6 \\ 1 & 4 & 2 \end{vmatrix} + (-1)^{3+3}.3. \begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} + (-1)^{3+4}.0. \begin{vmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \\ 1 & 6 & 4 \end{vmatrix} \\
&= 3. \begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} = 3. \begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ -2 & 0 & -10 \end{vmatrix} \quad L_3 \leftarrow L_3 - 3L_1 \\
&= 3.(-1)^{1+2}.2. \begin{vmatrix} 4 & 6 \\ -2 & -10 \end{vmatrix} = -6(4.(-10) - 6.(-2)) = 168.
\end{aligned}$$

IV. Application des déterminants

Nous verrons dans le chapitre suivant que les déterminants vont nous être très utiles pour calculer le polynôme caractéristique d'un endomorphisme ou d'une matrice. Ici nous donnons d'autres applications.

IV.1. Calcul de l'inverse d'une matrice.

Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On peut lui associer n^2 cofacteurs Δ_{ij} pour $1 \leq i, j \leq n$. Il est donc possible de construire une matrice carrée d'ordre n dont les coefficients sont les cofacteurs de M, c'est la matrice des cofacteurs dont la définition est donnée ci-dessous.

Définition 49. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On appelle* comatrice *de M la matrice notée* $\mathrm{Com}(M)$ *telle que le coefficient de la i ième ligne et de la j ième colonne soit exactement le cofacteur Δ_{ij} de M.*

Nous donnons maintenant une relation entre une matrice et sa comatrice, ainsi qu'une expression de l'inverse d'une matrice.

Théorème 50.

Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On a

$$M\,{}^t\mathrm{Com}(M) = \det(M)I_n \ \text{ et } \ {}^t\mathrm{Com}(M)\,M = \det(M)I_n.$$

De plus si M est inversible $(\det(M) \neq 0)$ *alors on a*

$$M^{-1} = \frac{1}{\det(M)}\,{}^t\mathrm{Com}(M).$$

DÉMONSTRATION. Le terme d'indice (i, j) du produit $M\,{}^t\mathrm{Com}(M)$ est $\sum_{k=1}^{n} m_{ik}\Delta_{jk}$ (on multiplie par la transposée). Distinguons deux cas :

— Si $j = i$ on reconnaît la deuxième formule du théorème 47 et donc on obtient $\det(M)$.
— Sinon c'est le développement suivant la j ième ligne, du déterminant de la matrice obtenue à partir de M en ayant remplacé la j ième ligne par la i ième. Cette matrice a deux lignes égales, donc son déterminant est nul. On obtient donc $M\,{}^t\mathrm{Com}(M) = \det(M)I_n$.

L'autre égalité s'obtient de la même manière en utilisant le développement suivant les colonnes.

Pour le dernier résultat, lorsque M est inversible il suffit de diviser tous les termes dans la première formule par $\det(M)$. □

Exemple 51. *Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible c'est-à-dire telle que* $\det(M) = ad - bc \neq 0$. *Alors l'inverse de M est donné par*

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

IV.2. Système de Cramer.

On considère ici des systèmes linéaires qui ont un nombre d'équations égal au nombre d'inconnues.

Définition 52. *Soit n un entier naturel non nul. On appelle* système d'équations linéaires *de n équations à n inconnues sur le corps K, le système :*

$$(S)\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right.$$

où les a_{ij} et b_i sont des éléments de K pour $i = 1, \ldots, n$ et $j = 1, \ldots, n$.

On peut écrire matriciellement ce système de la manière suivante. Soit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

la matrice du système. On pose encore

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Alors le système est équivalent à l'équation matricielle $AX = B$.

Théorème 53.

On dit qu'un système linéaire de n équations à n inconnues est de Cramer*, si sa matrice est inversible. Alors le système a une unique solution qui est*

$$X = A^{-1}B.$$

DÉMONSTRATION. Le résultat est immédiat. □

Proposition 54. *Soit (S) un système de Cramer, d'équation matricielle $AX = B$. Pour i, $1 \leq i \leq n$, on définit la matrice A_i obtenue en remplaçant dans A la i ième colonne par le second membre B. Alors la solution du système est donnée par : pour tout i compris entre 1 et n*

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

DÉMONSTRATION. Comme le système est de Cramer, il admet une unique solution. Si on note $C_1, \dots, C_n$ les vecteurs colonnes de la matrice A du système, on peut écrire

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = B.$$

Maintenant, si A_i est la matrice obtenue en remplaçant dans A la i ième colonne par le second membre B, on a

$$\begin{aligned} \det(A_i) &= \det(C_1, \dots, C_{i-1}, B, C_{i+1}, \dots, C_n) \\ &= \det(C_1, \dots, C_{i-1}, x_1C_1 + x_2C_2 + \cdots + x_nC_n, C_{i+1}, \dots, C_n) \\ &= \sum_{k=1}^{n} x_k \det(C_1, \dots, C_{i-1}, C_k, C_{i+1}, \dots, C_n) \\ &= x_i \det(C_1, \dots, C_{i-1}, C_i, C_{i+1}, \dots, C_n) = x_i \det(A) \end{aligned}$$

car l'application déterminant est une forme n-linéaire alternée. Comme $\det(A) \neq 0$, ceci montre le résultat. □

Exemple 55. *On souhaite résoudre le système :*

$$(S) \left\{ \begin{array}{l} 2x_1 + 3x_2 + x_3 = 9 \\ x_1 + 2x_2 + 3x_3 = 6 \\ 3x_1 + x_2 + 2x_3 = 8 \end{array} \right. .$$

La matrice du système est $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. *Son déterminant est égal à* 18*, c'est donc un système de Cramer. La solution est donc*

$$\begin{aligned} x_1 &= \frac{1}{18}\begin{vmatrix} 9 & 3 & 1 \\ 6 & 2 & 3 \\ 8 & 1 & 2 \end{vmatrix} = \frac{35}{18}, \\ x_2 &= \frac{1}{18}\begin{vmatrix} 2 & 9 & 1 \\ 1 & 6 & 3 \\ 3 & 8 & 2 \end{vmatrix} = \frac{29}{18}, \\ x_3 &= \frac{1}{18}\begin{vmatrix} 2 & 3 & 9 \\ 1 & 2 & 6 \\ 3 & 1 & 8 \end{vmatrix} = \frac{5}{18}. \end{aligned}$$

IV.3. Orientation de l'espace.

Soit E un espace vectoriel de dimension finie sur $\mathbb{R}$. On définit sur l'ensemble des bases de E une relation $\mathcal{R}$ par :

- deux bases $\mathcal{B}$ et $\mathcal{B}'$ de E sont en relation par $\mathcal{R}$ si et seulement si le déterminant de la matrice de passage $P_{\mathcal{B}\to\mathcal{B}'}$ est strictement positif, c'est-à-dire

$$\mathcal{B}\ \mathcal{R}\ \mathcal{B}' \iff \det(P_{\mathcal{B}\to\mathcal{B}'}) > 0.$$

Proposition 56. *$\mathcal{R}$ est une relation d'équivalence qui a deux classes d'équivalences.*

Orienter l'espace E, c'est convenir que l'une des classes d'équivalence est constituée de *bases directes* et l'autre de *bases indirectes*. En général on convient que la base canonique est directe ce qui fixe l'orientation de l'espace.

Par exemple si $E = \mathbb{R}^3$, avec la convention que la base canonique $(\vec{i}, \vec{j}, \vec{k})$ est directe, alors $(\vec{j}, \vec{k}, \vec{i})$ et $(\vec{k}, \vec{i}, \vec{j})$ sont directes mais $(\vec{i}, \vec{k}, \vec{j})$, $(\vec{j}, \vec{i}, \vec{k})$ et $(\vec{k}, \vec{j}, \vec{i})$ sont indirectes.

DÉMONSTRATION. Commençons par montrer que $\mathcal{R}$ est une relation d'équivalence.

— $\mathcal{R}$ est réflexive car $P_{\mathcal{B}\to\mathcal{B}} = \mathrm{Id}$, donc $\det(P_{\mathcal{B}\to\mathcal{B}}) = 1 > 0$ ce qui donne $\mathcal{B}\ \mathcal{R}\ \mathcal{B}$.

— $\mathcal{R}$ est symétrique car $\mathcal{B}\ \mathcal{R}\ \mathcal{B}' \iff \det(P_{\mathcal{B}\to\mathcal{B}'}) > 0$, et comme $\det(P_{\mathcal{B}'\to\mathcal{B}})$ est égal à l'inverse de $\det(P_{\mathcal{B}\to\mathcal{B}'})$ ceci est équivalent à $\det(P_{\mathcal{B}'\to\mathcal{B}}) > 0$ c'est-à-dire $\mathcal{B}'\ \mathcal{R}\ \mathcal{B}$.

— $\mathcal{R}$ est transitive car si $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ sont trois bases de E alors on a

$$P_{\mathcal{B}\to\mathcal{B}''} = P_{\mathcal{B}\to\mathcal{B}'}P_{\mathcal{B}'\to\mathcal{B}''},$$

De plus si $\mathcal{B}\ \mathcal{R}\ \mathcal{B}'$ et $\mathcal{B}'\ \mathcal{R}\ \mathcal{B}''$ alors on obtient

$$\det(P_{\mathcal{B}\to\mathcal{B}''}) = \det(P_{\mathcal{B}\to\mathcal{B}'}P_{\mathcal{B}'\to\mathcal{B}''}) = \det(P_{\mathcal{B}\to\mathcal{B}'})\det(P_{\mathcal{B}'\to\mathcal{B}''}) > 0,$$

et donc $\mathcal{B}\ \mathcal{R}\ \mathcal{B}''$.

Ces trois points montrent que $\mathcal{R}$ est une relation d'équivalence.

De plus, il y a au moins deux classes d'équivalence car si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E alors $\mathcal{B}' = (e_2, e_1, \dots, e_n)$ est aussi une base de E et $\det(P_{\mathcal{B}\to\mathcal{B}}) = -1 < 0$.

Il n'y a que deux classes d'équivalence car étant donné deux bases $\mathcal{B}$ et $\mathcal{B}'$ de E, la matrice de passage $P_{\mathcal{B}\to\mathcal{B}'}$ est inversible donc soit $\det(P_{\mathcal{B}\to\mathcal{B}'}) > 0$ et, $\mathcal{B}$ et $\mathcal{B}'$ sont dans la même classe, soit $\det(P_{\mathcal{B}\to\mathcal{B}'}) < 0$ et, $\mathcal{B}$ et $\mathcal{B}'$ ne sont pas dans la même classe. $\square$

Chapitre 2

RÉDUCTION D'ENDOMORPHISMES

Pour décrire un endomorphisme f d'un espace vectoriel E, on cherche une base de E dans laquelle la matrice de f soit la plus simple possible. Pour diverses raisons, on voudrait que cette matrice soit *diagonale*, c'est-à-dire que les coefficients en dehors de la diagonale soient nuls. En d'autre termes, les vecteurs de cette base sont tels que leur image par f leur est colinéaire. On les appelle des *vecteurs propres* de f. L'avantage d'avoir une matrice diagonale est qu'il est alors plus facile de faire des calculs faisant intervenir f, par exemple le calcul des itérés de f.

Il n'existe pas toujours de base de E formée de vecteurs propres de f. Par exemple la rotation R_α dans le plan réel orienté, de matrice

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

dans la base canonique, n'a pas de vecteur propre si $\alpha \not\equiv 0 \mod \pi$, puisqu'aucun vecteur non nul x n'est colinéaire à $R_\alpha(x)$. Le problème dans ce cas se résout en passant du corps $\mathbb{R}$ au corps $\mathbb{C}$: nous verrons qu'un endomorphisme d'un espace vectoriel complexe a toujours un vecteur propre. Néanmoins, même sur $\mathbb{C}$, il peut ne pas y avoir "assez" de vecteurs propres pour former une base ; c'est par exemple le cas pour l'endomorphisme de $\mathbb{C}^2$ de matrice

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

dont tous les vecteurs propres sont colinéaires à e_1 et donc ne peuvent pas former une base. On cherchera pour ces endomorphismes une base dans laquelle leur matrice est aussi simple que possible, c'est-à-dire aussi proche que possible d'une forme diagonale.

Dans tout ce chapitre, les espaces vectoriels considérés sont des espaces vectoriels sur K de dimension finie, où K est le corps des nombres réels $\mathbb{R}$ ou le corps des nombres complexes $\mathbb{C}$. En fait les résultats sont encore vrais si K est un corps commutatif contenant $\mathbb{Q}$.

I. Diagonalisation

I.1. Valeur propre - Vecteur propre.

Définition 1. *Soit $f \in \mathcal{L}(E)$. On dit que $\lambda \in K$ est* valeur propre *de f s'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$; x est alors appelé* vecteur propre *de f associé à la valeur propre λ.*

Remarque 2. *Tous les multiples non nuls d'un vecteur propre de f sont encore des vecteurs propres de f pour la même valeur propre. L'ensemble des valeurs propres d'un endomorphisme f s'appelle le spectre* de f et est noté $\mathrm{Sp}(f)$.

Exemple 3. *Si f est une homothétie d'un espace vectoriel E, $f = a\mathrm{Id}_E$, alors tout vecteur non nul est un vecteur propre associé à la valeur propre a.*

Exemple 4. *Comme on l'a vu dans l'introduction il existe des endomorphismes qui n'ont pas de valeur propre ni de vecteur propre ; par exemple les rotations d'angle différent de* $0 \mod \pi$ *dans le plan réel.*

Exemple 5. *L'exemple que nous donnons ici concerne un espace vectoriel réel de dimension infinie.*
Soit E l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, indéfiniment dérivables. L'application $u : E \to E$ qui a

une fonction associe sa dérivée est un endomorphisme de E. Alors, pour tout réel λ, la fonction $f_\lambda(t) = e^{\lambda t}$ est un vecteur propre associé à la valeur propre λ, car $f_\lambda \neq 0$ et $u(f_\lambda) = f'_\lambda = \lambda f_\lambda$.

Dans le théorème suivant nous caractérisons de façon plus précise les valeurs propres d'un endomorphisme.

Théorème 6.

Soient $f \in \mathcal{L}(E)$ et λ un scalaire. Les assertions suivantes sont équivalentes :
(i) *λ est valeur propre de f ;*
(ii) *l'endomorphisme $f - \lambda \mathrm{Id}_E$* n'est pas injectif, *i.e. son* noyau *vérifie*

$$\mathrm{Ker}(f - \lambda \mathrm{Id}_E) = \{x \in E,\ (f - \lambda \mathrm{Id}_E)(x) = 0\} \neq \{0\};$$

(iii) $\det(f - \lambda \mathrm{Id}_E) = 0$ *;*
(iv) $\det(M - \lambda \mathrm{I}_n) = 0$ *où M est la matrice de f dans n'importe quelle base de E.*

DÉMONSTRATION. Pour que λ soit valeur propre de f il faut et il suffit qu'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$, c'est-à-dire que l'on ait $(f - \lambda \mathrm{Id}_E)(x) = 0$, ou encore que le noyau $\mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0\}$. Ceci entraîne l'équivalence de (i) et (ii). Pour que l'endomorphisme $f - \lambda \mathrm{Id}_E$ de l'espace vectoriel *de dimension finie* E ne soit pas injectif il faut et il suffit qu'il ne soit pas bijectif, c'est-à-dire que son déterminant soit nul, d'où l'équivalence de (ii) et (iii). Enfin par définition du déterminant d'un endomorphisme, (iii) et (iv) sont équivalentes. □

Ce qui précède montre que l'ensemble des vecteurs propres associés à une valeur propre λ auquel on ajoute le vecteur nul est exactement $\mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.

Définition 7. *Soit λ une valeur propre d'un endomorphisme f. On appelle* sous-espace propre *associé à λ, le sous-espace vectoriel de E défini par $E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.*

Remarque 8. *C'est en cherchant le noyau de l'application $f - \lambda \mathrm{Id}_E$ que l'on détermine les vecteurs propres associés à la valeur propre λ.*

I.2. Polynôme caractéristique.

Définition 9. *Le* polynôme caractéristique *de $f \in \mathcal{L}(E)$ est défini par $\chi_f(X) = \det(f - X \mathrm{Id}_E)$.*

Si E est un K-espace vectoriel alors $\chi_f(X) \in K[X]$. De plus, si M est la matrice de f dans une base quelconque $\mathcal{B}$ de E, alors

$$\chi_f(X) = \det(M - X \mathrm{I}_n).$$

Théorème 10.

Les valeurs propres d'un endomorphisme f sur un K-espace vectoriel E sont exactement les racines de son polynôme caractéristique qui sont dans K.

DÉMONSTRATION. On a les équivalences suivantes :

$$\begin{aligned} \lambda \in K \text{ est valeur propre de } f \quad &\Longleftrightarrow \quad (f - \lambda \mathrm{Id}_E) \text{ est non injectif } \Longleftrightarrow \det(f - \lambda \mathrm{Id}_E) = 0 \\ &\Longleftrightarrow \quad \chi_f(\lambda) = 0 \Longleftrightarrow \lambda \text{ est une racine de } \chi_f \text{ dans } K \end{aligned}$$

□

Soit M une matrice d'ordre n à coefficients dans K. Soit X une indéterminée, alors on peut écrire :

$$M - X\mathrm{I}_n = \begin{pmatrix} m_{11}-X & m_{12} & \cdots & m_{1n} \\ m_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & m_{(n-1)n} \\ m_{n1} & \cdots & m_{n(n-1)} & m_{nn}-X \end{pmatrix}$$

Le déterminant de cette matrice est une somme de produits de coefficients de la matrice, c'est donc bien un polynôme en X à coefficients dans K. Chacun des produits comporte n termes, qui sont des coefficients de chacune des colonnes ; ce sont donc des polynômes de degré au plus n. De plus un seul produit est de degré n, c'est le produit de tous les coefficients diagonaux (il correspond à la permutation $\sigma = \mathrm{Id}$). Comme la signature de l'identité est égale à $+1$, le terme de plus haut degré est $(-1)^n X^n$. On a donc montré le résultat suivant.

Proposition 11. *Le polynôme caractéristique d'une matrice d'ordre n (ou d'un endomorphisme d'un espace vectoriel de dimension n) est un polynome de degré n dont le terme dominant est $(-1)^n$.*

Comme un polynôme de degré n a au plus n racines on obtient :

Corollaire 12. *En dimension n un endomorphisme (ou une matrice d'ordre n) a au plus n valeurs propres distinctes.*

Exemple 13. *Le polynôme caractéristique de la matrice*

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

est

$$\chi_M(X) = \begin{vmatrix} -X & 1 \\ 1 & -X \end{vmatrix} = X^2 - 1.$$

Les valeurs propres sont ± 1. Les sous-espaces propres associés à ces valeurs propres se déterminent alors de la manière suivante :

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_1 \iff \begin{cases} -x+y=0 \\ x-y=0 \end{cases} \iff \{x=y \iff E_1 =< \begin{pmatrix} 1 \\ 1 \end{pmatrix} > .$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_{-1} \iff \begin{cases} x+y=0 \\ x+y=0 \end{cases} \iff \{y=-x \iff E_1 =< \begin{pmatrix} 1 \\ -1 \end{pmatrix} > .$$

Exemple 14. *La rotation plane R_θ a pour matrice dans la base canonique*

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

Son polynôme caractéristique est

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

Ce polynôme n'a pas de racines réelles si $\sin\theta$ est non nul. Par contre dans $\mathbb{C}$ il a deux racines $e^{\pm i\theta}$, qui sont donc les valeurs propres de R_θ. On peut vérifier que les vecteurs propres associés sont

$$\begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

Exemple 15. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice triangulaire. Alors $M - X\mathrm{I}_n$ est aussi une matrice triangulaire et le polynôme caractéristique (déterminant d'une matrice triangulaire) est donc le produit des coefficients diagonaux i.e. $\chi_M(X) = (m_{11}-X)\cdots(m_{nn}-X)$. Les valeurs propres de M sont donc les coefficents diagonaux de M. En particulier ce résultat est vrai pour une matrice diagonale.*

Définition 16. *Soit f un endomorphisme d'un K-espace vectoriel E.*
L'ordre de multiplicité d'une valeur propre λ *de f est l'ordre de multiplicité de la racine λ dans le polynôme caractéristique de f.*
Un polynôme de $K[X]$ est dit scindé *si toutes ses racines sont dans K autrement dit s'il se décompose dans $K[X]$ comme produit de polynômes de degré 1. Par extension on dira qu'un endomorphisme est scindé si son polynôme caractéristique l'est.*

On peut remarquer que si un endomorphisme f d'un espace vectoriel E de dimension n, est scindé, son polynôme caractéristique a la forme suivante :

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où r, $1 \leq r \leq n$, représente le nombre de valeurs propres distinctes, les $(\lambda_i)_{1\leq i\leq r}$ sont les différentes valeurs propres et les $(\alpha_i)_{1\leq i\leq r}$ sont leurs ordres de multiplicité respectifs. On a de plus $\displaystyle\sum_{i=1}^{r} \alpha_i = n$.

I.3. Étude des sous-espaces propres.

Théorème 17.

> *Soit $f \in \mathcal{L}(E)$. On considère $\{\lambda_1, \ldots, \lambda_r\}$ l'ensemble des valeurs propres deux à deux distinctes de f. Alors les sous-espaces propres $(E_{\lambda_i})_{1\leq i\leq r}$ sont en* somme directe, *i.e.* $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.

Rappelons que r sous-espaces vectoriels $(E_{\lambda_i})_{1\leq i\leq r}$ sont dits en somme directe si et seulement si l'une deux propriétés équivalentes suivantes est vérifiée :

(a) $x_1 + x_2 + \cdots + x_r = 0$, avec pour $i = 1, \ldots, r$, $x_i \in E_{\lambda_i} \Rightarrow \forall\ 1 \leq i \leq r$, $x_i = 0$;
(b) pour $1 \leq k \leq r$, $E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}$.

DÉMONSTRATION. On va établir ce résultat par récurrence sur r le nombre de valeurs propres distinctes. Si $r = 2$, alors on a $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. En effet soit $x \in E_{\lambda_1} \cap E_{\lambda_2}$, alors $f(x) = \lambda_1 x$ et $f(x) = \lambda_2 x$. On obtient donc $(\lambda_1 - \lambda_2)x = 0$ c'est-à-dire $x = 0$ car $\lambda_1 \neq \lambda_2$. Un vecteur propre est donc toujours associé qu'à une seule valeur propre.

Supposons maintenant la propriété vraie jusqu'au rang $k - 1$, c'est-à-dire que pour tout l entre 2 et $k - 1$,

$$E_{\lambda_l} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{l-1}}) = \{0\}.$$

Montrons alors que :

$$E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}.$$

Soit

$$x \in E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}).$$

On a $x = x_1 + \cdots + x_{k-1}$ avec pour $1 \leq i \leq k - 1$, $x_i \in E_{\lambda_i}$. En prenant l'image par f de cet élément, on obtient :

$$f(x) = \lambda_k x = \lambda_k(x_1 + \cdots + x_{k-1}),$$

et d'autre part

$$f(x) = f(x_1) + \cdots + f(x_{k-1}) = \lambda_1 x_1 + \cdots + \lambda_{k-1} x_{k-1}.$$

Ceci entraîne que

$$(\lambda_k - \lambda_1)x_1 + \cdots + (\lambda_k - \lambda_{k-1})x_{k-1} = 0.$$

Or par hypothèse de récurrence $E_{\lambda_1}, \ldots, E_{\lambda_{k-1}}$ sont en somme directe, et comme les valeurs propres sont deux à deux distinctes, on obtient $x_i = 0$ pour $1 \leq i \leq k-1$, c'est-à-dire $E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}$.
□

Théorème 18.

> *Soient $f \in \mathcal{L}(E)$ et λ une valeur propre de f. Alors la dimension du sous-espace propre associé à la valeur propre λ est inférieure ou égale à la multiplicité de la valeur propre λ, c'est-à-dire :* $\dim(E_\lambda) \leq \alpha_\lambda$.
> *En particulier, si λ est une valeur propre simple (multiplicité égale à 1) alors* $\dim(E_\lambda) = 1$

DÉMONSTRATION. On considère $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ une base de E telle que $(e_1, \dots, e_k)$ soit une base de E_λ. Dans cette base la matrice de f s'écrit

$$M = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & \mathcal{M}_1 \\ & & \lambda & \\ \hline 0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & \mathcal{M}_2 \\ 0 & \cdots & 0 & \end{array}\right)$$

Le polynôme caractéristique de f est donc

$$\chi_f(X) = \det(M - X\mathrm{I}_n) = (\lambda - X)^k \det(\mathcal{M}_2 - X\mathrm{I}_{n-k}).$$

Ainsi λ est racine de χ_f d'ordre supérieur ou égal à $k = \dim(E_\lambda)$. Pour le second point un sous-espace propre contient au-moins un vecteur propre (vecteur non nul) donc on a $1 \leq \dim(E_\lambda)$. Si on applique alors le premier point avec $\alpha_\lambda = 1$, on obtient le résultat. $\square$

I.4. Endomorphismes diagonalisables.

Définition 19. *On dit que $f \in \mathcal{L}(E)$ est* diagonalisable *s'il existe une base de E constituée de vecteurs propres.*

Remarque 20. *Dans une telle base la matrice de f s'écrit*

$$M = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Théorème 21.

> *Soit $f \in \mathcal{L}(E)$, dont $\{\lambda_1, \dots, \lambda_r\}$ est l'ensemble des valeurs propres deux à deux distinctes. f est diagonalisable si et seulement si ses sous-espaces propres vérifient : $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.*

DÉMONSTRATION. Supposons que f soit diagonalisable. Quitte à réordonner les vecteurs, la base de vecteurs propres de f est de la forme

$$(e_{1,\lambda_1}, \dots, e_{\alpha_1,\lambda_1}, \dots, e_{1,\lambda_r}, \dots, e_{\alpha_r,\lambda_r})$$

où pour $1 \leq i \leq r$ et pour $1 \leq j \leq \alpha_i$, e_{j,λ_i} est un vecteur propre associé à la valeur propre λ_i. Si $\alpha_i < \dim(E_{\lambda_i}) = \beta_i$ alors E_{λ_i} aurait une base de la forme $(e_{1,\lambda_i}, \dots, e_{\beta_i,\lambda_i})$. De plus, d'après le théorème 9,

$$(e_{1,\lambda_1}, \dots, e_{\alpha_1,\lambda_1}, \dots, e_{1,\lambda_i}, \dots, e_{\beta_i,\lambda_i}, \dots, e_{1,\lambda_r}, \dots, e_{\alpha_r,\lambda_r})$$

serait une famille libre. Ainsi on obtient une contradiction (car la famille libre aurait plus de vecteurs qu'une base). Donc pour tout $1 \leq i \leq r$, $\alpha_i = \dim(E_{\lambda_i})$, et comme les sous-espaces propres sont en somme

directe, on aboutit au résultat.
Réciproquement, si les sous-espaces propres de f vérifient

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}.$$

Alors en notant une base de E_{λ_i}, $\mathcal{B}_i = (e_{1,\lambda_i}, \ldots, e_{\alpha_i,\lambda_i})$, on a $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E formée de vecteurs propres. □

Théorème 22.

> *Soit $f \in \mathcal{L}(E)$, avec E un $K-$espace vectoriel. f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :*
> **(a)** *f est scindé, i.e. toutes ses valeurs propres sont dans K ;*
> **(b)** *Pour toute valeur propre λ de f on a $\dim(E_\lambda) = \alpha_\lambda$ qui est l'ordre de multiplicité de λ.*

Démonstration. Si f est diagonalisable alors il existe, d'après le théorème 12, une base de vecteurs propres, $(e_{1,\lambda_1}, \ldots, e_{\alpha_1,\lambda_1}, \ldots, e_{1,\lambda_r}, \ldots, e_{\alpha_r,\lambda_r})$, dans laquelle sa matrice s'écrit

$$M = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix}$$

Son polynôme caractéristique est égal à $(-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}$. Il est donc scindé, et pour $1 \leq i \leq r$, $\dim(E_{\lambda_i}) = \alpha_i =$ l'ordre de multiplicité de λ_i.
Réciproquement, pour tout sous-espace propre E_{λ_i} on a d'après la propriété (b), $\dim(E_{\lambda_i}) = \alpha_i$, et d'après (a), $\sum_{i=1}^{r} \alpha_i = n = \dim(E)$. Donc comme les sous-espaces propres sont en somme directe, on obtient que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

Corollaire 23. *Soit $f \in \mathcal{L}(E)$. Une condition suffisante pour que f soit diagonalisable est que f ait exactement n valeurs propres deux à deux distinctes.*

Démonstration. Si f possède n valeurs propres deux à deux distinctes, f est scindé, et de plus tout sous-espace propre est de dimension 1. f est donc diagonalisable d'après le théorème 13. □

Attention ce n'est pas une condition nécessaire. En effet, les homothéties $\lambda \mathrm{Id}_E$, par exemple, sont diagonalisables mais n'ont qu'une seule valeur propre λ. Ce sont d'ailleurs les seuls endomorphismes diagonalisables qui ont une seule valeur propre.

I.5. Exemple de diagonalisation.

Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{pmatrix}.$$

Le polynôme caractéristique de f est alors $\chi_f(X) = (1-X)^2(2-X)$. f a donc une valeur propre double $\lambda = 1$, et une valeur propre simple $\lambda = 2$. Recherchons maintenant les sous-espaces propres. E_1 le sous-espace propre associé à la valeur propre $\lambda = 1$ est déterminé par :

$$(M - 1\mathrm{I}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{l} 4x - 2y + 3z = 0 \\ 4x - 2y + 3z = 0 \end{array} \right. \iff \left\{ \; 4x - 2y + 3z = 0 \right.$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 2, engendré par deux vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Le sous-espace propre E_2 associé à la valeur propre $\lambda = 2$ est déterminé par :

$$(M - 2\mathrm{I}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} -x & = & 0 \\ 4x - 3y + 3z & = & 0 \\ 4x - 2y + 2z & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} x & = & 0 \\ y & = & z \end{array} \right.$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 1, engendré par un vecteur u_3 de coordonnées dans la base canonique

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

La somme des dimensions des sous-espaces propres est égale à 3, c'est-à-dire à celle de l'espace vectoriel E (ici $\mathbb{R}^3$), f est donc diagonalisable et sa matrice dans la base (u_1, u_2, u_3) est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

II. Trigonalisation

II.1. Endomorphismes trigonalisables.

Définition 24. *Une matrice carrée M d'ordre n est dite* trigonalisable *si elle est semblable à une matrice triangulaire T, supérieure ou inférieure, c'est-à-dire s'il existe une matrice inversible P telle que $M = P^{-1}TP$.*

Définition 25. *Soit $f \in \mathcal{L}(E)$. f est dit* trigonalisable *s'il existe une base B de E dans laquelle sa matrice est triangulaire supérieure ou inférieure.*

On peut remarquer que tout endomorphisme (ou matrice) diagonalisable est trigonalisable. D'autre part, si $f \in \mathcal{L}(E)$ est trigonalisable cela signifie qu'il existe une base B de E dans laquelle la matrice de f s'écrit

$$M = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Son polynôme caractéristique est donc égal à $\chi_f(X) = (-1)^n \prod_{i=1}^{n}(X - a_{ii})$. Il est donc scindé et ses valeurs propres sont exactement les $(a_{ii})_{1\leq i\leq n}$.

Théorème 26.

$f \in \mathcal{L}(E)$ est trigonalisable si et seulement si f est scindé.

DÉMONSTRATION. Si f est trigonalisable, d'après la remarque ci-dessus, f est scindé. Réciproquement supposons que f soit scindé et montrons par récurrence sur la dimension n de E que f est trigonalisable. Si $n = 1$, la propriété est vraie. Supposons qu'elle soit vraie jusqu'au rang $n-1$. Comme χ_f est scindé, il existe au moins une valeur propre λ. Soit alors u_1 un vecteur propre associé à cette valeur propre. On considère les vecteurs $(u_2, \dots, u_n)$ de façon à ce que $\mathcal{B} = (u_1, u_2, \dots, u_n)$ soit une base de E. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & M_1 & \\ 0 & & & \end{pmatrix}$$

M_1 est la matrice d'un endomorphisme g égal à la composée de f avec la projection de E sur le sous-espace de E de dimension $n-1$ engendré par $(u_2, \dots, u_n)$. On a

$$\chi_f(X) = (\lambda - X)\det(M_1 - X\mathrm{I}_{n-1}) = (\lambda - X)\chi_g(X),$$

et comme χ_f est scindé, χ_g l'est aussi. Donc par hypothèse de récurrence, il existe une base $(u'_2, \dots, u'_n)$ de ce sous-espace dans laquelle la matrice de g est triangulaire supérieure. Ainsi dans la base $\mathcal{B}' = (u_1, u'_2, \dots, u'_n)$ de E la matrice de f est triangulaire supérieure. □

Corollaire 27. *Tout endomorphisme sur un espace vectoriel complexe est trigonalisable.*

DÉMONSTRATION. Cela provient du fait que tout polynôme sur $\mathbb{C}[X]$ est scindé. □

II.2. Exemple de trigonalisation.

On considère f l'endomorphisme de $E = \mathbb{R}^4$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}.$$

Le calcul du polynôme caractéristique donne $\chi_f(X) = \chi_M(X) = (1-X)^4$. Il y a donc une seule valeur propre $\lambda = 1$ d'ordre 4. On détermine maintenant le sous-espace propre E_1 associé à cette valeur propre.

$$(M - 1\mathrm{I}_4)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} -3y + 3t & = & 0 \\ -2x - 7y + 13t & = & 0 \\ -3y + 3t & = & 0 \\ -x - 4y + 7t & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} x & = & 3t \\ y & = & t \end{array} \right. .$$

L'unique sous-espace propre E_1 est donc de dimension 2 ($< 4 = \dim(\mathbb{R}^4)$ donc f n'est pas diagonalisable), il est engendré par les vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

On complète (u_1, u_2) par (u_3, u_4) pour obtenir une base de $\mathbb{R}^4$. Ici on peut par exemple choisir $u_3 = e_1$ et $u_4 = e_2$. On a donc les relations suivantes :

$$e_1 = u_3,\ e_2 = u_4,\ e_3 = u_2 \text{ et } e_4 = u_1 - 3u_3 - u_4.$$

Ainsi :

$$\begin{cases} f(u_3) &= f(e_1) = e_1 - 2e_2 - e_4 = -u_1 + 4u_3 - u_4 \\ f(u_4) &= f(e_2) = -3e_1 - 6e_2 - 3e_4 - 4e_4 = -4u_1 - 3u_2 + 9u_3 - 2u_4 \end{cases}$$

La matrice de f dans la base (u_1, u_2, u_3, u_4) s'écrit alors

$$\widetilde{M} = \begin{pmatrix} 1 & 0 & -1 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & -1 & -2 \end{pmatrix}.$$

On considère la sous-matrice

$$M_1 = \begin{pmatrix} 4 & 9 \\ -1 & -2 \end{pmatrix}$$

qui est la matrice de la projection de f sur le sous-espace engendré par u_3 et u_4, dans la base canonique de ce sous-espace. On va maintenant trigonaliser cette matrice. Le polynôme caractéristique de cette matrice est

$$\chi_{M_1}(X) = (1 - X)^2.$$

Cette matrice n'a qu'une seule valeur propre double qui est 1. Le sous-espace propre associé a cette valeur propre est déterminé par

$$\begin{cases} 3x + 9y &= 0 \\ -x - 3y &= 0 \end{cases} \iff \{x = -3y\ .$$

Sa dimension est donc 1, et il est engendré par v_1 de coordonnées dans la base canonique du sous-espace

$$\begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

On le complète en une base du sous-espace avec un vecteur v_2 de coordonnées

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Les vecteurs correspondants dans l'espace $E = \mathbb{R}^4$ sont donc $v_1' = -3u_3 + u_4$ et $v_2' = u_4$. Ainsi

$$\begin{cases} f(v_1') &= -u_1 - 3u_2 + v_1' \\ f(v_2') &= -4u_1 - 3u_2 - 3v_1' + v_2' \end{cases}.$$

La matrice de f dans la base (u_1, u_2, v_1', v_2') s'écrit alors

$$\begin{pmatrix} 1 & 0 & -1 & -4 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

III. Polynômes d'endomorphismes - Polynôme minimal

III.1. Polynômes d'endomorphismes.

Soit E un $K-$espace vectoriel de dimension n et f un endomorphisme de E. On introduit les notations suivantes :

$$f^0 = \mathrm{Id}_E, \quad f^1 = f, \quad f^2 = f \circ f, \quad f^k = \overbrace{f \circ f \circ \cdots \circ f}^{k \text{ fois}}.$$

Plus généralement, si $Q(X) = a_0 + a_1 X + \cdots + a_k X^k$ est un polynôme de $K[X]$, alors on définit le *polynôme d'endomorphismes* $Q(f)$ par :

$$Q(f) = a_0 \mathrm{Id}_E + a_1 f + \cdots + a_k f^k.$$

Si M est la matrice de f dans une base $\mathcal{B}$ de E alors le *polynôme de matrices* $Q(M)$ défini par :

$$Q(M) = a_0 \mathrm{I}_n + a_1 M + \cdots + a_k M^k,$$

est la matrice de $Q(f)$ dans la base $\mathcal{B}$.

Proposition 28. *Pour tout endomorphisme f de E, il existe un polynôme non nul Q de $K[X]$ tel que $Q(f) = 0$.*

DÉMONSTRATION. E est un $K-$espace vectoriel de dimension n, donc $\mathcal{L}(E)$ est un $K-$espace vectoriel de dimension n^2. Par conséquent, les $n^2 + 1$ endomorphismes $\mathrm{Id}_E,\ f,\ f^2, \ldots,\ f^{n^2}$ sont liés. Il existe donc des coefficients $a_0,\ a_1, \ldots,\ a_{n^2}$ de K non tous nuls, tels que $a_0 \mathrm{Id}_E + a_1 f + \cdots + a_{n^2} f^{n^2} = 0$. C'est-à-dire que le polynôme non nul

$$Q(X) = a_o + a_1 X + \cdots + a_{n^2} X^{n^2}$$

de $K[X]$ vérifie $Q(f) = 0$. □

Définition 29. *On dit qu'un sous-ensemble non vide $\mathfrak{I}$ de $K[X]$ est un* idéal *s'il vérifie les deux propriétés suivantes :*

(a) $\forall Q_1 \in \mathfrak{I}, \forall Q_2 \in \mathfrak{I}, \quad Q_1 - Q_2 \in \mathfrak{I}.$
(b) $\forall Q_1 \in \mathfrak{I}, \forall B \in K[X], \quad Q_1 B \in \mathfrak{I}.$

Proposition 30. *Soit $\mathfrak{I}$ un idéal de $K[X]$. Il existe un polynôme $\mathfrak{P}$ tel que $\mathfrak{I}$ soit l'ensemble des polynômes multiples de ce polynôme, i.e. $\mathfrak{I} = \{\mathfrak{P}Q,\ Q \in K[X]\}$. On dit que $\mathfrak{P}$* engendre *$\mathfrak{I}$ ou qu'il est un* générateur *de $\mathfrak{I}$.*

DÉMONSTRATION. Si $\mathfrak{I} = \{0\}$ il suffit de prendre $\mathfrak{P} = 0$. Sinon, soit $\mathfrak{P}$ un polynôme non nul dans $\mathfrak{I}$ et de degré minimal. Comme $\mathfrak{I}$ est un idéal, il contient tous les multiples de $\mathfrak{P}$. Inversement si A est un élément quelconque de $\mathfrak{I}$, on fait la division euclidienne $A = \mathfrak{P}Q + R$. Comme $\mathfrak{I}$ est un idéal, $R = A - \mathfrak{P}Q$ appartient à $\mathfrak{I}$, et son degré est strictement inférieur à celui de $\mathfrak{P}$. Ainsi R est nul, à cause du choix de $\mathfrak{P}$, et A est donc un multiple de $\mathfrak{P}$. □

Remarque 31. *On peut remarquer qu'un idéal de $K[X]$, $\mathfrak{I} \neq \{0\}$, admet toujours un* unique générateur unitaire*, c'est-à-dire un générateur dont le coefficient dominant vaut 1. En effet si $\mathfrak{P}$ est un générateur de $\mathfrak{I}$ dont le coefficient dominant vaut $a \neq 0$ alors d'après la définition d'un idéal $\mathfrak{P}' = \mathfrak{P}/a$ appartient à $\mathfrak{I}$ et de plus est unitaire. Or, comme les degrés de $\mathfrak{P}'$ et $\mathfrak{P}$ sont égaux, $\mathfrak{P}'$ est encore un générateur de $\mathfrak{I}$. Maintenant si $\mathfrak{P}''$ est un autre générateur unitaire de $\mathfrak{I}$, alors $\mathfrak{P}''$ appartient à l'idéal $\mathfrak{I}$. Il existe donc $Q \in K[X]$ tel que $\mathfrak{P}'' = \mathfrak{P}'Q$. Mais comme $\mathfrak{P}''$ et $\mathfrak{P}'$ sont deux générateurs unitaires, ils ont même degré et même coefficient dominant. Donc $Q = 1$ et $\mathfrak{P}'' = \mathfrak{P}'$.*

III.2. Polynôme minimal.

Théorème 32.

Soient E un $K-$espace vectoriel et f un endomorphisme de E. L'ensemble $\mathfrak{I} = \{P \in K[X],\ P(f) = 0\}$ des polynômes annulateurs de f est un idéal différent de $\{0\}$. Cet idéal est appelé idéal annulateur *de f. Le générateur unitaire de cet idéal s'appelle le* polynôme minimal *de f et est noté $m_f(X)$.*

DÉMONSTRATION. D'après la Proposition 16 cet ensemble contient un élément non nul. Donc $\mathfrak{I}$ est non vide et est différent de $\{0\}$. Soient P_1 et P_2 deux éléments de $\mathfrak{I}$. Alors $(P_1 - P_2)(f) = P_1(f) - P_2(f) = 0$, c'est-à-dire que $P_1 - P_2$ appartient à $\mathfrak{I}$. De plus si $P \in \mathfrak{I}$ et si $Q \in K[X]$ alors on a $(PQ)(f) = P(f)Q(f) = 0$. Ainsi $PQ \in \mathfrak{I}$. Ainsi $\mathfrak{I}$ est un idéal de $K[X]$. □

Remarque 33. *Le polynôme minimal d'un endomorphisme f est donc l'unique polynôme de plus bas degré et unitaire (coefficient dominant est égal à 1), vérifiant $m_f(f) = 0$. De plus si M est la matrice de f dans n'importe quelle base $\mathcal{B}$ de E, $m_M(X) = m_f(X)$ et donc $m_f(M) = 0$. Comme un générateur d'un idéal est de degré supérieur ou égal à 1, il en est de même pour le polynôme minimal .*

Exemple 34. *Soit f une homothétie d'un $K-$espace vectoriel E. On a $f = \lambda \mathrm{Id}_E$, avec $\lambda \in K$. Donc le polynôme $X - \lambda$ est un polynôme annulateur de f. Comme il est de degré 1 et unitaire, c'est le générateur unitaire de l'idéal annulateur de f, c'est-à-dire que c'est le polynôme minimal de f. Réciproquement, si le polynôme minimal de f est $X - \lambda$, alors on a $f - \lambda \mathrm{Id}_E = 0$, ou encore $f = \lambda \mathrm{Id}_E$. Par conséquent, f est une homothétie. Nous venons donc de montrer que le polynôme minimal d'un endomorphisme est de degré 1 si et seulement si cet endomorphisme est une homothétie.*

III.3. Théorème de Cayley-Hamilton.

Théorème 35.

Soient E un $K-$espace vectoriel et f un endomorphisme de E. Le polynôme minimal de f divise le polynôme caractéristique de f.

Démonstration. D'après la remarque ci-dessus il suffit de montrer que $\chi_f(f) = 0$. Soit M la matrice de f dans une base $\mathcal{B}$ de E. C'est une matrice carrée d'ordre n. On pose $N = M - X\mathrm{I}_n$ et on note $\mathrm{Com}(N)$ sa comatrice. Les coefficients de $\mathrm{Com}(N)$ et donc de ${}^t\mathrm{Com}(N)$ sont des polynômes de degré inférieur ou égal à $n - 1$. On peut donc écrire

$$ {}^t\mathrm{Com}(N) = \sum_{j=0}^{n-1} A_j X^j $$

où les A_j sont des matrices carrées à coefficients dans K. D'après le théorème 24 du chapitre précédent on a

$$ N\, {}^t\mathrm{Com}(N) = \det(N)\mathrm{I}_n = \chi_f(X)\mathrm{I}_n, $$

c'est-à-dire

$$ (M - X\mathrm{I}_n) \sum_{j=1}^{n-1} A_j X^j = \chi_f(X)\mathrm{I}_n. $$

En notant maintenant

$$ \chi_f(X) = \sum_{j=0}^{n} a_j X^j $$

(avec $a_n = (-1)^n$) on obtient en identifiant les coefficients de même degré : $MA_0 = a_o\mathrm{I}_n,\ MA_1 - A_0 = a_1\mathrm{I}_n$, $\ldots,\ MA_{n-1} - A_{n-2} = a_{n-1}\mathrm{I}_n,\ -A_{n-1} = a_n\mathrm{I}_n$.
On en déduit

$$ \chi_f(M) = \sum_{j=0}^{n} a_j M^j = MA_0 + M(MA_1 - A_0) + \cdots + M^{n-1}(MA_{n-1} - A_{n-2}) + M^n(-A_{n-1}) = 0, $$

car les termes se regroupent et s'annulent. □

Remarque 36. *Un énoncé équivalent de ce résultat est que χ_f appartient à l'idéal annulateur*

$$ \mathfrak{I} = \{P \in K[X],\ P(f) = 0\}, $$

ou encore que $\chi_f(f) = 0$. Comme le polynôme caractéristique est de degré n, on en déduit que le polynôme minimal est au plus de degré n.

Exemple 37. *Soit la matrice*
$$M = \begin{pmatrix} 0 & 3 \\ -2 & -5 \end{pmatrix}.$$
Le polynôme caractéristique de cette matrice est
$$\begin{aligned} \chi_M(X) &= \begin{vmatrix} -X & 3 \\ -2 & -X-5 \end{vmatrix} = \begin{vmatrix} -X-2 & -X-2 \\ -2 & -X-5 \end{vmatrix} = (-X-2)\begin{vmatrix} 1 & 1 \\ -2 & -X-5 \end{vmatrix} \\ &= -(X+2)\begin{vmatrix} 1 & 0 \\ -2 & -X-3 \end{vmatrix} = -(X+2)(-X-3) = (X+2)(X+3). \end{aligned}$$
Comme le polynôme minimal, divise le polynôme caractéristique, et qu'il est unitaire de degré au moins 1, *on en déduit qu'il est égal à* $X+2$, $X+3$ *ou* $(X+2)(X+3)$. *Comme* $M + 2\mathrm{I}_2 \neq 0$ *et que* $M + 3\mathrm{I}_2 \neq 0$, *on obtient que* $m_M(X) = (X+2)(X+3)$.

Proposition 38. λ *est valeur propre de* f *si et seulement si* λ *est une racine de son polynôme minimal.*

DÉMONSTRATION. Si λ est racine de m_f alors comme m_f divise χ_f, λ est racine de χ_f, et donc d'après le théorème 5, λ est une valeur propre de f.
Réciproquement, soit λ une valeur propre de f dont on note x un vecteur propre associé. On effectue la division euclidienne de m_f par $X - \lambda$, et on obtient $m_f(X) = Q(X)(X-\lambda) + c$, où $\deg(c) \leq 0$, i.e. $c \in K$. On en déduit que
$$0 = m_f(f) = Q(f) \circ (f - \lambda \mathrm{Id}_E) + c\mathrm{Id}_E.$$
Si on applique cette relation au vecteur x, on trouve
$$0 = Q(f) \circ (f - \lambda \mathrm{Id}_E)(x) + c\mathrm{Id}_E(x) = Q(f)(f(x) - \lambda x) + cx = Q(f)(0) + cx = cx.$$
Or comme x n'est pas nul, ceci entraîne que $c = 0$, de sorte que le polynôme minimal s'écrit
$$m_f(X) = Q(X)(X - \lambda).$$
Cela signifie que λ est racine de m_f. □

Exemple 39. *Si on reprend l'exemple précédent, on a trouvé que le polynôme caractéristique de la matrice* M *est* $\chi_M(X) = (X+2)(X+3)$. *C'est-à-dire que les valeurs propres de* M *sont* -2 *et* -3. *Donc d'après ce qui précède, ce sont des racines du polynôme minimal. D'autre part comme il est unitaire et qu'il divise le polynôme caractéristique, on obtient directement que* $m_M(X) = (X+2)(X+3)$.

Exemple 40. *Soit la matrice*
$$M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 2 \end{pmatrix}.$$
Comme la matrice est triangulaire, le polynôme caractéristique de cette matrice est $\chi_M(X) = (1-X)(1-X)(2-X)$. *Les valeurs propres de* M *sont donc* 1 *et* 2. *Par définition du polynôme minimal et d'après le théorème de Cayley-Hamilton, on a* $m_M(X) = (X-1)(X-2)$ *ou* $(X-1)^2(X-2)$, *car toute valeur propre est racine de ce polynôme. Or lorsqu'on effectue le calcul suivant :*
$$(M - \mathrm{I}_3)(M - 2\mathrm{I}_3) = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & -5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -10 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$
Donc le polynôme minimal est $m_M(X) = (X-1)^2(X-2)$.

III.4. Lemme de décomposition des noyaux.

Les résultats que l'on démontre ici reposent sur l'énoncé suivant.

Lemme 41. *Soit f un endomorphisme d'un $K-$espace vectoriel E et P un polynôme de $K[X]$.*
(a) *Le sous-espace vectoriel $\mathrm{Ker}P(f)$ est stable par f, c'est-à-dire*

$$f\left(\mathrm{Ker}P(f)\right) \subset \mathrm{Ker}P(f).$$

(b) *Si $P = P_1P_2$ avec P_1 et P_2 deux polynômes premiers entre eux, alors*

$$\mathrm{Ker}P(f) = \mathrm{Ker}P_1(f) \oplus \mathrm{Ker}P_2(f).$$

Démonstration. Remarquons d'abord que f et $P(f)$ commutent. Si x appartient au noyau de $P(f)$ on a $P(f)\big(f(x)\big) = f\big(P(f)(x)\big) = 0$, de sorte que $f(x)$ est encore dans le noyau de $P(f)$. Cela montre a).

Comme P_1 et P_2 sont premiers entre eux, d'après le théorème de Bézout il existe Q_1 et Q_2 deux polynômes de $K[X]$ tels que

$$Q_1P_1 + Q_2P_2 = 1.$$

En appliquant cette relation à f on obtient

$$Q_1(f)P_1(f) + Q_2(f)P_2(f) = \mathrm{Id}_E.$$

Soit $x \in \mathrm{Ker}P_1(f) \cap \mathrm{Ker}P_2(f)$, on a $P_1(f)(x) = P_2(f)(x) = 0$, et donc d'après l'égalité ci-dessus

$$x = \mathrm{Id}_E(x) = Q_1(f)P_1(f)(x) + Q_2(f)P_2(f)(x) = 0.$$

$\mathrm{Ker}P_1(f)$ et $\mathrm{Ker}P_2(f)$ sont donc en somme directe. D'autre part, pour tout x dans $\mathrm{Ker}P(f)$ on a :

$$x = \mathrm{Id}_E(x) = Q_1(f)P_1(f)(x) + Q_2(f)P_2(f)(x).$$

Or, comme

$$P_2(f)\big(Q_1(f)P_1(f)(x)\big) = Q_1(f)\big(P(f)(x)\big) = 0,$$

$Q_1(f)P_1(f)(x)$ est dans le noyau de $P_2(f)$; de même le vecteur $Q_2(f)P_2(f)(x)$ est dans le noyau de $P_1(f)$. Cela montre b). □

Corollaire 42. *Soit f un endomorphisme d'un $K-$espace vectoriel E et l polynômes $P_1, \ldots, P_l$ de $K[X]$ premiers entre eux deux à deux. On pose $P = P_1P_2 \ldots P_l$. Alors*

$$\mathrm{Ker}P(f) = \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_l(f).$$

Démonstration. On procède par récurrence sur l. Le cas $l = 1$ est trivial. Le cas $l = 2$ est exactement celui du Lemme 25. On suppose la propriété vraie au rang $l - 1$ et on montre qu'elle l'est encore au rang l. Comme les polynômes sont premiers entre eux deux à deux P_l est premier avec $P_1P_2 \ldots P_{l-1}$. Donc par l'hypothèse de récurrence on a :

$$\begin{aligned}\mathrm{Ker}P(f) &= \mathrm{Ker}(P_1P_2 \ldots P_{l-1})(f) \oplus \mathrm{Ker}P_l(f) = \big(\mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f)\big) \oplus \mathrm{Ker}P_l(f)\\ &= \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f) \oplus \mathrm{Ker}P_l(f).\end{aligned}$$

□

III.5. Diagonalisation à l'aide du polynôme minimal.

Théorème 43.

Soient E un $K-$espace vectoriel et f un endomorphisme de E de valeurs propres distinctes $\{\lambda_1, \dots \lambda_r\}$. Les assertions suivantes sont équivalentes :
(i) *f est diagonalisable ;*
(ii) $m_f(X) = \prod_{i=1}^{r}(X - \lambda_i)$;
(iii) *m_f est scindé et a racines simples ;*
(iv) *f admet un polynôme annulateur scindé à racines simples.*

DÉMONSTRATION. $(i) \Rightarrow (ii)$: d'après la proposition 24 les valeurs propres sont les racines de m_f. Donc le polynome

$$M(X) = \prod_{i=1}^{r}(X - \lambda_i)$$

divise m_f.
Si x_i est un vecteur propre associé à λ_i, alors il est clair que $(f - \lambda_i \mathrm{Id}_E)(x_i) = 0$ et donc

$$M(f)(x_i) = \left(\prod_{j \neq i}(\lambda_j \mathrm{Id}_E - f)\right)(\lambda_i \mathrm{Id}_E - f)(x_i) = 0.$$

Si f est diagonalisable, alors en vertu du théorème 12 $E = \bigoplus_{i=1}^{r} E_{\lambda_i}$. Donc tout $x \in E$ se décompose en $x = x_1 + \cdots + x_r$ avec $x_i \in E_{\lambda_i}$, puis d'après le résultat précédent

$$M(f)(x) = \sum_{i=1}^{r} M(f)(x_i) = 0.$$

On obtient ainsi (ii).
$(ii) \Rightarrow (iii) \Rightarrow (iv)$: trivial.
$(iv) \Rightarrow (i)$: soit P un polynôme annulateur de f scindé à racines simples, que l'on peut toujours supposer unitaire (quitte à le multiplier par une constante) :

$$P(X) = \prod_{i=1}^{s}(X - \mu_i) \quad \text{avec les } \mu_i \text{ deux à deux distincts.}$$

Comme les polynômes $X - \mu_i$ sont premiers entre eux on peut appliquer le corollaire 26, ce qui donne

$$E = \mathrm{Ker}[P(f)] = \bigoplus_{i=1}^{s} \mathrm{Ker}(f - \mu_i \mathrm{Id}_E).$$

Si dans cette décomposition, on ne garde que les μ_i tels que $\mathrm{Ker}(f - \mu_i \mathrm{Id}_E)$ ne soit pas réduit à $\{0\}$, on a obtenu une décomposition de E en somme de sous-espaces propres de f. f est donc diagonalisable. □

Corollaire 44. *Soit $f \in \mathcal{L}(E)$ diagonalisable. Soit F un sous-espace de E, stable par f. Alors* l'endomorphisme induit *f_F est diagonalisable.*

DÉMONSTRATION. Si f est diagonalisable, alors d'après le résultat précédent, il existe $P \in K[X]$ scindé à racines simples tel que $P(f) = 0$. On a bien sur aussi $P(f_F) = 0$. Donc à nouveau d'après le théorème 27, f_F est diagonalisable. □

III.6. Diagonalisation simultanée.

Théorème 45.

Soient f et g deux endomorphismes diagonalisables de E qui commutent. Alors il existe une base de diagonalisation commune pour f et g.
De même, si A et B sont deux matrices diagonalisables qui commutent, alors il existe une matrice inversible P et deux matrices diagonales D et D' telles que

$$A = PDP^{-1} \quad B = PD'P^{-1}$$

Démonstration. Si f est diagonalisable, alors d'après le théorème 12, on a $E = \oplus_{i=1}^{r} E_{\lambda_i}$. Comme f et g commutent il est facile de voir que chaque E_{λ_i} est stable par g. L'endomorphisme g_i, induit par g sur E_{λ_i} est d'après le résultat précédent diagonalisable. On peut donc trouver une base $(e_1^i, \dots, e_{\gamma_i}^i)$ de E_{λ_i} qui est une base de vecteurs propres pour g_i. Bien sûr, chaque e_k^i est un vecteur propre de f associé à la valeur propre λ_i.
En rassemblant les bases obtenues pour chaque E_{λ_i}, on obtient une base de E, qui est une base commune de vecteurs propres pour f et g. □

Ce résultat peut s'étendre sans trop de difficultés au cas de p endomorphismes diagonalisables qui commutent.

IV. Sous-espaces caractéristiques

Dans cette partie on suppose que f est un endomorphisme scindé d'un espace vectoriel E de dimension n. On notera $\{\lambda_1, \dots \lambda_r\}$ l'ensemble de ses valeurs propres distinctes. Son polynôme caractéristique s'écrit donc

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où α_i est l'ordre de multiplicité de la valeur propre λ_i. De plus, par le Théorème de Cayley-Hamilton son polynôme minimal s'écrit

$$m_f(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i},$$

avec pour $1 \leq i \leq r$, $1 \leq \beta_i \leq \alpha_i$.

IV.1. Définition.

Définition 46. *On reprend les notations ci-dessus. On appelle* sous-espace caractéristique *de f relatif à la valeur propre λ_i le sous-espace vectoriel de E défini par $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$, où α_i est l'ordre de multiplicité de la valeur propre λ_i.*

Remarque 47. *Si $\alpha_i = 1$ alors $N_{\lambda_i} = E_{\lambda_i}$, autrement dit pour une valeur propre simple le sous-espace caractéristique est égal au sous-espace propre.*

Exemple 48. *Soient E un espace vectoriel de dimension 3 et f l'endomorphisme de E de matrice*

$$M = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

dans une base (e_1, e_2, e_3) de E. Son polynôme caractéristique est $\chi_f(X) = (2 - X)^2(3 - X)$, de sorte que son polynôme minimal est $m_f(X) = (X - 2)^k(X - 3)$ avec $k = 1$ ou $k = 2$. Par le calcul on vérifie que la matrice $(M - 2\mathrm{I}_3)(M - 3I_3)$ n'est pas nulle, donc $m_f(X) = (X - 2)^2(X - 3)$. Comme les racines de

m_f ne sont pas toutes simples, f n'est pas diagonalisable. Comme 3 est une valeur propre simple, E_3 est un sous-espace de dimension 1, et d'après la forme de la matrice on peut dire qu'il est engendré par e_3. L'endomorphisme f n'étant pas diagonalisable, la dimension de E_3 nous permet d'affirmer que E_2 est de dimension 1, et d'après M qu'il est engendré par e_1. Pour déterminer N_2 il faut résoudre $(M-2\mathrm{I}_3)^2v=0$ et on trouve que c'est le plan engendré par e_1 et e_2. D'autre part, on a $N_3=E_3$, car 3 est une valeur propre simple. On peut remarquer que (e_1,e_2,e_3) est une base de E, et par conséquent que $E=N_2\oplus N_3$. Ce résultat va être montré de manière générale dans la suite.

IV.2. Comparaison avec les sous-espaces propres.

Proposition 49. *Le sous-espace propre E_{λ_i} associé à la valeur propre λ_i est inclus dans le sous-espace caractéristique N_{λ_i}.*

DÉMONSTRATION. Soit $x\in E_{\lambda_i}$, c'est-à-dire tel que $(f-\lambda_i\mathrm{Id}_E)(x)=0$. On a

$$(f-\lambda_i\mathrm{Id}_E)^{\alpha_i}(x)=(f-\lambda_i\mathrm{Id}_E)^{\alpha_i-1}\big((f-\lambda_i\mathrm{Id}_E)(x)\big)=0.$$

Ainsi x appartient à N_{λ_i}. □

En particulier ceci entraîne que la dimension de N_{λ_i} est toujours supérieure ou égale à 1, car celle de E_{λ_i} l'est.

IV.3. Stabilité des sous-espaces caractéristiques.

Proposition 50. *Les sous-espaces caractéristiques sont stables par f.*

DÉMONSTRATION. Soit $x\in N_{\lambda_i}=\mathrm{Ker}\big[(f-\lambda_i\mathrm{Id}_E)^{\alpha_i}\big]$, c'est-à-dire tel que $\big[(f-\lambda_i\mathrm{Id}_E)^{\alpha_i}\big](x)=0$. On a alors

$$\big[(f-\lambda_i\mathrm{Id}_E)_i^{\alpha}\big]f(x)=f\big(\big[(f-\lambda_i\mathrm{Id}_E)^{\alpha_i}\big](x)\big)=0.$$

Donc $f(x)\in N_{\lambda_i}$. □

IV.4. Théorème de décomposition en sous-espaces caractéristiques.

Théorème 51.

On a

$$E=N_{\lambda_1}\oplus\cdots\oplus N_{\lambda_r}.$$

DÉMONSTRATION. On a

$$\chi_f(X)=(-1)^n\prod_{i=1}^{r}(X-\lambda_i)^{\alpha_i}.$$

D'après le Théorème de Cayley-Hamilton, $\chi_f(f)=0$, c'est-à-dire que $E=\mathrm{Ker}\big[\chi_f(f)\big]$. De plus si $i\neq j$, alors $(X-\lambda_i)^{\alpha_i}$ et $(X-\lambda_j)^{\alpha_j}$ sont des polynômes premiers entre eux. Donc le corollaire 26 entraîne que

$$\mathrm{Ker}\big[\chi_f(f)\big]=N_{\lambda_1}\oplus\cdots\oplus N_{\lambda_r},$$

et donc le résultat. □

IV.5. Autre définition des sous-espaces caractéristiques.

Proposition 52. *Pour tout $1 \leq i \leq r$, on a $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$, où β_i est l'ordre de multiplicité de la racine λ_i dans le polynôme minimal.*

Démonstration. On observe les faits suivants :

— Comme $\beta_i \leq \alpha_i$ la même démonstration que pour la proposition 31 montre que :

$$\mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big] \subset \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big].$$

— On a $m_f(f) = 0$, c'est-à-dire $E = \mathrm{Ker}\big[m_f(f)\big]$. De plus, si $i \neq j$, alors $(X - \lambda_i)^{\beta_i}$ et $(X - \lambda_j)^{\beta_j}$ sont des polynômes premiers entre eux. Donc par le corollaire 26 on obtient

$$E = \mathrm{Ker}\big[m_f(f)\big] = \mathrm{Ker}\big[(f - \lambda_1 \mathrm{Id}_E)^{\beta_1}\big] \oplus \cdots \oplus \mathrm{Ker}\big[(f - \lambda_r \mathrm{Id}_E)^{\beta_r}\big].$$

— La même démarche appliquée au polynôme caractéristique montre que

$$E = \mathrm{Ker}\big[\chi_f(f)\big] = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}.$$

En combinant ceci on déduit que $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$ pour tout $1 \leq i \leq r$.

□

IV.6. Nouveau théorème de diagonalisation.

Théorème 53.

f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :
(i) f est scindé.
(ii) Pour toute valeur propre λ_i de f on a $N_{\lambda_i} = E_{\lambda_i}$.

Démonstration. On a déjà vu que si f est diagonalisable alors f est scindé et de plus on a

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}.$$

D'autre part, f étant scindé, d'après le théorème 34 on a aussi $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$ et d'après la proposition 31, $E_{\lambda_i} \subset N_{\lambda_i}$. Tout ceci entraîne que $E_{\lambda_i} = N_{\lambda_i}$.

Réciproquement, si f est scindé d'après le théorème 34, $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$. Ainsi comme $E_{\lambda_i} = N_{\lambda_i}$, $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

IV.7. Applications linéaires restreintes.

Proposition 54. *La restriction $f_i = f|_{N_{\lambda_i}}$ de f au sous-espace caractéristique N_{λ_i} est un endomorphisme de N_{λ_i}. Les polynômes minimal et caractéristique de f_i sont*

$$m_{f_i}(X) = (X - \lambda_i)^{\beta_i} \text{ et } \chi_{f_i}(X) = (-1)^{\alpha_i}(X - \lambda_i)^{\alpha_i}.$$

De plus on a $\dim(N_{\lambda_i}) = \alpha_i$.

Démonstration. La restriction de f à N_{λ_i} est une application linéaire de N_{λ_i} dans $f\big(N_{\lambda_i}\big)$. Or d'après la proposition 33, N_{λ_i} est stable par f et donc par sa restriction. On en déduit que $f_i \in \mathcal{L}\big(N_{\lambda_i}\big)$.

On sait que

$$N_{\lambda_i} = \mathrm{Ker}\big[(f-\lambda_i \mathrm{Id}_E)^{\alpha_i}\big] = \mathrm{Ker}\big[(f-\lambda_i \mathrm{Id}_E)^{\beta_i}\big].$$

On en déduit donc $(f_i - \lambda_i \mathrm{Id}_{N_{\lambda_i}})^{\beta_i} = 0$, c'est-à-dire que le polynôme minimal de f_i divise le polynôme $(X-\lambda_i)^{\beta_i}$. Supposons $\deg(m_{f_i}) = \gamma_i < \beta_i$, alors le polynôme

$$Q(X) = (X-\lambda_i)^{\gamma_i} \prod_{j \neq i} (X-\lambda_j)^{\beta_j}$$

s'annule en f et

$$\gamma_i + \sum_{j\neq i} \beta_j < \sum_{j=1}^{r} \beta_j = \deg(m_f).$$

Ceci est impossible car le polynôme minimal de f est le polynôme unitaire de plus bas degré vérifiant cette propriété. Ainsi on a $\gamma_i = \beta_i$ et $m_{f_i}(X) = (X-\lambda_i)^{\beta_i}$.

Supposons que $\dim(N_{\lambda_i}) = \gamma_i$. Comme E est la somme directe des N_{λ_i} il existe une base $\mathcal{B}$ de E telle que $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, où $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\gamma_i,i})$ est une base de N_{λ_i}. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}$$

où les M_i sont les matrices des f_i dans les bases $\mathcal{B}_i$. Or le polynôme minimal de f_i est $m_{f_i}(X) = (X-\lambda_i)^{\beta_i}$. Donc la seule valeur propre de M_i (c'est-à-dire de f_i) est λ_i et son polynôme caractéristique est égal à $(-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i}$. On obtient donc l'égalité suivante :

$$\begin{aligned} (-1)^n \prod_{i=1}^{r} (X-\lambda_i)^{\alpha_i} &= \chi_f(X) = \prod_{i=1}^{r} \det(M_i - X\mathrm{I}_{\gamma_i}) = \prod_{i=1}^{r} \chi_{f_i}(X) = \prod_{i=1}^{r} (-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i} \\ &= (-1)^n \prod_{i=1}^{r} (X-\lambda_i)^{\gamma_i} \end{aligned}$$

car

$$\sum_{i=1}^{r} \gamma_i = \sum_{i=1}^{r} \dim(N_{\lambda_i}) = \dim(E) = n.$$

On en déduit que $\gamma_i = \alpha_i$, c'est-à-dire que $\dim(N_{\lambda_i}) = \alpha_i$. De plus on obtient que $\chi_{f_i}(X) = (-1)^{\alpha_i}(X-\lambda_i)^{\alpha_i}$. □

IV.8. Trigonalisation des matrices en blocs relatifs aux sous-espaces caractéristiques.

On a vu que si on prend $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ avec $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\alpha_i,i})$ une base de N_{λ_i}, alors la matrice de f dans $\mathcal{B}$ s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}.$$

Pour trigonaliser M il suffit alors de trigonaliser chaque M_i.

Exemple : On veut trigonaliser l'endomorphisme f dont la matrice dans la base canonique est
$$\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}.$$
On commence par chercher le polynôme caractéristique de f. On obtient
$$\chi_f(X) = (X-1)^2(X+1)^2.$$
f a donc deux valeurs propres doubles, 1 et -1. On détermine alors les sous-espaces caractéristiques.
Pour $\lambda = -1$, on obtient :
$$N_{-1} = \mathrm{Ker}[(f + 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M + \mathrm{I}_4)^2].$$
On a donc
$$(M+\mathrm{I}_4)^2 \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} 12z - 8t & = & 0 \\ 8z - 4t & = & 0 \\ 12z - 8t & = & 0 \\ 8z - 4t & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} 3z & = & 4t \\ 2z & = & t \end{array} \right.$$
Une base de N_{-1} est donc donnée par u_1 et u_2 de coordonnées dans la base canonique
$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$
Pour $\lambda = 1$, $N_1 = \mathrm{Ker}[(f - 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M - \mathrm{I}_4)^2]$. On a donc
$$(M-\mathrm{I}_4)^2 \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{l} -12x + 16y + 12z - 16t = 0 \\ -16x + 20y + 16z - 20t = 0 \end{array} \right. \iff \left\{ \begin{array}{l} x = z \\ y = t \end{array} \right. .$$
Une base de N_1 est donc donnée par u_3 et u_4 de coordonnées dans la base canonique $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

La matrice de passage de la base canonique à la nouvelle base (u_1, u_2, u_3, u_4), ainsi que son inverse, sont données par
$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
Donc la matrice de f dans la nouvelle base s'écrit :
$$\widetilde{M} = P^{-1}MP = \begin{pmatrix} 3 & -4 & 0 & 0 \\ 4 & -5 & 0 & 0 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$
On trigonalise alors chacune des sous matrices
$$M_1 = \begin{pmatrix} 3 & -4 \\ 4 & -5 \end{pmatrix} \quad \text{et} \quad M_2 = \begin{pmatrix} 3 & -2 \\ 2 & 1 \end{pmatrix}.$$

Le polynôme caractéristique de M_1 est égal à $(X+1)^2$. Le sous-espace propre de M_1 associé à la valeur propre -1 est déterminé par

$$(M_1 + \mathrm{I}_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} 4x_1 - 4x_2 & = & 0 \\ 4x_1 - 4x_2 & = & 0 \end{array} \right. \iff x_1 = x_2.$$

Une base de cet espace est donc $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ dans (u_1, u_2), i.e.

$$v_1 = u_1 + u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

On complète par un second vecteur de N_{-1} non colinéaire à v_1, par exemple $v_2 = u_1$.
Le polynôme caractéristique de M_2 est égal à $(X-1)^2$. Le sous-espace propre de M_2 associé à la valeur propre 1 est déterminé par

$$(M_2 - \mathrm{I}_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} 2x_3 - 2x_4 & = & 0 \\ 2x_3 - 2x_4 & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} x_3 & = & x_4 \end{array} \right.$$

Une base de cet espace est donc $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ dans (u_3, u_4), i.e.

$$v_3 = u_3 + u_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

On complète par un second vecteur de N_1 non colinéaire à v_3, par exemple $v_4 = u_3$.
La matrice de passage de la base (u_1, u_2, u_3, u_4) à la base (v_1, v_2, v_3, v_4) et son inverse, sont égales à

$$P' = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad P'^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Donc la matrice de f dans la base (v_1, v_2, v_3, v_4) s'écrit :

$$M' = (P')^{-1}\widetilde{M}(P') = \begin{pmatrix} -1 & 4 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

V. Endomorphismes nilpotents

V.1. Caractérisation des endomorphismes nilpotents.

Définition 55. *On dit qu'un endomorphisme f (respectivement une matrice M) est* nilpotent *(respectivement* nilpotente*) si et seulement si il existe un entier $k \geq 1$ tel que $f^k = 0$, l'endomorphisme identiquement nul, (respectivement $M^k = 0$, la matrice nulle). Le plus petit entier $k \geq 1$ vérifiant ceci, est alors appelé l'*indice de nilpotence *de f (respectivement de M).*

Proposition 56. *Un endomorphisme d'un espace vectoriel de dimension n est nilpotent si et seulement si son polynôme caractéristique est $(-1)^n X^n$.*

DÉMONSTRATION. Si f est un endomorphisme nilpotent, alors il existe un entier $k \geq 1$ tel que $f^k = 0$, c'est-à-dire que le polynôme X^k appartient à l'idéal annulateur de f. Donc par définition, le polynôme minimal de f divise X^k ; autrement dit il est de la forme X^j avec j un entier tel que $1 \leq j \leq k$. L'unique racine de ce polynôme est donc 0, et donc l'unique valeur propre de f est 0. Comme l'espace vectoriel est de dimension n, le polynôme caractéristique de f est donc $(-1)^n X^n$. Réciproquement, si le polynôme caractéristique de f est $(-1)^n X^n$, d'après le Théorème de Cayley-Hamilton $\chi_f(f) = f^n = 0$, c'est-à-dire que f est nilpotent. □

Une conséquence de ce résultat est que le polynôme minimal d'un endomorphisme nilpotent est égal à X^p avec $p \leq n$ (ceci montre au passage que l'indice de nilpotence est inférieur ou égal à la dimension de E). D'autre part, comme il n'a qu'une seule valeur propre qui est zéro, il n'est pas diagonalisable sauf s'il est nul. Mais il est trigonalisable.

V.2. Décomposition "diagonale + nilpotent" ou décomposition de Dunford.

Théorème 57.

Soit f un endomorphisme d'un espace vectoriel E. Si f est scindé alors il existe un unique couple (u, v) d'endomorphismes tel que
- $f = u + v$.
- *u est diagonalisable et v est nilpotent.*
- *u et v commutent.*

DÉMONSTRATION. Montrons que le couple (u, v) existe. Comme f est scindé d'après le théorème 34 on a $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$, où les $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$ sont les sous-espaces caractéristiques de f correspondant aux différentes valeurs propres $(\lambda_i)_{1 \leq i \leq r}$ de f. On définit alors u et v par leurs restrictions aux sous-espaces caractéristiques de la manière suivante :

$$u_{|N_{\lambda_i}} = \lambda_i \mathrm{Id}_{N_{\lambda_i}} \text{ et } v_{|N_{\lambda_i}} = f_{|N_{\lambda_i}} - \lambda_i \mathrm{Id}_{N_{\lambda_i}}.$$

Comme les sous-espaces caractéristiques sont supplémentaires et que chaque restriction de u est diagonale, u l'est aussi. D'autre part, par définition chaque restriction de v est nilpotente d'ordre $n_i \leq \beta_i$ et donc v sera aussi nilpotente d'ordre $\max\limits_{1 \leq i \leq r}(n_i)$. Enfin il est clair que pour tout $1 \leq i \leq r$, $u_{|N_{\lambda_i}}$ commute avec $v_{|N_{\lambda_i}}$ car l'identité commute avec tout endomorphisme, donc u et v commutent. Nous admettons l'unicité de la décomposition. □

Exemple 58. *Soit la matrice*

$$M = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}.$$

On vérifie par le calcul que son polynôme caractéristique est $\chi_M(X) = -X^2(X+1)$. Les valeurs propres sont donc : -1 valeur propre simple et 0 valeur propre double. Le sous-espace caractéristique associé à la valeur propre -1, est égal au sous-espace propre associé à cette même valeur propre. Pour le déterminer on résout $(M + \mathrm{I}_3)v = 0$. On trouve qu'il est engendré par $e_1 = (1, -1, -2)$. Pour le sous-espace caractéristique associé à la valeur propre 0 on résout $(M - 0.\mathrm{I}_3)^2 v = M^2 v = 0$ et on constate qu'il est engendré par $e_2 = (1, 2, 0)$ et $e_2 = (0, 1, 1)$. Nous allons maintenant écrire la matrice dans la nouvelle base. La matrice de passage de la base canonique de $\mathbb{R}^3$ à la base (e_1, e_2, e_3) est

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \text{ et son inverse } P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}.$$

On en déduit que

$$\widetilde{M} = P^{-1}MP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \widetilde{D} + \widetilde{N}.$$

La décomposition de Dunford de M est donc $M = D + N$ avec

$$D = P\widetilde{D}P^{-1} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix} \quad et \quad N = P\widetilde{N}P^{-1} = M - D = \begin{pmatrix} 4 & -2 & 3 \\ 8 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}.$$

On peut vérifier que $DN = ND$.

V.3. Décomposition de Jordan.

Définition 59. *On appelle* bloc de Jordan *de taille p et de valeur propre λ une matrice carrée d'ordre p de la forme*

$$J_{\lambda,p} = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \ldots & 0 & \lambda \end{pmatrix}.$$

Théorème 60.

Soit f un endomorphisme scindé de E. On suppose que le polynôme caractéristique de f est

$$(-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}, \qquad avec \ \sum_{i=1}^{r} \alpha_i = n,$$

où les λ_i, $1 \leq i \leq r$ sont deux à deux distincts. Il existe alors une base de E dans laquelle la matrice de f est diagonale par blocs de la forme :

$$\begin{pmatrix} J_{\lambda_1,n_{1,1}} & 0 & \cdots & 0 \\ 0 & J_{\lambda_1,n_{1,2}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & & J_{\lambda_r,n_{r,s_r}} \end{pmatrix},$$

où pour $1 \leq i \leq r$ et $1 \leq j \leq s_i$, le bloc diagonal $J_{\lambda_i,n_{i,j}}$ est un bloc de Jordan de taille $n_{i,j}$ de valeur propre λ_i. Pour $1 \leq i \leq r$, on a $\sum_{j=1}^{s_i} n_{i,j} = \alpha_i$, et on impose $n_{i,1} \geq n_{i,2} \geq \cdots \geq n_{i,s_i}$. L'entier $n_{i,1}$ est égal à la multiplicité de λ_i comme racine du polynôme minimal.

On peut remarquer que la matrice de Jordan donnée par le théorème est triangulaire supérieure puisque les blocs de Jordan le sont. Sur la diagonale figurent les valeurs propres, avec la multiplicité qu'elles ont dans le polynôme caractéristique, et les coefficients de la diagonale secondaire $\{(i,j), j = i+1\}$ sont égaux à 1 ou 0. Tous les autres coefficients sont nuls.
Les valeurs propres étant données et indexées, la matrice de Jordan est entièrement déterminée par la donnée des entiers $n_{i,1}, \ n_{i,2}, \ \ldots, \ n_{i,s_i}$. Avec ces hypothèses, elle est unique et est appelée *forme normale de Jordan.*

DÉMONSTRATION. En procédant comme dans la démonstration du théorème de la décomposition de Dunford, f étant scindé, on peut écrire pour chaque restriction f_i à un sous-espace caractéristique N_{λ_i}, $f_i = u_i + v_i$ avec $u_i = \lambda_i \mathrm{Id}_{N_{\lambda_i}}$ et v_i nilpotente. Pour montrer le résultat il suffit de démontrer que l'on peut mettre chaque v_i sous la forme voulue. On va utiliser le lemme suivant :

Lemme 61. *Soit v un endomorphisme nilpotent d'indice q sur E de dimension n. Alors on a :*
$$\{0\} = \mathrm{Ker}(v^0) \subsetneq \mathrm{Ker}(v^1) \subsetneq \cdots \subsetneq \mathrm{Ker}(v^q) = E$$

DÉMONSTRATION. On a $\mathrm{Ker}(v^0) = \mathrm{Ker}(\mathrm{Id_E}) = \{0\}$ et $\mathrm{Ker}(v^q) = E$.

On doit montrer que

i) $x \in \mathrm{Ker}(v^p) \Rightarrow x \in \mathrm{Ker}(v^{p+1})$ pour tout $1 \leq p \leq q-1$.

ii) $\mathrm{Ker}(v^p) \neq \mathrm{Ker}(v^{p+1})$ pour tout $1 \leq p \leq q-1$.

Si $x \in \mathrm{Ker}(v^p)$ alors $v^p(x) = 0$ et par conséquent
$$v^{p+1}(x) = v(v^p(x)) = f(0) = 0,$$
et donc on obtient le premier résultat.

Pour le second supposons par l'absurde que
$$\mathrm{Ker}(v^p) = \mathrm{Ker}(v^{p+1})$$
pour un certain $1 \leq p \leq q-1$. Soit $x \in \mathrm{Ker}(v^{p+2})$, on a $v^{p+2}(x) = 0$, c'est-à-dire $v^{p+1}(v(x)) = 0$. Donc $v(x)$ appartient à $\mathrm{Ker}(v^{p+1})$ et par conséquent à $\mathrm{Ker}(v^p)$. Ainsi, $v^{p+1}(x) = v^p(v(x)) = 0$ c'est-à-dire que $x \in \mathrm{Ker}(v^{p+1})$. On en déduit que
$$\mathrm{Ker}(v^{p+2}) = \mathrm{Ker}(v^{p+1}).$$

On montre ainsi par récurrence que
$$\mathrm{Ker}(v^p) = \mathrm{Ker}(v^{p+1}) = \cdots = \mathrm{Ker}(v^q) = E,$$
et donc que v est nilpotente d'indice inférieur à q, ce qui est contraire aux hypothèses. □

Le résultat de ce lemme montre en particulier que $\dim \mathrm{Ker}(v^i) \geq i$ pour $1 \leq i \leq q$. On va construire par récurrence descendante sur $i \in \{1, \ldots q\}$ des sous-espaces vectoriels F_i de E tels que
$$\mathrm{Ker}(v^i) = \mathrm{Ker}(v^{i-1}) \oplus F_i$$
et vérifiant $v(F_i) \subset F_{i-1}$.
La construction que nous allons mettre en œuvre va entraîner que :
$$\begin{array}{lclcl}
\mathrm{Ker}(v^i) & = & \mathrm{Ker}(v^{i-1}) & \oplus & F_i \\
v\downarrow & & v\downarrow & & v\downarrow \\
\mathrm{Ker}(v^{i-1}) & = & \mathrm{Ker}(v^{i-2}) & \oplus & F_{i-1}
\end{array}$$
ainsi que
$$E = \underbrace{F_1 \oplus F_2 \oplus \cdots \oplus F_i}_{\mathrm{Ker}(v^i)} \oplus F_{i+1} \oplus \cdots \oplus F_q.$$
La construction se fait de la manière suivante. On prend pour F_q n'importe quel supplémentaire de $\mathrm{Ker}(v^{q-1})$ dans E. Si F_i $(i \geq 2)$ est construit, on remaque les résultats suivants. Comme $v(\mathrm{Ker}(v^i)) \subset \mathrm{Ker}(v^{i-1})$ car si $x \in v(\mathrm{Ker}(v^i))$, il existe $y \in \mathrm{Ker}(v^i)$ tel que $x = v(y)$ et donc
$$v^{i-1}(x) = v^{i-1}(v(y)) = v^i(y) = 0.$$
Ainsi, on obtient le premier résultat :
$$v(F_i) \subset \mathrm{Ker}(v^{i-1}).$$
D'autre part, par construction on a $F_i \cap \mathrm{Ker}(v^{i-1}) = \{0\}$ et donc
$$v(F_i) \cap \mathrm{Ker}(v^{i-2}) = \{0\}.$$
En effet s'il existait $x \neq 0 \in v(F_i) \cap \mathrm{Ker}(v^{i-2})$ cela entrainerait qu'il existe $y \neq 0 \in F_i$ tel que $v(y) = x \in v(F_i)$. De plus
$$0 = v^{i-2}(x) = v^{i-2}(v(y)) = v^{i-1}(y),$$
c'est-à-dire que $y \in \mathrm{Ker}(v^{i-1})$. On aurait donc trouvé $y \neq 0$ dans $F_i \cap \mathrm{Ker}(v^{i-1})$ ce qui contredit les hypothèses.

D'après ces deux résultats il suffit donc de prendre pour F_{i-1} un supplémentaire de $\mathrm{Ker}(v^{i-2})$ dans $\mathrm{Ker}(v^{i-1})$ contenant $v(F_i)$. D'autre part, la restriction de v à F_i est injective pour $i \geq 2$ puisque $F_i \cap \mathrm{Ker}(v) = \{0\}$.

On construit alors une base de E de la façon suivante

$$\begin{array}{ccccccc}
F_q & \overset{v}{\hookrightarrow} & F_{q-1} & \overset{v}{\hookrightarrow} & \dots & \overset{v}{\hookrightarrow} & F_1 \\
e_{q,1} & & v(e_{q,1}) & & & & v^{q-1}(e_{q,1}) \\
\vdots & & \vdots & & & & \vdots \\
e_{q,s_q} & & v(e_{q,s_q}) & & & & v^{q-1}(e_{q,s_q}) \\
 & & e_{q-1,s_q+1} & & & & v^{q-2}(e_{q-1,s_q+1}) \\
 & & \vdots & & & & \vdots \\
 & & e_{q-1,s_{q-1}} & & & & v^{q-2}(e_{q-1,s_{q-1}}) \\
 & & & & & & \vdots \\
 & & & & & & e_{1,s_2+1} \\
 & & & & & & \vdots \\
 & & & & & & e_{1,s_1}
\end{array}$$

avec $s_i = \dim F_i$. La famille de vecteurs qui apparaissent ci-dessus forme une base de E que l'on ordonne en partant du coin supérieur droit avec $v^{q-1}(e_{q,1})$ en lisant la première ligne vers la gauche, puis la seconde ligne aussi vers la gauche et ainsi de suite jusqu'à la dernière ligne, qui est le seul vecteur e_{1,s_1}. La matrice de v dans cette base est diagonale par blocs (de blocs $J_{0,k}$) avec s_q blocs de type $J_{0,q}$, puis $s_{q-1} - s_q$ blocs de type $J_{0,q-1}$ et ainsi de suite jusqu'à $m_1 - m_2$ bloc de type $J_{0,1}$ (c'est à dire la matrice carrée d'ordre 1 nulle !) $\square$

Chapitre 3

FORMES BILINÉAIRES et QUADRATIQUES

On fixe un corps K, en général $\mathbb{R}$ ou $\mathbb{C}$. Tous les espaces vectoriels considérés seront des espaces vectoriels sur K de dimension finie.

I. Rappel sur les formes linéaires

Définition 1. *Soit E un K- espace vectoriel. On appelle* forme linéaire *sur E une application linéaire f de E dans K (K étant considéré comme un espace vectoriel sur lui-même).*

Proposition 2. *Soient E un K-espace vectoriel de dimension n et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Une application f de E dans K est une forme linéaire si et seulement si il existe n scalaires $a_1, \ldots, a_n$ tels que pour tout $x = x_1e_1 + \cdots + x_ne_n$,*

$$f(x) = a_1x_1 + \cdots + a_nx_n.$$

Ceci se vérifie facilement en remarquant que $a_i = f(e_i)$.

Remarque 3. *L'expression ci-dessus est une expression polynomiale homogène de degré 1 par rapport aux coordonnées de x dans la base $\mathcal{B}$ (ce résultat pourra être rapproché d'une caractérisation des formes quadratiques sur un espace de type fini).*

Exemple 4. *Si E est l'espace des polynômes P de degré $\leq n$ et à coefficients dans K, et si a est un élément de K alors l'application $f : P \mapsto P(a)$ est une forme linéaire sur E.*

Définition 5. *Soit E un espace vectoriel. On appelle* espace dual *(ou simplement* dual*) de E l'espace vectoriel des formes linéaires sur E, c'est-à-dire $\mathcal{L}(E, K)$, et on le note E^*.*

Définition 6. *Soit E un espace vectoriel. On appelle* hyperplan *de E un sous-espace vectoriel qui est le noyau d'une forme linéaire non nulle.*

Proposition 7. *Une forme linéaire non nulle a pour noyau un sous-espace vectoriel de dimension $n - 1$ de E. Réciproquement, tout sous-espace vectoriel H de dimension $n - 1$ de E est le noyau d'au moins une forme linéaire sur E, qu'on appellera équation de H. L'ensemble des formes linéaires nulles sur H est un sous-espace vectoriel de E^* de dimension 1.*

Démonstration. Une forme linéaire est de rang 0 si elle est nulle ou de rang 1 sinon. Dans ce dernier cas, son noyau (thèorème du rang) est un sous-espace vectoriel de E de dimension $n - 1$. Réciproquement, si H est un tel sous-espace, on en choisit une base que l'on complète en une base $\mathcal{B}$ de E par un vecteur e (il suffit pour cela que $e \notin H$). Il existe une unique forme linéaire qui vaut 1 sur e et 0 sur H. Comme elle n'est pas nulle son noyau est de dimension $n - 1$ et contient H, donc c'est H.
Soit g une autre forme linéaire nulle sur H. On a $g = g(e)f$. En effet cette égalité est vraie pour tous les vecteurs de $\mathcal{B}$, donc sur E par linéarité. □

Définition 8. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Pour tout $j = 1, \ldots, n$ on définit une forme linéaire e_j^* sur E en posant*

$$\textit{pour } 1 \leq k \leq n, \quad e_j^*(e_k) = \delta_{jk} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

δ_{jk} *est appelé le* symbole de Kronecker.

Proposition 9. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. La base $\mathcal{B}^* = (e_1^*, \ldots, e_n^*)$ de E^* définie comme ci-dessus est une base de E^*.*

DÉMONSTRATION. Si la combinaison linèaire de formes linèaires

$$L := \sum_{j=1}^{n} \lambda_j e_j^*$$

est nulle, où les λ_j sont des rèels, alors on obtient en l'appliquant à e_k, pour tout $1 \leq k \leq n$,

$$0 = L(e_k) = \sum_{j=1}^{n} \lambda_j e_j^*(e_k) = \sum_{j=1}^{n} \lambda_j \delta_{jk} = \lambda_k.$$

En reportant dans la dèfinition de L on obtient donc que L est nulle. Ceci montre que la famille $\mathcal{B}^*$ est libre.
Soit l une forme linéaire quelconque sur E. La forme linéaire

$$\tilde{l} := \sum_{j=1}^{n} l(e_j) e_j^*$$

vérifie pour tout $1 \leq k \leq n$, Ainsi les formes linéaires l et $\tilde{l}$ prennent les mêmes valeurs sur une base de E, elles sont donc égales. Ceci montre que la famille $\mathcal{B}^*$ est génératrice. Donc $\mathcal{B}^*$ est une base de E^*. $\square$

Définition 10. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. La base $\mathcal{B}^* = (e_1^*, \ldots, e_n^*)$ de E^* définie comme ci-dessus est appelée* base duale *de B.*

Remarquons que $\mathcal{B}$ et $\mathcal{B}^*$ ont le même nombre d'éléments de sorte que E et E^* ont la même dimension.

II. Formes bilinéaires

Définition 11. *Soit φ une application de $E \times E$ dans K qui à (x, y) fait correspondre $\varphi(x, y)$. On dit que φ est* une forme bilinéaire *sur E si* les applications partielles *$\varphi_x \; : y \mapsto \varphi(x, y)$, où x est fixé, et $\varphi_y \; : x \mapsto \varphi(x, y)$, où y est fixé, sont des formes linéaires sur E.*

Remarque 12. *La propriété de la définition se traduit par la condition :*

$$\forall (x, x') \in E \times E, \quad \forall (y, y') \in E \times E, \quad \forall (\lambda, \mu) \in K \times K, \quad \forall (\lambda', \mu') \in K \times K,$$
$$\varphi(\lambda x + \mu y, \lambda' x' + \mu' y') = \lambda\lambda' \varphi(x, x') + \lambda\mu' \varphi(x, y') + \mu\lambda' \varphi(y, x') + \mu\mu' \varphi(y, y'),$$

ou encore par : pour tout x, y et z dans E et tout λ dans K,

$$\begin{array}{rcll} \varphi(\lambda x, y) & = & \lambda\varphi(x, y), & \varphi(x, \lambda y) = \lambda\varphi(x, y), \\ \varphi(x + z, y) & = & \varphi(x, y) + \varphi(z, y), & \varphi(x, y + z) = \varphi(x, y) + \varphi(x, z). \end{array}$$

Les formes bilinéaires sur $E \times E$ constituent un espace vectoriel sur K, noté $\mathcal{L}(E, E, K)$ ou bien $\mathcal{L}^2(E)$, de dimension $(\dim(E))^2 = n^2$.
Attention, en général une application bilinéaire n'est pas une application linéaire.

Exemple 13. *La relation $\varphi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1 y_1 + \cdots + x_n y_n$ définit une forme bilinéaire sur K^n (à vérifier en exercice).*

A une application linéaire il est toujours possible d'associer une matrice dans une base. Voyons ce qu'il en est pour une forme bilinéaire. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base d'un K-espace vectoriel E, et soient

$$x = \sum_{i=1}^{n} x_i e_i \text{ et } y = \sum_{j=1}^{n} y_j e_j$$

deux éléments de E. On a

$$\varphi(x,y)=\varphi\left(\sum_{i=1}^{n}x_ie_i,\sum_{j=1}^{n}y_je_j\right)=\sum_{1\leq i,j\leq n}\varphi(e_i,e_j)x_iy_j.$$

Ainsi, la forme bilinéaire φ est déterminée de façon unique par la matrice suivante dans la base $\mathcal{B}$:

$$M_{\varphi,\mathcal{B}}=M_{\mathcal{B}}=(\varphi(e_i,e_j))_{1\leq i,j\leq n}$$

.

Si on note

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\text{ et }Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

les matrices colonnes des vecteurs x et y dans la base $\mathcal{B}$, alors on a

$$\varphi(x,y)={}^tXM_{\mathcal{B}}Y.$$

Nous allons maintenant voir comment se traduit un changement de bases, pour la matrice d'une forme bilinéaire. Considérons maintenant $\mathcal{B}'=(e'_1,\ldots,e'_n)$ une autre base de E et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, définie par

$$e'_j=\sum_{i=1}^{n}p_{ij}e_i,$$

pour tout $1\leq j\leq n$. Notons encore X et X' les matrices colonnes du vecteur x dans les bases $\mathcal{B}$ et $\mathcal{B}'$ (on a $X=PX'$), et Y et Y' les matrices colonnes du vecteur y dans $\mathcal{B}$ et $\mathcal{B}'$. On obtient alors

$$\varphi(x,y)={}^tXM_{\mathcal{B}}Y={}^t(PX')M_{\mathcal{B}}PY'={}^tX'\,{}^tPM_{\mathcal{B}}PY'={}^tX'M_{\mathcal{B}'}Y';$$

de sorte que la matrice de φ dans la base $\mathcal{B}'$ s'écrit :

$$M_{\mathcal{B}'}={}^tPM_{\mathcal{B}}P.$$

Remarque 14. *Attention à na pas confondre avec la formule de changement de base pour une application linéaire : $M_{\mathcal{B}'}=P^{-1}M_{\mathcal{B}}P$. On constate en particulier que le déterminant de la matrice $M_{\mathcal{B}}$ dépend de la base choisie.*

Ce qui précède nous donne la caractérisation suivante d'une forme bilinéaire sur un espace vectoriel E de dimension n :

Proposition 15. *Soit $\mathcal{B}=(e_1,\ldots,e_n)$ une base de E. Une application f de $E\times E$ dans K est une forme bilinéaire sur E si et seulement si il existe des scalaires a_{ij}, pour $1\leq i\leq n$ et $1\leq j\leq n$, tels que pour tout*

$$x=x_1e_1+\cdots+x_ne_n$$

et tout

$$y=y_1e_1+\cdots+y_ne_n,$$

$f(x,y)$ s'écrive de la manière suivante :

$$f(x,y)=\sum_{1\leq i,j\leq n}a_{ij}x_iy_j.$$

Définition 16. *On dit que deux matrices carrées d'ordre n à coefficients dans K, M et M', sont* congruentes, *s'il existe une matrice inversible P telle que $M'={}^tPMP$.*

Remarque 17. *La relation de congruence dans l'ensemble des matrices carrées d'ordre n à coefficients dans K est une relation d'équivalence dans laquelle toutes les matrices d'une même classe ont même rang. On peut donc définir le rang d'une forme bilinéaire de la manière suivante.*

Définition 18. *Le* rang d'une forme bilinéaire *est le rang de sa matrice dans une base quelconque de E.*

Terminons par une dernière définition qui caractérise des formes bilinéaires particulières.

Définition 19. *On dit qu'une forme bilinéaire φ est* définie *si*

$$\forall x \in E, \quad \varphi(x,x) = 0 \iff x = 0.$$

Définition 20. *On dit qu'une forme bilinéaire φ est* positive *si*

$$\forall x \in E, \quad \varphi(x,x) \geq 0.$$

Exemple 21. *En reprenant l'exemple précédent, on remarque que la forme bilinéaire est définie positive.*

III. Formes bilinéaires symétriques et orthogonalité

III.1. Formes bilinéaires symétriques.

Définition 22. *On dit qu'une forme bilinéaire φ est* symétrique *si*

$$\forall (x,y) \in E \times E, \quad \varphi(x,y) = \varphi(y,x).$$

Exemple 23. *Le premier exemple de ce chapitre est une forme bilinéaire symétrique.*

Proposition 24. *Pour qu'une forme bilinéaire soit symétrique il faut et il suffit que sa matrice dans une base donnée soit symétrique (i.e. $a_{ij} = a_{ji}$).*

DÉMONSTRATION. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. La matrice de la forme bilinéaire φ dans cette base est $(\varphi(e_i,e_j))_{1 \leq i,j \leq n}$. Si cette matrice est symétrique on a pour tout $1 \leq i \leq n$ et $1 \leq j \leq n$, $\varphi(e_i,e_j) = \varphi(e_j,e_i)$.

Ceci est équivalent à, pour tout x et y des vecteurs de E,

$$\varphi(x,y) = \sum_{1 \leq i,j \leq n} \varphi(e_i,e_j) x_i y_j = \sum_{1 \leq i,j \leq n} \varphi(e_j,e_i) y_j x_i = \varphi(y,x),$$

c'est-à-dire que la forme bilinéaire φ est symétrique. $\square$

Remarque 25. *Les formes bilinéaires symétriques sur $E \times E$ constituent un sous-espace vectoriel de $\mathcal{L}(E,E,K)$, de dimension $\dfrac{n(n+1)}{2}$. Cela peut se montrer en utilisant le résultat précédant, et en remarquant que pour définir une matrice symétrique d'ordre n, il faut choisir $\dfrac{n(n+1)}{2}$ coefficients.*

Définition 26. *Soit E un K-espace vectoriel. On appelle* noyau *d'une forme bilinéaire symétrique φ de E, le sous-espace vectoriel de E défini par*

$$\operatorname{Ker}(\varphi) = \{y \in E \mid \forall x \in E,\ \varphi(x,y) = 0\}.$$

Définition 27. *On dit qu'une forme bilinéaire symétrique φ est* non dégénérée *si son noyau est nul, i.e.* $\operatorname{Ker}(\varphi) = \{0\}$.

Exemple 28. *Reprenons la forme bilinéaire symétrique*

$$\varphi\left((x_1, \dots, x_n), (y_1, \dots, y_n)\right) = x_1 y_1 + \cdots + x_n y_n$$

sur $E = \mathbb{R}^n$*. Soit* $y \in \mathrm{Ker}(\varphi)$*. On a pour tout* $x \in E$*,* $\varphi(x, y) = 0$*. En particulier si* $x = y$*, on obtient*

$$\varphi(y, y) = y_1^2 + \cdots + y_n^2 = 0.$$

Une somme de carrés (de réels) est nulle, que si chacun de ses carrés est nul, c'est-à-dire que $y_1 = \cdots = y_n = 0$*. On en déduit que* $y = 0$ *et donc que le noyau de* φ *est* $\{0\}$*. Par conséquent,* φ *est une forme bilinéaire non dégénérée.*
Cet exemple nous permet d'affirmer que si une forme bilinéaire symétrique est définie alors elle est non dégénérée. La réciproque est en général fausse.

Exemple 29. *Soit l'application définie sur* $\mathbb{R}^2 \times \mathbb{R}^2$ *par,*

$$\varphi((x_1, x_2), (y_1, y_2)) = x_1 y_1.$$

On peut vérifier aisément que c'est une forme bilinéaire symétrique. Remarquons que $(x_1, x_2) = x_1(1, 0) + x_2(0, 1)$*, et donc que*

$$\varphi((x_1, x_2), (y_1, y_2)) = x_1\varphi((1, 0), (y_1, y_2)) + x_2\varphi((0, 1), (y_1, y_2)).$$

Par conséquent, on obtient

$$\begin{aligned} y \in \mathrm{Ker}(\varphi) &\iff \forall (x_1, x_2) \in \mathbb{R}^2,\ \varphi((x_1, x_2), (y_1, y_2)) = 0 \\ &\iff \varphi((1, 0), (y_1, y_2)) = \varphi((0, 1), (y_1, y_2)) = 0 \iff y_1 = 0 \end{aligned}$$

C'est-à-dire que $\mathrm{Ker}(\varphi) = \{(0, y_2),\ y_2 \in \mathbb{R}\}$*.*

Remarque 30. *En d'autres termes, une forme bilinéaire symétrique* φ *est non dégénérée si, pour tout vecteur non nul* x *dans* E*, il existe* y *dans* E *tel que* $\varphi(x, y) \neq 0$*. Du point de vue matriciel cela signifie que si* $M_{\mathcal{B}}$ *est la matrice de* φ *dans une base quelconque* $\mathcal{B}$ *de* E*, alors* $M_{\mathcal{B}}$ *est inversible ou encore* $\det(M_{\mathcal{B}}) \neq 0$*.*

Remarque 31. *L'exemple ci-dessus nous donne une méthode pour déterminer le noyau d'une forme bilinéaire symétrique. En effet si* φ *est une forme bilinéaire symétrique sur un* K*-espace vectoriel* E*, et si* $(e_1, \ldots, e_n)$ *est une base de* E *alors*

$$y \in \mathrm{Ker}(\varphi) \iff \varphi(e_1, y) = \cdots = \varphi(e_n, y) = 0.$$

Si on note M_φ *la matrice de* φ *dans la base* $\mathcal{B}$*, on traduit matriciellement cette propriété par :*

$$e_1^t M_\varphi Y = \cdots = e_n^t M_\varphi Y = 0,$$

oİ Y *est le vecteur colonne dont les coordonnées sont données par la décomposition de* y *dans la base* $(e_1, \ldots, e_n)$*. Ceci est équivalent à dire que toutes les coordonnées de* $M_\varphi y$ *sont nulles. Autrement dit on a la caractérisation suivante :*

$$y \in \mathrm{Ker}(\varphi) \iff M_\varphi y = 0.$$

Exemple 32. $\varphi((x_1, x_2), (y_1, y_2)) = x_1y_1 - x_2y_2$ *est une forme bilinéaire symétrique sur* $\mathbb{R}^2$*. Si on détermine le noyau de* φ *on a* $\varphi((1, 0), (y_1, y_2)) = y_1 = 0$ *et* $\varphi((0, 1), (y_1, y_2)) = -y_2 = 0$*, c'est-à-dire* $y = (y_1, y_2) = 0$*. Ainsi* $\mathrm{Ker}(\varphi) = \{0\}$*. D'autre part, on constate que* $\varphi((1, 1), (1, 1)) = 0$*. Cette forme bilinéaire symétrique est donc non dégénérée, mais n'est pas définie.*

Proposition 33. *Soit* φ *une forme bilinéaire sur un espace vectoriel* E*. On a*

$$rg(\varphi) + \dim(\mathrm{Ker}(\varphi)) = \dim(E)$$

où $rg(\varphi)$ *est le rang de* φ*.*

DÉMONSTRATION. La preuve découle du théoréme de rang pour les applications linéaires en réinterprétant la matrice associée à φ comme matrice d'une application linéaire. Voici une seconde preuve autonome. Notons $p = \dim(\mathrm{Ker}(\varphi))$. On considère $(e_1, \ldots, e_p)$ une base de $\mathrm{Ker}(\varphi)$, que l'on complète en une base $(e_1, \ldots, e_n)$ de E. Dans cette base la matrice de φ est de la forme

$$A = \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix}$$

avec B une matrice de dimension $n-p$. Il est donc clair que $rg(\varphi) \leq n-p$.
Supposons par l'absurde qu'il existe une relation de dépendance linéaire entre les colonnes de B. Quitte à changer l'ordre des $n-p$ derniers vecteurs de base, on peut toujours supposer que la dernière colonne de B (ou de A) est combinaison linéaire des $n-p-1$ précédentes. Il existe donc des scalaires $(\lambda_1, \ldots, \lambda_{n-p-1})$ tels que :

$$\forall i \in 1, .., n,\ \varphi(e_i, e_n) = \sum_{j=1}^{n-p-1} \lambda_j \varphi(e_i, e_{j+p}).$$

On en déduit que $e_n - \sum_{j=1}^{n-p-1} \lambda_j e_{j+p}$ appartient à $\mathrm{Ker}(\varphi)$. Ce qui n'est pas le cas par construction de la base $(e_1, \ldots, e_n)$. Ainsi on aboutit à une contraction et on a donc $rg(\varphi) = n-p$. □

III.2. Orthogonalité.

Définition 34. *Soit φ une forme bilinéaire symétrique sur un espace vectoriel E. On dit que deux vecteurs x et y de E sont* orthogonaux *(pour φ) si $\varphi(x, y) = 0$.*

Définition 35. *Soit φ une forme bilinéaire symétrique sur un espace vectoriel E. Si x est un vecteur de E tel que $\varphi(x, x) = 0$ on dit que x est* isotrope.

Remarque 36. *L'ensemble des vecteurs isotropes n'est pas en général un sous-espace vectoriel. En effet cet ensemble est stable par multiplication par un scalaire, mais pas nécessairement par addition. Soit φ une forme bilinéaire symétrique sur un K-espace vectoriel E. Si $\varphi(x, x) = \varphi(y, y) = 0$ alors on a*

$$\begin{aligned} &\forall \lambda \in K,\ \varphi(\lambda x, \lambda x) = \lambda^2 \varphi(x, x) = 0 \\ &\varphi(x+y, x+y) = \varphi(x, x) + \varphi(y, y) + 2\varphi(x, y) = 2\varphi(x, y) \end{aligned}$$

qui n'est pas forcément nul.

Définition 37. *Si U est une partie de E, l'ensemble*

$$U^\perp = \{y \in E \mid \forall x \in U,\ \varphi(x, y) = 0\}$$

est un sous-espace vectoriel de E que l'on appelle orthogonal *de U.*

Proposition 38. *Soient φ une forme bilinéaire symétrique sur un espace vectoriel E. Alors on a les propriétés suivantes :*
a. $\{0\}^\perp = E$
b. $E^\perp = \mathrm{Ker}(\varphi)$.

Remarque 39. *Attention, on a $E^\perp = \{0\}$ si et seulement si φ est non dégénérée.*

DÉMONSTRATION. Pour tout $x \in E$ on a $\varphi(x, 0) = 0$, c'est-à-dire $x \in \{0\}^\perp$, et donc on obtient le résultat. Pour le second point on a directement $\mathrm{Ker}(\varphi) = \{y \in E \mid \forall x \in E,\ \varphi(x, y) = 0\} = E^\perp$. □

Caractérisons un peu plus précisément l'orthogonal d'un sous-espace vectoriel.

Théorème 40.

Soit φ une forme bilinéaire symétrique sur un espace vectoriel E. On note F un sous-espace vectoriel E. Alors on a

$$\dim(F) + \dim(F^\perp) \geq \dim(E) \quad \text{et} \quad F \subset (F^\perp)^\perp.$$

De plus, lorsque φ est non dégénérée il y a égalité dans les deux relations ci-dessus.

Remarque 41. *On constate qu'un sous-espace vectoriel et son orthogonal ne sont pas nécessairement supplémentaires même pour une forme bilinéaire non dégénérée. Ceci va être illustré dans les exemples après la démonstration.*

Démonstration. Soit F un sous-espace vectoriel de E ($\dim(E) = n$) dont on note $(e_1, \dots, e_p)$ une base. Le sous-espace vectoriel $F^\perp$ de E est l'ensemble des $y \in E$ qui vérifient les p équations linéaires $\varphi(e_1, y) = 0, \dots, \varphi(e_p, y) = 0$. C'est un système linéaire de p équations à n inconnues. L'ensemble des solutions est donc un espace vectoriel de dimension supérieure ou égale à $n - p$. Ce qui conduit à la première relation. Pour la seconde on a : si $x \in F$, alors pour tout $y \in F^\perp$, $\varphi(x, y) = 0$ et par conséquent $x \in (F^\perp)^\perp$. Le cas oİ φ est non dégénérée est laissé au lecteur. □

Exemple 42. *Soit $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ telle que $\varphi(x, y) = x_1y_1 - x_2y_2$. On vérifie que φ est une forme bilinéaire symétrique de noyau engendré par le vecteur $(0, 0, 1)$. Soit $H = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = x_2 \text{ et } x_3 = 0\}$. On a $\dim H = 1$ et $u(1, 1, 0)$ est une base de H. Déterminons l'orthogonal de H :*

$$y \in H^\perp \iff y \in \{u\}^\perp \iff \varphi(u, y) = 0 \iff y_1 - y_2 = 0.$$

Ainsi on a $H^\perp = \{(y_1, y_2, y_3) \in \mathbb{R}^3 \mid y_1 = y_2\}$, c'est-à-dire que $H \subset H^\perp$. Donc que H et $H^\perp$ ne sont pas supplémentaires. Dans cet exemple φ était dégénérée. Mais dans l'exemple suivant elle ne l'est plus.

Exemple 43. *Soit $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ telle que $\varphi(x, y) = x_1y_1 - x_2y_2$. On vérifie que φ est une forme bilinéaire symétrique de noyau $\{0\}$. Soit H le sous-espace vectoriel de $\mathbb{R}^2$ engendré par $u(1, 1)$. Déterminons l'orthogonal de H :*

$$y \in H^\perp \iff y \in \{u\}^\perp \iff \varphi(u, y) = 0 \iff y_1 - y_2 = 0.$$

Ainsi on a $H^\perp = \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 = y_2\}$, c'est-à-dire que $H^\perp = H$ et donc H et $H^\perp$ ne sont pas supplémentaires.

En fait on a le résultat suivant.

Proposition 44. *Soient φ une forme bilinéaire symétrique non dégénérée sur un espace vectoriel E. Soit F un sous-espace vectoriel de E. Alors on a :*

$$\forall x \in F,\, x \neq 0,\, \varphi(x, x) \neq 0 \Rightarrow E = F \oplus F^\perp$$

.

Démonstration. Puisque φ est non dégénérée, on a $\dim(F) + \dim(F^\perp) = \dim(E)$. On en déduit que :

$$\forall x \in F,\, x \neq 0,\, \varphi(x, x) = 0 \Rightarrow F \cap F^\perp = \{0\} \iff E = F \oplus F^\perp.$$

□

Donnons enfin quelques dernières propriétés de l'orthogonal d'une partie de E.

Proposition 45. *Soient φ une forme bilinéaire symétrique sur un espace vectoriel E et A et B deux parties de E. Alors on a les propriétés suivantes :*

— $(A + B)^\perp = (A \cup B)^\perp = A^\perp \cap B^\perp$.

— $A^\perp + B^\perp \subset (A \cap B)^\perp$.
— $A \subset B \Rightarrow B^\perp \subset A^\perp$.

Remarque 46. *Il est évident que pour A une partie quelconque de E on a encore $A \subset (A^\perp)^\perp$. Il suffit de reprendre la démostration faite dans le cadre d'un sous-espace vectoriel.*

DÉMONSTRATION. La première égalité du point a provient du fait que $\mathrm{Vect}(A \cup B) = A + B$. De plus, on a

$$\begin{aligned} x \in (A \cup B)^\perp \quad &\Longleftrightarrow \quad \forall y \in A \cup B,\ \varphi(x,y) = 0 \\ &\Longleftrightarrow \quad \forall y \in A,\ \varphi(x,y) = 0 \text{ et } \forall y \in B,\ \varphi(x,y) = 0 \quad \text{car } A \subset A \cup B \text{ et } B \subset A \cup B \\ &\Longleftrightarrow \quad x \in A^\perp \ \text{ et } x \in B^\perp \iff x \in A^\perp \cap B^\perp. \end{aligned}$$

Pour la propriété suivante on peut écrire :

$$x \in A^\perp + B^\perp \iff \exists a \in A^\perp \text{ et } \exists b \in B^\perp \mid x = a + b.$$

On obtient donc

$$\forall y \in A \cap B,\ \varphi(x,y) = \varphi(a+b,y) = \varphi(a,y) + \varphi(b,y) = 0;$$

en effet on a $\varphi(a,y) = 0$ car $y \in A$ et $a \in A^\perp$, et $\varphi(b,y) = 0$ car $y \in B$ et $b \in B^\perp$. Ainsi $x \in (A \cap B)^\perp$, ce qui montre le résultat.

Enfin on a :

$$\begin{aligned} y \in B^\perp \quad &\Longleftrightarrow \quad \forall x \in B,\ \varphi(x,y) = 0 \\ &\Rightarrow \quad \forall x \in A,\ \varphi(x,y) = 0 \text{ car } A \subset B \\ &\Rightarrow \quad y \in A^\perp \end{aligned}$$

ce qui démontre le dernier point. □

IV. Formes quadratiques

IV.1. Généralités.

Définition 47. *On dit qu'une application $Q : E \to K$ est une* forme quadratique *sur un espace vectoriel E s'il existe une forme bilinéaire symétrique φ sur $E \times E$ vérifiant $Q(x) = \varphi(x,x)$ pour tout x dans E. φ est appelée* la forme bilinéaire associée *à Q.*

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E, alors la matrice de Q dans cette base est exactement celle de sa forme bilinéaire associée φ. Si on note $M_\mathcal{B}$ cette matrice, et si de plus X est la matrice colonne (dans $\mathcal{B}$) d'un vecteur x de E, alors

$$Q(x) = {}^tXM_\mathcal{B}X = \sum_{1 \le i,j \le n} m_{ij}x_ix_j = \sum_{i=1}^n m_{ii}x_i^2 + 2\sum_{i<j} m_{ij}x_ix_j,$$

car φ étant symétrique on a $m_{ij} = \varphi(e_i,e_j) = \varphi(e_j,e_i) = m_{ji}$. Une forme quadratique s'écrit donc comme un polynôme homogène de degré 2 (tous les monomes sont de degré 2).

Réciproquement soit le polynôme en $x_1, \dots, x_n$ homogène de degré 2 suivant :

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} a_{ij}x_ix_j.$$

On a

$$p(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{i<j} m_{ij}x_ix_j = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{i \ne j} a_{ij}(x_ix_j/2 + x_jx_i/2) = \varphi(x,x),$$

où φ est la forme bilinéaire symétrique de matrice $(m_{ij})_{1 \le i,j \le n}$ avec $m_{ii} = a_{ii}$ et si $i \ne j$, $m_{ij} = m_{ji} = a_{ij}/2$. Ce qui démontre la réciproque.

Remarque 48. *Le rang d'une forme quadratique Q* est le rang de sa matrice dans une base quelconque. Comme les matrices de Q et de sa forme bilinéaire symétrique associée sont les mêmes, on en déduit que le rang d'une forme quadratique est égal au rang de sa forme bilinéaire symétrique associée. De plus la proposition 9 entraîne que : $\mathrm{rg}(Q) = \dim(E) - \dim(\mathrm{Ker}(\varphi))$.

Remarque 49. *Soit Q une forme quadratique et φ sa forme bilinéaire symétrique associée. Alors φ est définie si et seulement si sa forme quadratique associée, Q, l'est aussi ; c'est-à-dire qu'elle vérifie : $\forall x \in E, \quad Q(x) = 0 \iff x = 0$. De plus φ est positive si et seulement si sa forme quadratique associée est aussi positive :*

$$\forall x \in E, \quad Q(x) \geq 0.$$

Proposition 50. *Formule de polarisation.*

Si Q est une forme quadratique sur E et si φ est sa forme bilinéaire symétrique associée alors

$$\forall (x, y) \in E \times E, \quad \varphi(x, y) = \frac{1}{2}\Big(Q(x + y) - Q(x) - Q(y)\Big).$$

Démonstration. Comme φ est une forme bilinéaire symétrique, on a

$$Q(x + y) = \varphi(x + y, x + y) = \varphi(x, x) + \varphi(x, y) + \varphi(y, x) + \varphi(y, y) = Q(x) + 2\varphi(x, y) + Q(y),$$

c'est-à-dire la formule de la proposition. □

Remarque 51. *Cette formule montre que pour toute forme quadratique il existe une unique forme bilinéaire symétrique associée.*

Attention, étant donné une forme quadratique Q, il existe en général une infinité de formes bilinéaires φ vérifiant $Q(x) = \varphi(x, x)$. Par exemple, si Q est la forme quadratique sur $\mathbb{R}^2$ définie par $Q((x_1, x_2)) = x_1x_2$ et si φ_λ est la forme bilinéaire définie par

$$\varphi_\lambda((x_1, x_2), (y_1, y_2)) = \lambda x_1y_2 + (1 - \lambda)x_2y_1,$$

alors on a $\varphi_\lambda(x, x) = Q(x)$, pour tout λ. Mais ces formes ne sont pas symétriques sauf pour $\lambda = 1/2$, qui correspond à la forme bilinéaire symétrique associée.

Corollaire 52. *Soit Q une forme quadratique sur E dont l'expression dans la base $(e_1, \dots, e_n)$ est pour tout $x = x_1e_1 + \cdots + x_ne_n$ est*

$$Q(x) = \sum_{i=1}^{n} \alpha_{ii}x_i^2 + \sum_{1 \leq i < j \leq n} \alpha_{ij}x_ix_j$$

avec α_{ij} des scalaires. Alors la forme bilinéaire symétrique associée à Q, a pour expression pour tout $x = x_1e_1 + \cdots + x_ne_n$ et tout $y = y_1e_1 + \cdots + y_ne_n$

$$\varphi(x, y) = \sum_{i=1}^{n} \alpha_{ii}x_iy_i + \frac{1}{2} \sum_{1 \leq i < j \leq n} \alpha_{ij}(x_iy_j + x_jy_i)$$

Ceci se vérifie facilement avec le résultat précédent.

IV.2. Réduction d'une forme quadratique.

Le problème de réduction d'une forme quadratique Q, ou d'une forme bilinéaire symétrique φ, consiste à chercher une base de E qui donne à l'expression de $Q(x)$ ou de $\varphi(x, y)$ la forme la plus simple possible ou la mieux adaptée à une situation donnée.

Théorème 53.

> *Soit φ une forme bilinéaire symétrique quelconque sur $E \times E$. Alors, il existe une base $(e_1, \dots, e_n)$ de E formée de vecteurs orthogonaux par rapport à φ. On a donc $\varphi(e_i, e_j) = 0$ si $i \neq j$.*

On remarque que la matrice de φ dans cette base est diagonale.

DÉMONSTRATION. Elle se fait par récurrence sur la dimension n de E. Si $n = 1$, il n'y a rien à démontrer. Supposons le théorème établi pour la dimension $n - 1$. Soit φ une forme bilinéaire symétrique définie sur l'espace E de dimension n, et Q la forme quadratique associée. Si $\varphi = 0$ toute base de E convient. Si $\varphi \neq 0$, on a aussi $Q \neq 0$ d'après la formule de polarisation, et il existe un vecteur e_1 dans E tel que $Q(e_1) = \varphi(e_1, e_1) \neq 0$. Considérons alors le sous-espace vectoriel E_1 de E constitué par les vecteurs $x \in E$ tels que $\varphi(e_1, x) = 0$. Alors E_1 et Ke_1 sont supplémentaires dans E, c'est-à-dire

$$E = Ke_1 \oplus E_1.$$

En effet, on a d'abord $E = Ke_1 + E_1$, car il suffit, $x \in E$ étant donné, de déterminer $\lambda \in K$ tel que $y = x - \lambda e_1$ vérifie $\varphi(e_1, y) = \varphi(e_1, x - \lambda e_1) = 0$; cette égalité équivaut à $\varphi(e_1, x) - \lambda\varphi(e_1, e_1) = 0$, et donc $\lambda = \varphi(e_1, x)/Q(e_1)$.

De plus $Ke_1 \cap E_1 = \{0\}$, car l'égalité $\varphi(e_1, x) = 0$, avec $x = \lambda e_1$, implique

$$\varphi(e_1, \lambda e_1) = \lambda\varphi(e_1, e_1) = \lambda Q(e_1) = 0,$$

et donc $\lambda = 0$ et $x = 0$ car $Q(e_1) \neq 0$.

E_1 et Ke_1 étant supplémentaires, l'espace E_1 est de dimension $n-1$. On peut alors appliquer l'hypothèse de récurrence à la restriction de Q à E_1. En prenant la base correspondante $(e_2, \dots, e_n)$, on obtient avec $(e_1, e_2, \dots, e_n)$ une base de E qui convient pour le théorème. □

Corollaire 54. *Soit Q une forme quadratique sur E. Il existe une base de E dans laquelle Q s'exprime par :*

$$Q(x) = \sum_{i=1}^{r} a_i x_i^2, \qquad a_i = Q(e_i) \neq 0,$$

où r est le rang de Q. Dans cette base, la forme bilinéaire symétrique associée φ s'écrit :

$$\varphi(x, y) = \sum_{i=1}^{r} a_i x_i y_i, \qquad a_i = \varphi(e_i, e_i) \neq 0.$$

DÉMONSTRATION. Soit Q une forme quadratique sur E et φ sa forme bilinéaire associée. D'après le théorème précédent il existe une base $(e_1, \dots, e_n)$ orthogonale pour φ. Dans cette base on a

$$Q(x) = \sum_{i=1}^{n} \varphi(e_i, e_i) x_i^2 + 2 \sum_{i<j} \varphi(e_i, e_j) x_i x_j = \sum_{i=1}^{n} \varphi(e_i, e_i) x_i^2,$$

car $\varphi(e_i, e_j) = 0$ si $i \neq j$. De plus comme Q est de rang r, il en est de même pour φ, donc il y a exactement r coefficients $\varphi(e_i, e_i)$ qui ne sont pas nuls. Après les avoir numérotés de 1 à r si nécessaire, on obtient le résultat du corollaire. □

Corollaire 55. *Toute matrice symétrique est congruente à une matrice diagonale.*

DÉMONSTRATION. Soit M une matrice symétrique. On peut la voir comme la matrice d'une forme bilinéaire symétrique φ. Donc d'après le théorème précédent il existe une base orthogonale pour φ. Dans cette base la matrice M' de φ est diagonale, et par la formule de changement de base on a $M' = {}^tPMP$. □

Dans les problèmes de réduction, on peut essayer d'aller plus loin en donnant une forme particulière aux coefficients a_i. Mais les résultats dépendent du corps de base. Nous allons préciser ceci dans les cas $K = \mathbb{C}$ et $K = \mathbb{R}$.

IV.3. Réduction sur un espace vectoriel complexe ($K = \mathbb{C}$).

Soit Q une forme quadratique de rang r sur un espace vectoriel complexe E. D'après le corollaire 19, il existe une base $(e_1, \dots, e_n)$ dans laquelle Q s'écrit

$$Q(x) = \sum_{i=1}^{r} a_i x_i^2, \qquad a_i = Q(e_i) \neq 0.$$

Comme les a_i sont des coefficients complexes, on peut poser :

$$a_i = b_i^2,\ b_i \neq 0,\ i = 1, \dots, r,$$

puis effectuer le changement de variable

$$y_i = b_i x_i,\ i = 1, \dots, r \quad ; \quad y_j = x_j,\ j = r+1, \dots, n,$$

provenant du changement de base :

$$e'_i = \frac{1}{b_i} e_i,\ i = 1, \dots, r \quad ; \quad e'_j = e_j,\ j = r+1, \dots, n.$$

La forme quadratique Q est alors donnée dans la base $(e'_1, \dots, e'_n)$ par

$$Q(y) = \sum_{i=1}^{r} y_i^2.$$

On peut remarquer que la nouvelle base reste orthogonale pour la forme bilinéaire associée à Q puisqu'on effectue une simple homothétie des vecteurs de la base de départ. Nous avons donc démontré le théorème suivant.

Théorème 56.

Pour toute forme quadratique Q de rang r sur un espace vectoriel complexe E il existe une base dans laquelle Q s'écrit :

$$Q(x) = \sum_{i=1}^{r} x_i^2.$$

On dit que Q est décomposable en somme de carrés de r formes linéaires indépendantes.

Une conséquence directe de ce résultat, est que toute matrice symétrique complexe d'ordre n et de rang r est congruente à la matrice

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

dont les r premiers coefficients diagonaux valent 1 et tous les autre sont nuls.

On en déduit que deux formes quadratiques de même rang sur un espace vectoriel complexe sont toujours congruentes.

IV.4. Réduction sur un espace vectoriel réel ($K = \mathbb{R}$).

Soit Q une forme quadratique de rang r sur un espace vectoriel réel E. D'après le corollaire 19, il existe une base $(e_1, \ldots, e_n)$ dans laquelle Q s'écrit

$$Q(x) = \sum_{i=1}^{r} a_i x_i^2, \qquad a_i = Q(e_i) \neq 0.$$

Comme on travaille sur le corps des réels, on est amené à distinguer dans cette décomposition les coefficients a_i qui sont positifs et ceux qui sont négatifs. Dans le premier cas on pose $a_i = b_i^2,\ b_i \neq 0$, et dans le deuxième $a_i = -b_i^2,\ b_i \neq 0$. Après le changement de base

$$e_i' = \frac{1}{b_i} e_i,\ i = 1, \ldots, r \quad ; \quad e_j' = e_j,\ j = r+1, \ldots, n,$$

qui se traduit par le changement de coordonnées

$$y_i = b_i x_i,\ i = 1, \ldots, r \quad ; \quad y_j = x_j,\ j = r+1, \ldots, n,$$

la forme quadratique prend l'expression réduite :

$$Q(y) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2.$$

Nous allons compléter ce résultat en démontrant que p (et par suite $p' = r - p$) ne dépend que de la forme quadratique Q et non pas de la base choisie.

Théorème 57.

> *Soit Q une forme quadratique de rang r sur un espace vectoriel réel E de dimension n. Il existe une base de E dans laquelle Q s'écrit*
> $$Q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2.$$
> *De plus p et $p' = r - p$ ne dépendent que de Q et non pas de la base choisie.*

Définition 58. *Le couple $s = (p, p')$, qui est formé par le nombre de carrés précédés du signe $+$ et le nombre de carrés précédés du signe $-$ s'appelle la* signature *de la forme quadratique Q à coefficients réels.*

Le théorème exprime donc que $s = (p, p')$, avec le rang $r = p + p'$ ne dépendent que de Q. C'est la *loi d'inertie de Sylvester*.

DÉMONSTRATION. On a déjà montré que Q se décompose sous la forme donnée dans le théorème. Il reste à démontrer l'unicité de p dans toutes les décompositions possibles. Soit donc

$$\begin{aligned} Q(x) &= y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 \\ &= z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2 \end{aligned}$$

où $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ et $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ sont les coordonnées du vecteur x dans les bases $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e_1', \ldots, e_n')$.
On sait que les sous-espaces $\mathbb{R}e_{r+1} + \cdots + \mathbb{R}e_n$ et $\mathbb{R}e_{r+1}' + \cdots + \mathbb{R}e_n'$ coïncident avec le noyau de la forme bilinéaire symétrique associée, $\mathrm{Ker}(\varphi)$.
Supposons $q > p$. De l'expression ci-dessus on déduit pour tout x dans E

$$y_1^2 + \cdots + y_p^2 + z_{q+1}^2 + \cdots + z_r^2 = z_1^2 + \cdots + z_q^2 + y_{p+1}^2 + \cdots + y_r^2.$$

Nous allons prendre un vecteur x tel que

$$y_1 = \cdots = y_p = 0 \quad \text{et} \quad z_{q+1} = \cdots = z_r = z_{r+1} = \cdots = z_n = 0,$$

donc dans l'intersection des sous-espaces vectoriels :

$$F = \mathbb{R}e_{p+1} + \cdots + \mathbb{R}e_r + \cdots \mathbb{R}e_n \quad \text{et} \quad G = \mathbb{R}e'_1 + \cdots + \mathbb{R}e'_q.$$

On a $F \cap G \neq \{0\}$; en effet sinon la dimension de l'espace $F+G$ vérifierait

$$\dim F + \dim G = n - p + q > n,$$

ce qui est impossible. Il existe donc bien $x \neq 0$ dans $F \cap G$. On déduit alors de l'égalité ci-dessus que

$$z_1^2 + \cdots + z_q^2 + y_{p+1}^2 + \cdots + y_r^2 = 0,$$

et comme les coordonnées sont réels que $z_1 = \cdots = z_q = 0 = y_{p+1} = \cdots = y_r$, c'est-à-dire contenu du choix pour x, que x est nul. L'hypothèse $q > p$ conduit donc à une contradiction.
On élimine de la même façon le cas $p > q$. Il en résulte que $p = q$ et le théorème est démontré. □

Corollaire 59. *Soit une forme quadratique Q à coefficients réels de signature $s = (p, p')$ dans un espace vectoriel E de dimension n. On a les prorpiétés suivantes :*

- *$p + p' = n$ implique que Q est non dégénérée.*
- *$p' = 0$ implique que Q est positive.*
- *$p = 0$ implique que Q est négative.*
- *$s = (n, 0)$ implique que Q est définie positive.*
- *$s = (0, n)$ implique que Q est définie négative.*

Démonstration. Cela se déduit des résultats précédents. □

IV.5. Méthode de Gauss de réduction des formes quadratiques.

Pour trouver le rang d'une forme quadratique on peut calculer celui de sa matrice dans une base donnée. En revanche, les résultats que nous avons démontrés ne permettent pas de déterminer pratiquement la signature dans le cas d'une forme quadratique réelle. La méthode de Gauss va nous permettre de décomposer effectivement toute forme quadratique, réelle ou complexe, en combinaison linéaire de carrés de formes linéaires indépendantes.

On considère la forme quadratique non nulle

$$Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

dont les coefficients appartiennent à un corps K. Deux cas se présentent.

Premier cas : Q contient un carré.
C'est-à-dire qu'il existe $a_{ii} \neq 0$ dans l'expression ci-dessus. Supposons par exemple que $a_{11} \neq 0$. On écrit alors :

$$\begin{aligned}
Q(x_1, \ldots, x_n) &= a_{11} \left[x_1^2 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_1 x_j \right] + \sum_{2 \leq i < j \leq n} a_{ij} x_i x_j \\
&= a_{11} \left(x_1 + \frac{1}{2} \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j \right)^2 + \sum_{2 \leq i < j \leq n} a_{ij} x_i x_j - \frac{1}{4} \left(\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j \right)^2 \\
&= a_{11} \left(x_1 + \frac{1}{2} \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j \right)^2 + \widetilde{Q}(x_2, \ldots, x_n).
\end{aligned}$$

où $\widetilde{Q}(x_2, \ldots, x_n)$ est une forme quadratique sur un sous-espace de dimension $n-1$ (indépendante du premier terme puisque ne dépendant pas de x_1). On recommence la méthode avec $\widetilde{Q}$. Si $\widetilde{Q}$ contient un carré, on opère de même sur $\widetilde{Q}$ et ainsi de suite jusqu'à ce qu'on tombe sur une forme dont tous les termes carrés sont nuls. On applique alors la méthode du deuxième cas. Si $\widetilde{Q}$ ne contient pas de carré on applique directement la méthode suivante.

DEUXIÈME CAS : Q ne contient pas de carré.
C'est-à-dire que tous les $a_{ii}=0$. Mais comme Q au départ n'est pas nulle il existe $a_{ij}\neq 0$ dans l'expression ci-dessus. Supposons par exemple que $a_{12}\neq 0$. On écrit alors :

$$\begin{aligned} Q(x_1,\dots,x_n) &= a_{12}x_1x_2 + x_1\sum_{j=3}^n a_{1j}x_j + x_2\sum_{j=3}^n a_{2j}x_j + \sum_{3\leq i<j\leq n} a_{ij}x_ix_j \\ &= a_{12}\left(x_1+\frac{1}{a_{12}}\sum_{j=3}^n a_{2j}x_j\right)\left(x_2+\frac{1}{a_{12}}\sum_{j=3}^n a_{1j}x_j\right) \\ &\quad + \sum_{3\leq i<j\leq n} a_{ij}x_ix_j - \frac{1}{a_{12}}\sum_{j=3}^n a_{2j}x_j\sum_{j=3}^n a_{1j}x_j \end{aligned}$$

et donc

$$\begin{aligned} Q(x_1,\dots,x_n) &= \frac{a_{12}}{4}\left(x_1+x_2+\frac{\sum_{j=3}^n a_{2j}x_j+\sum_{j=3}^n a_{1j}x_j}{a_{12}}\right)^2 \\ &\quad -\frac{a_{12}}{4}\left(x_1-x_2+\frac{\sum_{j=3}^n a_{2j}x_j-\sum_{j=3}^n a_{1j}x_j}{a_{12}}\right)^2 \\ &\quad +\widetilde{Q}(x_3,\dots,x_n) \end{aligned}$$

où $\widetilde{Q}(x_3,\dots,x_n)$ est une forme quadratique sur un sous-espace de dimension $n-2$ (indépendante des premiers termes puisque ne dépendant pas de x_1 ni de x_2 ; les deux premiers termes étant eux aussi indépendants puisque l'un dépend de x_1+x_2 et l'autre de x_1-x_2). Si $\widetilde{Q}$ contient un carré on opère comme au premier cas, sinon on applique la méthode du second cas.

On procède ainsi jusqu'au moment où la forme quadratique $\widetilde{Q}$ que l'on obtient à chaque étape soit nulle. Au final Q s'écrit comme une combinaison linéaire de r carrés de formes linéaires indépendantes ; il en résulte immédiatement que le rang de Q est r.
Attention la décomposition par cette méthode n'est pas nécessairement unique.

IV.6. Exemple.
Appliquons la méthode à la forme quadratique

$$Q(x_1,x_2,x_3,x_4) = x_1^2+x_3^2+x_4^2-2x_1x_3+2x_1x_4+4x_2x_3+6x_3x_4.$$

Q contient un carré, on commence donc par appliqué le premier cas.

$$\begin{aligned} Q(x_1,x_2,x_3,x_4) &= x_1^2+2x_1(-x_3+x_4)+x_3^2+x_4^2+4x_2x_3+6x_3x_4 \\ &= (x_1-x_3+x_4)^2-(-x_3+x_4)^2+x_3^2+x_4^2+4x_2x_3+6x_3x_4 \\ &= (x_1-x_3+x_4)^2+4x_2x_3+8x_3x_4. \end{aligned}$$

On obtient $\widetilde{Q}(x_2,x_3,x_4)=4x_2x_3+8x_3x_4$, qui ne contient pas de carré. On applique donc la méthode du deuxième cas.

$$\begin{aligned} \widetilde{Q}(x_2,x_3,x_4) &= 4x_2x_3+8x_3x_4 = 4(x_2+2x_4)(x_3+0) \\ &= (x_2+x_3+2x_4)^2-(x_2-x_3+2x_4)^2+0 \end{aligned}$$

où la dernière forme quadratique est nulle. Le procédé est donc terminé et on obtient :

$$Q(x_1,x_2,x_3,x_4) = (x_1-x_3+x_4)^2+(x_2+x_3+2x_4)^2-(x_2-x_3+2x_4)^2.$$

Si Q est une forme quadratique sur un espace vectoriel réel la décomposition est achevée. Par contre si c'est une forme quadratique sur un espace vectoriel complexe on peut encore “enlever” les signes moins et on obtient :

$$Q(x_1, x_2, x_3, x_4) = (x_1 - x_3 + x_4)^2 + (x_2 + x_3 + 2x_4)^2 + (ix_2 - ix_3 + 2ix_4)^2.$$

Chapitre 4

ESPACES VECTORIELS EUCLIDIENS

Dans ce chapitre tous les espaces vectoriels sont réels et en général de dimension finie.

I. Produit scalaire et norme euclidienne

I.1. Définitions.

Définition 1. *On appelle* produit scalaire *une forme bilinéaire symétrique définie positive sur un espace vectoriel réel.*

Un produit scalaire φ sur un espace vectoriel E est donc une application de $E \times E$ dans $\mathbb{R}$ qui vérifie :

(1) φ est une forme bilinéaire sur $E \times E$;

(2) $\varphi(x, y) = \varphi(y, x)$ (symétrique) ;

(3) $\varphi(x, x) = 0 \iff x = 0$ (définie) ;

(4) $\varphi(x, x) \geq 0$ (positive).

En général on note un produit scalaire par $\varphi(x, y) = (x, y)$ ou $\varphi(x, y) = x.y$.

Exemple 2. *Dans le chapitre précédent, on a vu que $\varphi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1y_1 + \cdots + x_ny_n$ est une forme bilinéaire symétrique définie positive sur $E = \mathbb{R}^n$. C'est donc un produit scalaire.*

Exemple 3. *Sur $E = \mathcal{C}^0[0, 1]$, l'espace vectoriel des fonctions continues sur $[0, 1]$. On peut montrer que $\varphi(f, g) = \int_0^1 f(t)g(t)dt$ définit un produit scalaire sur E.*

Définition 4. *Un espace vectoriel réel de dimension finie muni d'un produit scalaire s'appelle un* espace vectoriel euclidien *ou plus simplement un* espace euclidien

Définition 5. *On dit que deux vecteurs x et y d'un espace euclidien E muni du produit scalaire φ sont* orthogonaux *si $(x, y) = 0$. On note $x \perp y$.*

Définition 6. *Si φ est un produit scalaire sur E, l'application $N : E \to \mathbb{R}^+$ définie par $N(x) = \sqrt{\varphi(x, x)}$ est appelée* norme euclidienne *sur E associée à φ.*

Remarque 7. *Comme un produit scalaire est une forme bilinéaire symétrique définie positive,*

$$N(x)^2 = \left(\sqrt{(x, x)}\right)^2 = (x, x)$$

est une forme quadratique définie positive sur E. On note habituellement une norme de la manière suivante : $||x||$. Il existe des tas de normes sur $\mathbb{R}^n$ qui ne sont pas euclidiennes (c'est-à-dire qui ne sont pas définies à partir d'un produit scalaire) ; par exemple la norme

$$||(x_1, \ldots, x_n)|| = |x_1| + \cdots + |x_n|.$$

On peut montrer (exercice) qu'une norme $||.||$ sur un espace vectoriel réel E est euclidienne si et seulement si elle vérifie pour tout x et tout y dans E :

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

I.2. Propriétés.

Nous allons maintenant démontrer quelques propriétés des produits scalaires et des normes sur un espace euclidien E.

Propriété 1. *Inégalité de Cauchy-Schwarz*

$$\forall x \in E,\ \forall y \in E, \quad |(x, y)| \leq ||x||\,||y||.$$

L'égalité n'a lieu que si les deux vecteurs sont colinéaires.

DÉMONSTRATION. Soit $\lambda \in \mathbb{R}$, on a

$$(x + \lambda y, x + \lambda y) = (x, x) + 2\lambda(x, y) + \lambda^2(y, y) = ||x||^2 + 2\lambda(x, y) + \lambda^2||y||^2 > 0$$

si $x + \lambda y \neq 0$. Donc le discriminant de cette équation du second degré en λ doit être strictement négatif, c'est-à-dire $\Delta = 4(x, y)^2 - 4||x||^2||y||^2 < 0$, ou encore que $(x, y)^2 < ||x||^2||y||^2$; et en prenant la racine on trouve l'inégalité. On remarque qu'il n'y égalité que si le discriminant est nul, ce qui entraîne que l'équation du second degré doit elle aussi être nulle. Comme un produit scalaire est défini ceci signifie que $x + \lambda y = 0$, ou encore $x = -\lambda y$, c'est-à-dire que x et y sont colinéaires. □

Propriété 2. Une norme euclidienne sur un espace euclidien E vérifie

i : $||x|| = 0 \iff x = 0$;

ii : $\forall x \in E,\ \forall \lambda \in \mathbb{R}, \quad ||\lambda x|| = |\lambda|\,||x||$;

iii : $\forall x \in E,\ \forall y \in E, \quad ||x+y|| \leq ||x||+||y||$; c'est *l'inégalité de Minkowsky* ou *l'inégalité triangulaire.*

DÉMONSTRATION. On a : $||x|| = 0 \iff ||x||^2 = 0 \iff (x, x) = 0 \iff x = 0$, car un produit scalaire est défini.

Pour le deuxième point on a $||\lambda x|| = \sqrt{(\lambda x, \lambda x)} = \sqrt{\lambda^2(x, x)} = |\lambda|\sqrt{(x, x)} = |\lambda|\,||x||$. Enfin pour le troisième point on a

$$||x + y||^2 = (x + y, x + y) = ||x||^2 + 2(x, y) + ||y||^2 \leq ||x||^2 + 2||x||\,||y|| + ||y||^2 = (||x|| + ||y||)^2$$

et donc le résultat. □

Remarque 8. *Dans la démonstration du troisième point, il y a égalité si et seulement si $(x, y) = ||x||\,||y||$. Or d'après la propriété précédente ceci ne peut se produire que si x et y sont colinéaires. Supposons donc $y = \lambda x$ avec $\lambda \in \mathbb{R}$. On a*

$$(x, y) = (x, \lambda x) = \lambda||x||^2 \text{ et } ||x||\,||y|| = ||x||\,||\lambda x|| = |\lambda|\,||x||^2.$$

Donc pour avoir une égalité il faut que $\lambda = |\lambda|$, c'est-à-dire que λ soit positif. On a donc égalité si et seulement si x et y sont colinéaires et de même sens.

Propriété 3. *Théorème de Pythagore*

Soient x et y deux vecteurs d'un espace euclidien E. On a

$$x \perp y \iff ||x + y||^2 = ||x||^2 + ||y||^2.$$

DÉMONSTRATION. On a

$$||x + y||^2 = (x + y, x + y) = ||x||^2 + 2(x, y) + ||y||^2,$$

et donc

$$||x + y||^2 = ||x||^2 + ||y||^2 \iff (x, y) = 0 \iff x \perp y.$$

□

II. Orthogonalité

II.1. Bases orthogonales et orthonormées.

Proposition 9. *Dans un espace euclidien de dimension n une famille de p ($\leq n$) vecteurs non nuls deux à deux orthogonaux est libre.*

DÉMONSTRATION. Soit $(e_1, \ldots, e_p)$ une famille de p vecteurs non nuls deux à deux orthogonaux dans un espace euclidien de dimension $n \geq p$. Supposons $\sum_{1 \leq j \leq p} \lambda_j e_j = 0$. En calculant le produit scalaire de ce vecteur avec e_k, pour tout $1 \leq k \leq p$, on obtient

$$0 = \sum_{1 \leq j \leq p} \lambda_j (e_j, e_k) = \lambda_k ||e_k||^2,$$

c'est-à-dire $\lambda_k = 0$ car $e_k \neq 0$. La famille est donc libre. □

Définition 10. *Soit E un espace euclidien. On dit qu'une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E est* orthogonale *si pour $i \neq j$, $e_i \perp e_j$. Si les vecteurs sont de plus unitaires (i.e. de norme égale à 1) on dit que la base est* orthonormée.

Remarquons que $\mathcal{B} = (e_1, \ldots, e_n)$ est une base orthonormée de E est équivalent à dire

$$\forall 1 \leq i \leq n, \ \forall 1 \leq j \leq n, \quad (e_i, e_j) = \delta_{ij}.$$

Théorème 11.

Dans tout espace euclidien E il existe une base orthonormée.

DÉMONSTRATION. Une première méthode pour démontrer ce théorème consisterait à utiliser les résultats qui ont été établis dans le chapitre précédent. Mais nous allons adopté une autre méthode qui sera plus effective (on construit explicitement une base orthogonale ou orthonormée).

PROCÉDÉ D'ORTHONORMALISATION DE SCHMIDT

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base d'un espace euclidien E. On construit deux bases $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ à partir de $\mathcal{B}$ par le procédé d'orthonormalisation de Schmidt, telle que la première base soit orthogonale et la seconde orthonormée. Pour ce faire on cherche les scalaires $\alpha_{i,j}$, $1 \leq j < i \leq n$, tels que

$$\begin{array}{ll}
u_1 = e_1 & v_1 = \frac{u_1}{||u_1||} \\
u_2 = \alpha_{2,1} v_1 + e_2 & v_2 = \frac{u_2}{||u_2||} \\
u_3 = \alpha_{3,1} v_1 + \alpha_{3,2} v_2 + e_3 & v_3 = \frac{u_3}{||u_3||} \\
\vdots & \vdots \\
u_n = \alpha_{n,1} v_1 + \cdots + \alpha_{n,n-1} v_{n-1} + e_n & v_n = \frac{u_n}{||u_n||}
\end{array}$$

$\alpha_{2,1}$ se calcule par $(u_1, u_2) = 0$. Ceci équivaut à

$$(v_1, u_2) = 0 \iff \alpha_{2,1} + (e_2, v_1) = 0 \iff \alpha_{2,1} = -(e_2, v_1)$$

d'où $u_2 = -(e_2, v_1) v_1 + e_2$ et $v_2 = \dfrac{u_2}{||u_2||}$.

Supposons que l'on a déterminé $u_1, \ldots, u_k$ et $v_1, \ldots, v_k$. Les coefficients de u_{k+1} se déterminent alors par $(u_i, u_{k+1}) = 0$ pour tout $1 \leq i \leq k$, ce qui revient à

$$(v_i, u_{k+1}) = 0 \iff \alpha_{k+1,i} + (e_{k+1}, v_i) = 0 \iff \alpha_{k+1,i} = -(e_{k+1}, v_i).$$

On obtient $u_{k+1} = -(e_{k+1}, v_1) v_1 - \cdots - (e_{k+1}, v_k) v_k + e_{k+1}$ et $v_{k+1} = \dfrac{u_{k+1}}{||u_{k+1}||}$. Par construction, les bases $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ que l'on obtient sont respectivement orthogonale et orthonormée. □

Exemple 12. *Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base de $\mathbb{R}^3$, avec $e_1 = (1, 1, 1)$, $e_2 = (1, 1, 0)$ et $e_3 = (1, 0, 0)$. On commence par poser $u_1 = e_1$ et $v_1 = \dfrac{u_1}{||u_1||} = (\dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}})$. Puis on a*

$$u_2 = -(e_2, v_1)v_1 + e_2 = -\frac{2}{\sqrt{3}}(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) + (1, 1, 0) = (\frac{1}{3}, \frac{1}{3}, -\frac{2}{3})$$

et $v_2 = \dfrac{u_2}{||u_2||} = (\dfrac{1}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}, -\dfrac{\sqrt{2}}{\sqrt{3}})$. Enfin pour le dernier vecteur on obtient

$$\begin{aligned} u_3 &= -(e_3, v_1)v_1 + -(e_3, v_2)v_2 + e_3 \\ &= -\frac{1}{\sqrt{3}}(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) - \frac{1}{\sqrt{6}}(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{\sqrt{2}}{\sqrt{3}}) + (1, 0, 0) = (\frac{1}{2}, -\frac{1}{2}, 0) \end{aligned}$$

et $v_3 = \dfrac{u_3}{||u_3||} = (\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, 0)$.

Proposition 13. *Soit E un espace euclidien. Si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base orthonormée de E alors les coordonnées d'un vecteur $x = (x_1, \ldots, x_n)$ dans cette base sont $x_i = (x, e_i)$ pour $1 \leq i \leq n$. De plus l'expression du produit scalaire est donné par*

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

et celle de la norme par

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{(x, e_1)^2 + \cdots + (x, e_n)^2}\,.$$

Du point de vue matriciel, si on note X et Y les matrices colonnes de x et y dans la base $\mathcal{B}$, on a

$$(x, y) = {}^tXY \quad \text{et} \quad ||X|| = \sqrt{{}^tXX}.$$

DÉMONSTRATION. Comme $\mathcal{B} = (e_1, \ldots, e_n)$ est une base orthonormée, pour le premier point on a :

$$(x, e_i) = \left(\sum_{j=1}^{n} x_j e_j, e_i\right) = \sum_{j=1}^{n} x_j (e_j, e_i) = x_i(e_i, e_i) = x_i.$$

On en déduit directement le deuxième et le troisième résultat. Les résultats matriciels sont une simple conséquence du fait que dans une base orthonormée la matrice du produit scalaire est l'identité. □

II.2. Sous-espaces vectoriels orthogonaux.

Définition 14. *Soit E un espace euclidien. On dit que deux sous-espaces vectoriels F et G de E sont* orthogonaux *si pour tout x dans F et pour tout y dans G on a $x \perp y$.*

Proposition 15. *Si F et G sont deux sous-espaces vectoriels de E orthogonaux, alors $F \cap G = \{0\}$.*

DÉMONSTRATION. Soit $x \in F \cap G$. Comme F et G sont orthogonaux on a $(x, x) = 0$, c'est-à-dire $x = 0$, d'où le résultat. □

On rappelle que si F un sous-espace vectoriel d'un espace euclidien E, alors l'ensemble des vecteurs de E orthogonaux à F est un sous-espace vectoriel de E appelé orthogonal de F et noté $F^\perp = \{y \in E \,|\, \forall x \in F,\ (x, y) = 0\}$. On a les propriétés suivantes.

Propriétés.

a) $E = F \oplus F^\perp$

b) $(F^\perp)^\perp = F$

Démonstration. Soit $(v_1, \dots, v_p)$ une base de F que l'on complète en une base de E, et que l'on transforme en une base orthonormée $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$, par le procédé de Schmidt. Comme les p premiers v_i sont des vecteurs orthonormés combinaisons linéaires des p premiers e_i, ils forment une base orthonormée de F. On considère un vecteur x de E, qui s'écrit dans cette base $x = \sum_{j=1}^{n} x_j e_j$. Alors $y = \sum_{i=1}^{p} x_i e_i$ appartient à F et on a

$$(x-y, y) = \sum_{i=p+1}^{n} \sum_{j=1}^{p} \lambda_i \lambda_j (e_i, e_j) = \sum_{i=p+1}^{n} \sum_{j=1}^{p} \lambda_i \lambda_j \delta_{ij} = 0,$$

car $i \neq j$. Le vecteur $x - y$ appartient donc à $F^\perp$ et $x = y + x - y$, c'est-à-dire $E = F + F^\perp$. De plus d'après la proposition 9, $F \cap F^\perp = \{0\}$; on en déduit a).
Pour b) on utilise a). On a $F \oplus F^\perp = E = F^\perp \oplus (F^\perp)^\perp$, d'où le résultat car on sait que pour une forme bilinéaire symétrique $F \subset (F^\perp)^\perp$.

□

II.3. Projections orthogonales.

Soit E un espace euclidien dont on note F un sous-espace vectoriel.
D'après les propriétés des sous-espaces orthogonaux on a, $E = F \oplus F^\perp$, c'est-à-dire que tout vecteur de x s'écrit $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$.

Définition 16. *L'application p de E dans E, qui à x associe $p(x) = x_1$ est appelée* projection orthogonale *sur F. On dit que $p(x)$ est le* projeté orthogonal *de x sur F.*

Remarque 17. *L'image de E par p est F et le noyau de p est $F^\perp$, c'est-à-dire :* $\mathrm{Im}(p) = F$, et, $\mathrm{Ker}(p) = F^\perp$, *car $x_2 = x - x_1 = x - p(x)$ appartient à $F^\perp$. Ici le théorème de Pythagore peut donc s'écrire*

$$||x||^2 = ||x - p(x)||^2 + ||p(x)||^2.$$

Une projection orthogonale est idempotente, c'est-à-dire que $p \circ p = p$.

Proposition 18. *Soit E un espace euclidien et F un sous-espace vectoriel de E. On note p la projection orthogonale de E sur F et x un vecteur de E. Alors pour tout y dans F on a $||y - x|| \geq ||x - p(x)||$.*

Démonstration. Soient $x \in E$ et $y \in F$. On a $y - x = (y - p(x)) + (p(x) - x)$, avec $y - p(x) \in F$ et $p(x) - x \in F^\perp$. Donc $y - p(x)$ et $p(x) - x$ sont deux vecteurs orthogonaux. Ainsi par le théorème de Pythagore, on obtient

$$||y - x||^2 = ||y - p(x)||^2 + ||p(x) - x||^2 \geq ||p(x) - x||^2,$$

c'est-à-dire $||y - x|| \geq ||p(x) - x|| = ||x - p(x)||$. □

Définition 19. *Soient A une partie non vide de E un espace euclidien et x un vecteur de E. La* distance *de x à A est définie par :* $\mathrm{d}(x, A) = \inf \{||x - y||, y \in A\}$.

Proposition 20. *Soit E un espace euclidien et F un sous-espace vectoriel de E. On note p la projection orthogonale de E sur F et x un vecteur de E. Alors on a*

$$\mathrm{d}(x, F) = ||x - p(x)||.$$

Démonstration. Cette proposition se déduit des deux résultats précédents. □

Théorème 21.

> *Soit E un espace euclidien. Toute projection orthogonale de E est diagonalisable dans une base orthonormée de E.*

DÉMONSTRATION. Soit p la projection sur F. On considère $(e_1, \ldots, e_m)$ et $(e_{m+1}, \ldots, e_n)$ des bases orthonormées (si nécessaire en appliquant le procédé de Schmidt) de F et $F^\perp$. Comme ces sous-espaces sont supplémentaires et orthogonaux $(e_1, \ldots, e_m, e_{m+1}, \ldots, e_n)$ est une base orthonormée de E. Dans cette base on a $p(e_i) = e_i$ si $1 \leq i \leq m$, car ces vecteurs appartiennent à F, et $p(e_j) = 0$ si $m + 1 \leq j \leq n$ car $e_j \in F^\perp$. Dans cette base la matrice de p est

$$\begin{array}{c} \overbrace{e_1 \ \ldots \ e_m}^{F} \ \overbrace{e_{m+1} \ \ldots \ e_n}^{F^\perp} \\ \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \end{array}$$

ce qui démontre le résultat. □

Remarque 22. *Dans la base orthonormée construite ci-dessus on a* $x = \sum_{i=1}^{n} (x, e_i) e_i$ *et donc* $p(x) = \sum_{i=1}^{m} (x, e_i) e_i$, *où* $(e_1, \ldots, e_m)$ *est la base orthonormée de* F.

II.4. Symétries orthogonales.

Soit E un espace euclidien dont on note F un sous-espace vectoriel. Pour tout x dans E on a donc $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$.

Définition 23. *L'application s de E dans E, qui à x associe $s(x) = x_1 - x_2$ est appelée* symétrie orthogonale *par rapport à F, où encore symétrie par rapport F parallèlement à $F^\perp$. On dit que $s(x)$ est le* symétrique orthogonal *de x par rapport à F .*

Remarque 24. *Si on note p la projection orthogonale sur F alors une définition équivalente du symétrique orthogonal de x est donné par $s(x) = 2p(x) - x$. En effet, on a $s(x) = x_1 - x_2 = 2x_1 - (x_1 + x_2) = 2p(x) - x$. On remarque encore que $x + s(x) = 2x_1 \in F$ et $x - s(x) = 2x_2 \in F^\perp$, c'est-à-dire que ces vecteurs sont orthogonaux. Enfin, on vérifie facilement à l'aide de la définition que $s \circ s = \mathrm{Id}_E$, on dit que s est une involution.*

Proposition 25. *Soit E un espace euclidien et F un sous-espace vectoriel de E. On note s la symétrie orthogonale de E par rapport à F. Alors on a*

$$\forall x \in E, \ \forall y \in E, \quad (s(x), s(y)) = (x, y),$$

et en particulier, $\forall x \in E, \ ||s(x)|| = ||x||$.

Démonstration. Soient $x = x_1 + x_2$ et $y = y_1 + y_2$ deux vecteurs de E décomposés suivant F et $F^\perp$. Par définition de s on a $s(x) = x_1 - x_2$ et $s(y) = y_1 - y_2$. De plus comme x_1 et y_1 sont dans F et x_2 et y_2 dans $F^\perp$ on a $(x_1, y_2) = (x_2, y_1) = 0$. On en déduit

$$\begin{aligned} (s(x), s(y)) &= (x_1 - x_2, y_1 - y_2) = (x_1, y_1) - (x_1, y_2) - (x_2, y_1) + (x_2, y_2) \\ &= (x_1, y_1) + (x_2, y_2) = (x_1, y_1) + (x_1, y_2) + (x_2, y_1) + (x_2, y_2) \\ &= (x_1 + x_2, y_1 + y_2) = (x, y). \end{aligned}$$

La seconde partie se démontre à partir de la première en prenant $y = x$. □

Théorème 26.

Soit E un espace euclidien. Toute symétrique orthogonale de E est diagonalisable dans une base orthonormée de E.

Démonstration. On peut refaire la même démonstration que dans le cas de la projection ou bien utiliser le fait que si s est la symétrie orthogonale par rapport à F et p la projection orthogonale sur F alors $s = 2p - \mathrm{Id}_E$. Or on sait qu'il existe une base orthonormée dans laquelle la matrice de p est diagonale. De plus la matrice de l'identité est diagonale dans n'importe quelle base. Donc la matrice de s est diagonale dans cette base, c'est-à-dire que s est diagonalisable dans une base orthonormée. □

Dans une telle base, comme pour tout $x \in F$ on a $s(x) = x$ et pour tout $x \in F^\perp$ on a $s(x) = -x$, la matrice de s s'écrit

$$\begin{array}{c} \overbrace{e_1 \ \dots \ e_m}^{F} \ \overbrace{e_{m+1} \ \dots \ e_n}^{F^\perp} \\ \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} \end{array}$$

III. Groupe orthogonal

III.1. Automorphismes orthogonaux.

Définition 27. *Soient E un espace euclidien et f un endomorphisme de E. On dit que f est* un opérateur orthogonal *si*

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

Remarque 28. *Par exemple, les symétries orthogonales sont des opérateurs orthogonaux, mais pas les projections.*
Si f est un opérateur orthogonal alors f est bijectif. En effet, soit x tel que $f(x) = 0$. Alors en appliquant la propriété ci-dessus pour $y = x$ on obtient $0 = ||f(x)||^2 = ||x||^2$, c'est-à-dire que $x = 0$ et donc f est injective et bijective car on travaille en dimension finie.

Définition 29. *Un endomorphisme bijectif sur un espace vectoriel E est appelé* automorphisme.

Définition 30. *L'ensemble des automorphismes orthogonaux sur un espace euclidien E s'appelle le* groupe orthogonal *de $\mathcal{L}(E)$ et est noté $O(E)$.*

Remarque 31. *$O(E)$ est un groupe pour la composition des applications c'est-à-dire que $O(E)$ vérifie :*

1) $\forall f \in O(E),\ \forall g \in O(E),\ f \circ g \in O(E)$ *;*

2) $\mathrm{Id}_E \in O(E)$*, avec* $\forall f \in O(E),\ \mathrm{Id}_E \circ f = f \circ \mathrm{Id}_E = f$ *;*

3) $\forall f \in O(E),\ f^{-1} \in O(E)$ *avec* $f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}_E$*.*

Théorème 32.

> *Soient E un espace euclidien et f un endomorphisme de E. Les assertions suivantes sont équivalentes :*
>
> *a) f est un automorphisme orthogonal ;*
>
> *b) L'image d'une base orthonormée quelconque de E est encore une base orthonormée de E ;*
>
> *c)* $\forall x \in E, \quad ||f(x)|| = ||x||$.

DÉMONSTRATION. Démontrons que a) $\Rightarrow$ b). Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E ; on note $\mathcal{B}' = (f(e_1), \dots, f(e_n))$ son image par f. Si f est orthogonal alors $(f(e_i), f(e_j)) = (e_i, e_j) = \delta_{ij}$, c'est-à-dire que $\mathcal{B}'$ est une base orthonormée de E.

Montrons maintenant que b) $\Rightarrow$ c). Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E dont l'image $\mathcal{B}' = (f(e_1), \dots, f(e_n))$ par f est encore une base orthonormée de E. Soit $x \in E$ qui s'écrit dans la première base $x = x_1 e_1 + \cdots + x_n e_n$, avec $x_i = (x, e_i)$. Son image par f s'écrit dans la seconde base $f(x) = x'_1 f(e_1) + \cdots + x'_n f(e_n)$, avec $x'_i = (f(x), f(e_i))$. Or on a

$$x'_i = (f(x), f(e_i)) = (x_1 f(e_1) + \cdots + x_n f(e_n), f(e_i)) = x_i (f(e_i), f(e_i)) = x_i$$

car $\mathcal{B}'$ est une base orthonormée. Donc $||f(x)|| = \sqrt{x_1^2 + \cdots + x_n^2} = ||x||$.

Finissons par établir que c) $\Rightarrow$ a). On suppose que f vérifie c). On a donc pour tout x dans E et tout y dans E,

$$\begin{aligned}(f(x), f(y)) &= \frac{1}{2}(||f(x)+f(y)||^2 - ||f(x)||^2 - ||f(y)||^2) = \frac{1}{2}(||f(x+y)||^2 - ||x||^2 - ||y||^2) \\ &= \frac{1}{2}(||x+y||^2 - ||x||^2 - ||y||^2) = (x, y),\end{aligned}$$

c'est-à-dire que f est orthogonal. □

Proposition 33. *Soient $f \in O(E)$ et F un sous-espace vectoriel stable par f alors $F^\perp$ est stable par f.*

DÉMONSTRATION. Soient $x \in F^\perp$ et $y \in F$, on veut montrer que $(f(x), y) = 0$. Comme F est stable par f, la restriction de f à F est une application linéaire, de F dans F, donc injective, et aussi bijective. Il existe donc $z \in F$ tel que $y = f(z)$. Ainsi

$$(f(x), y) = (f(x), f(z)) = (x, z) = 0$$

car $x \in F^\perp$ et $z \in F$. □

Finissons ce paragraphe par deux résultats concernant la réduction des endomorphismes orthogonaux.

Proposition 34. *Si f appartient à $O(E)$ et si λ est une valeur propre de f alors $|\lambda| = 1$.*

DÉMONSTRATION. Soit x un vecteur propre de f associé à λ. On a :

$$||x|| = ||f(x)|| = ||\lambda x|| = |\lambda| \, ||x||,$$

c'est-à-dire $|\lambda| = 1$ car $x \neq 0$. □

Proposition 35. *Si f appartient à $O(E)$ alors on a*

$$f \text{ est diagonalisable } \iff f \text{ est une symétrie orthogonale.}$$

Démonstration. On a déjà vu que si f est une symétrie orthogonale alors f est diagonalisable et appartient à $O(E)$.

Réciproquement si $f \in O(E)$ alors d'après la proposition précédente ses seules valeurs propres sont 1 et -1. De plus si f est diagonalisable on a $E = \mathrm{Ker}(f - \mathrm{Id}_E) \oplus \mathrm{Ker}(f + \mathrm{Id}_E)$. Pour tout $x \in E$ on a $x = x_1 + x_2$ avec $x_1 \in \mathrm{Ker}(f - \mathrm{Id}_E)$ $(f(x_1) = x_1)$ et $x_2 \in \mathrm{Ker}(f + \mathrm{Id}_E)$ $(f(x_2) = -x_2)$ et donc $f(x) = x_1 - x_2$.

Il reste à vérifier que les deux sous-espaces sont orthogonaux. Soient $x \in \mathrm{Ker}(f - \mathrm{Id}_E)$ et $y \in \mathrm{Ker}(f + \mathrm{Id}_E)$, on a $(x, y) = (f(x), f(y)) = -(x, y)$, donc $(x, y) = 0$ et x et y sont orthogonaux. Autrement dit les deux sous-espaces propres sont bien orthogonaux, et f est la symétrie orthogonale par rapport à $\mathrm{Ker}(f - \mathrm{Id}_E)$. □

III.2. Matrices orthogonales.

Définition 36. *Soit M une matrice de $\mathcal{M}_n(\mathbb{R})$. On dit que M est* orthogonale *si $M\,{}^tM = \mathrm{I}_n$. L'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$ s'appelle le* groupe orthogonal *de $\mathcal{M}_n(\mathbb{R})$ et est noté $O(n)$.*

Rappelons que dire que $O(n)$ est un groupe pour la multiplication des matrices signifie que les propriétés suivantes sont satisfaites :

— $\forall M_1 \in O(n),\ \forall M_2 \in O(n),\ M_1M_2 \in O(n)$;
— $\mathrm{I}_n \in O(n)$, et $\forall M \in O(n),\ M\mathrm{I}_n = \mathrm{I}_nM = M$;
— $\forall M \in O(n),\ M^{-1} \in O(n)$ avec $MM^{-1} = M^{-1}M = \mathrm{I}_n$.

Remarque 37. *Si M est orthogonale alors ${}^tM = M^{-1}$ et donc ${}^tMM = M^{-1}M = \mathrm{I}_n$. De plus comme ${}^t({}^tM) = M$, on en déduit que*

$$M \text{ est orthogonale } \Rightarrow {}^tM \text{ est orthogonale.}$$

Exemple 38. *La matrice*

$$M = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

est orthogonale. En effet, on a

$${}^tM = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

et $M\,{}^tM = \mathrm{I}_3$.

Proposition 39. *Soit M une matrice de $\mathcal{M}_n(\mathbb{R})$. Les assertions suivantes sont équivalentes :*

a) M est orthogonale ;
b) Les vecteurs lignes de M forment une base orthonormée de $\mathbb{R}^n$;
c) Les vecteurs colonnes de M forment une base orthonormée de $\mathbb{R}^n$.

Démonstration. Soit M une matrice de $\mathcal{M}_n(\mathbb{R})$. On note $e_1, \ldots, e_n$ ses n vecteurs lignes. Alors les vecteurs colonnes de ${}^tM$ sont exactement $e_1, \ldots, e_n$. Soit $A = M\,{}^tM$. Les coefficients de la matrice A sont égaux au produit scalaire des vecteurs lignes de M et des vecteurs colonnes de ${}^tM$, c'est-à-dire que $a_{ij} = (e_i, e_j)$ pour $1 \leq i \leq n$ et $1 \leq j \leq n$. On obtient

$$M \text{ est orthogonale } \iff A = M\,{}^tM = \mathrm{Id}_n \iff \forall 1 \leq i \leq n,\ \forall 1 \leq j \leq n, \quad (e_i, e_j) = \delta_{ij}$$

ce qui est équivalent à $(e_1, \ldots, e_n)$ est une base orthonormée de $\mathbb{R}^n$.
Comme les vecteurs colonnes de M sont exactement les vecteurs lignes de ${}^tM$ on obtient la deuxième équivalence par la même méthode en utilisant le fait que M est orthogonale équivaut à ${}^tM$ est orthogonale. □

Proposition 40. *Soient E un espace euclidien et $\mathcal{B}$ une base orthonormée de E. L'endomorphisme f de E est un automorphisme orthogonal si et seulement si sa matrice M dans la base $\mathcal{B}$ est une matrice orthogonale.*

Remarque 41. *La base orthonormée $\mathcal{B}$ étant quelconque, le résultat reste vrai dans toutes les bases orthonormées, c'est-à-dire que si f est orthogonal sa matrice dans toute base orthonormée est orthogonale.*
Ce résultat permet d'identifier les groupes $O(E)$ et $O(n)$ (ils sont isomorphes). Cela signifie que la structure de $O(E)$ ne dépend ni de E, ni du produit scalaire choisi, mais uniquement de la dimension n de E.

DÉMONSTRATION. Soit M la matrice de f dans une base $\mathcal{B} = (e_1, \ldots, e_n)$ orthonormée de E. Si M est orthogonale, alors d'après la proposition 27, ceci est équivalent à : les vecteurs colonnes $(f(e_1), \ldots, f(e_n))$ forment une base orthonormée de E ou encore à f est un automorphisme orthogonal, d'après le théorème 22. □

Proposition 42. *Le déterminant d'une matrice orthogonale est soit 1 soit -1.*

DÉMONSTRATION. Soit M une matrice orthogonale. On a donc $M\,{}^tM = \mathrm{I}_n$. On en déduit que

$$\det(M\,{}^tM) = \det(M)\det({}^tM) = \det(\mathrm{I}_n) = 1,$$

et comme $\det(M) = \det({}^tM)$ on obtient $(\det(M))^2 = 1$ c'est-à-dire $\det(M) = \pm 1$. □

Attention la réciproque de ce résultat est fausse.

Définition 43. *On appelle* groupe spécial orthogonal *de $O(n)$ le sous-groupe de $O(n)$ constitué des matrices de déterminant 1, et on le note $SO(n)$.*

Si E est un espace euclidien, $SO(E)$, le groupe spécial orthogonal *de $O(E)$, est l'ensemble des automorphismes orthogonaux de déterminant 1.*

Remarque 44. *Une propriété essentielle des endomorphismes qui appartiennent à $SO(E)$ est la suivante : $f \in SO(E) \iff$ l'image d'une base orthonormée directe de E par f est encore une base orthonormée directe de E, c'est alors vrai pour toute base orthonormée directe, et on a aussi l'image d'une base orthonormée indirecte est une base orthonormée indirecte. On dit que les éléments de $SO(E)$ conservent l'orientation de l'espace.*

Remarque 45. *Les espaces $SO(E)$ et $SO(n)$ sont isomorphes, c'est-à-dire que $SO(E)$ ne dépend que de la dimension n de l'espace euclidien E considéré.*

III.3. Changement de bases orthonormées.

Soient f un endomorphisme d'un espace euclidien E et $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées de E. On note M la matrice de f dans $\mathcal{B}$ et M' la matrice de f dans $\mathcal{B}'$. Alors le changement de base classique pour un endomorphisme d'un espace vectoriel s'écrit

$$M' = P_{\mathcal{B}' \to \mathcal{B}} M P_{\mathcal{B} \to \mathcal{B}'} = P^{-1}_{\mathcal{B} \to \mathcal{B}'} M P_{\mathcal{B} \to \mathcal{B}'}.$$

Ici on a de plus supposé que E est un espace euclidien et que les deux bases sont orthonormées. La matrice $P_{\mathcal{B} \to \mathcal{B}'}$ est celle d'un endomorphisme de E qui transforme une base orthonormée en une autre. C'est donc une matrice orthogonale. Elle vérifie par conséquent ${}^tP_{\mathcal{B} \to \mathcal{B}'} = P^{-1}_{\mathcal{B} \to \mathcal{B}'}$. La formule de changement de base devient donc

$$M' = {}^tP_{\mathcal{B} \to \mathcal{B}'} M P_{\mathcal{B} \to \mathcal{B}'}.$$

Elle est beaucoup plus pratique car il est plus facile de calculer la tranposée d'une matrice que son inverse. Mais attention, il faut l'utiliser uniquement sous les hypothèses données ci-dessus.

IV. Endomorphismes symétriques

Définition 46. *On dit qu'un endomorphisme f de E est :*
- symétrique *: si* $\forall x \in E,\ \forall y \in E, \quad (f(x), y) = (x, f(y))$ *;*
- antisymétrique *: si* $\forall x \in E,\ \forall y \in E, \quad (f(x), y) = -(x, f(y))$.

Par exemple, les projections et les symétries orthogonales sont symétriques.

Proposition 47. *Si f est un endomorphisme symétrique de E et si F est un sous-espace vectoriel de E stable par f alors $F^\perp$ est stable par f.*

Démonstration. Soit $x \in F^\perp$. Pour tout $y \in F$ on a $(f(x), y) = (x, f(y))$. Or F est stable par f donc $f(y)$ appartient à F et on en déduit que $(f(x), y) = (x, f(y)) = 0$, c'est-à-dire que $f(x) \in F^\perp$. □

Proposition 48. *Soit $f \in \mathcal{L}(E)$. f est symétrique si et seulement si sa matrice dans une base orthonormée de E est symétrique.*

Remarque 49. *De ce résultat on déduit que l'ensemble des opérateurs symétriques de E est un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension $n(n+1)/2$.*

Démonstration. Soit M la matrice de f dans une base orthonormée de E. Pour tout x et y dans E on a :

$$(f(x), y) = {}^t(MX)Y = {}^tX\,{}^tMY \text{ et } (x, f(y)) = {}^tX(MY) = {}^tXMY.$$

On en déduit que f est symétrique si et seulement si ${}^tM = M$, c'est-à-dire que M est une matrice symétrique. □

Proposition 50. *Si f est un endomorphisme symétrique (respectivement si M est une matrice symétrique) alors le polynôme caractéristique de f (respectivement de M) est scindé sur $\mathbb{R}$.*

Démonstration. Soit M la matrice de f dans une base orthonormée de E, c'est-à-dire que M est une matrice symétrique réelle. Soit $\lambda \in \mathbb{C}$ une valeur propre de M. Il existe un vecteur non nul x dans $\mathbb{C}^n$ de matrice colonne X dans la base de départ tel que $MX = \lambda X$. On a alors

$$\begin{aligned} (MX, \overline{X}) = {}^t(MX)\overline{X} &= \lambda\,{}^tX\overline{X} \\ = {}^t(MX)\overline{X} &= {}^tX\,{}^tM\overline{X} = {}^tXM\overline{X} = {}^tX\overline{MX} = \overline{\lambda}\,{}^tX\overline{X}, \end{aligned}$$

et donc $\overline{\lambda} = \lambda$ c'est-à-dire $\lambda \in \mathbb{R}$. □

Théorème 51.

> *Si f est un endomorphisme symétrique il existe une base orthonormée formée de vecteurs propres de f. Autrement dit f est diagonalisable dans une base orthonormée.*

Démonstration. On va procéder par récurrence sur la dimension de l'espace E. Si $\dim(E) = 1$ le résultat est trivial. On suppose que le résultat a été établi si $\dim(E) = n$. On considère donc f un endomorphisme symétrique sur un espace de dimension $n+1$. Le polynôme caractéristique de f étant scindé sur $\mathbb{R}$ il existe donc au moins une valeur propre λ dans $\mathbb{R}$. Soit alors $x \neq 0$ tel que $f(x) = \lambda x$. On pose $e_1 = x/||x||$. Soit $F = \mathbb{R}e_1$ le sous-espace vectoriel de E engendré par e_1. F est stable par f et donc $F^\perp$ aussi, et $\dim(F^\perp) = n$. La restriction de f à $F^\perp$ est un endomorphisme symétrique sur un espace de dimension n, donc d'après l'hypothèse de récurrence il existe une base orthonormée $(e_2, \ldots, e_n)$ de $F^\perp$ qui diagonalise cette restriction. Alors $(e_1, e_2, \ldots, e_n)$ est une base orthonormée de E qui diagonalise f. □

Corollaire 52. *Si M est une matrice symétrique, il existe $P \in O(n)$ telle que ${}^tPMP$ soit diagonale.*

Remarque 53. *Si x et y sont deux vecteurs propres associées à des valeurs propres λ et μ différentes alors x et y sont orthogonaux. En effet, on a*

$$\lambda(x,y) = (f(x),y) = (x,f(y)) = \mu(x,y).$$

et donc $(\lambda-\mu)(x,y) = 0$ c'est-à-dire $(x,y) = 0$ car $\lambda \neq \mu$.

V. Adjoint d'un endomorphisme

Proposition 54. *Si u est un endomorphisme de E, il existe un unique endomorphisme v de E tel que $\forall x \in E,\ \forall y \in E, \quad (u(x),y) = (x,v(y))$. Cet endomorphisme est appelé* adjoint *de u et est noté u^*.*

DÉMONSTRATION. Soit $(e_1,\ldots,e_n)$ une base orthonormée de E. Pour tout $y \in E$ on veut $(u(e_i),y) = (e_i,v(y))$. Donc si v existe il est déterminé de manière unique par

$$\forall y \in E, v(y) = \sum_{i=1}^n (e_i,v(y))e_i = \sum_{i=1}^n (u(e_i),y)e_i.$$

Supposons maintenant que v soit défini comme ci-dessus. Alors v est un endomorphisme de E et on a

$$\begin{aligned}
(u(x),y) &= \left(\sum_{i=1}^n (x,e_i)u(e_i),y\right) = \sum_{i=1}^n (x,e_i)(u(e_i),y) \\
(x,v(y)) &= \left(\sum_{i=1}^n (x,e_i)e_i, \sum_{i=1}^n (u(e_i),y)e_i\right) = \sum_{i=1}^n (x,e_i)(u(e_i),y)
\end{aligned}$$

c'est-à-dire $(u(x),y) = (x,v(y))$, ce qui termine la démonstration. □

Proposition 55. *Si $\mathcal{B}$ est une base orthonormée alors la matrice de u^* dans cette base est égale à la transposée de la matrice de u c'est-à-dire : $\mathcal{M}_{\mathcal{B}}(u^*) = {}^t\mathcal{M}_{\mathcal{B}}(u)$.*

DÉMONSTRATION. Soit $\mathcal{B} = (e_1,\ldots e_n)$ une base orthonormée de E. On note $\mathcal{M}_{\mathcal{B}}(u^*) = (m^*_{ij})$ et $\mathcal{M}_{\mathcal{B}}(u) = (m_{ij})$ les matrices de u^* et u dans cette base. On a $m^*_{ij} = (e_i,u^*(e_j)) = (u(e_i),e_j) = m_{ji}$, et donc le résultat. □

Remarque 56. *On a $\mathcal{M}_{\mathcal{B}}\left((u^*)^*\right) = {}^t\mathcal{M}_{\mathcal{B}}(u^*) = {}^t\left({}^t\mathcal{M}_{\mathcal{B}}(u)\right) = \mathcal{M}_{\mathcal{B}}(u)$, et donc $(u^*)^* = u$.*

Corollaire 57.
u est orthogonal si et seulement si $u^* = u^{-1}$.
u est symétrique si et seulement si $u^* = u$.

DÉMONSTRATION. Ce corollaire se déduit aisément de la proposition ci-dessus et des propriétés des endomorphismes orthogonaux et symétriques. □

Proposition 58.

- **i :** *Si u est un endomorphisme de E alors $\mathrm{Ker}(u) = (\mathrm{Im}(u^*))^\perp$ et $\mathrm{Im}(u) = (\mathrm{Ker}(u^*))^\perp$.*
- **ii :** *Si F est sous-espace vectoriel stable par u alors $F^\perp$ est stable par u^**
- **iii :** *Si u et v sont deux endomorphismes de E alors $(u+v)^* = u^* + v^*$ et $(u \circ v)^* = v^* \circ u^*$.*

Démonstration. Soit y dans $\mathrm{Ker}(u)$. Pour tout $z \in \mathrm{Im}(u^*)$ il existe $x \in E$ tel que $z = u^*(x)$. On a donc

$$(y, z) = (y, u^*(x)) = (u(y), x) = (0, x) = 0$$

et $\mathrm{Ker}(u) \subset (\mathrm{Im}(u^*))^\perp$. Si $y \in (\mathrm{Im}(u^*))^\perp$, alors pour tout $x \in E$ on a $(u(y), x) = (y, u^*(x)) = 0$ c'est-à-dire $u(y) = 0$ et $(\mathrm{Im}(u^*))^\perp \subset \mathrm{Ker}(u)$, ce qui termine la démonstration du premier point du i. D'après ce résultat en passant à l'orthogonal on obtient $(\mathrm{Ker}(u))^\perp = \mathrm{Im}(u^*)$. En appliquant ceci à u^* et en utilisant le fait que $(u^*)^* = u$ on trouve le deuxième point.
Pour le ii on considère $x \in F^\perp$. Alors on a $(y, u^*(x)) = (u(y), x) = 0$ car F est stable par u. Donc $F^\perp$ est stable par u^*.
Le troisième point se montre facilement à l'aide des matrices. □

VI. Espaces euclidiens de dimension 2

Dans ce qui suit on se fixe une base orthonormée $\mathcal{B}_0$ (en général la base canonique) qui détermine l'orientation des espaces euclidiens que l'on considère.

VI.1. Étude de $O_2(\mathbb{R})$.

Théorème 59.

Soit M une matrice de $\mathcal{M}_2(\mathbb{R})$. Alors les assertions suivantes sont équivalentes :

a) $$M \in O_2(\mathbb{R}),$$

b) $$\exists \alpha \in \mathbb{R} \ / \ M = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{ou} \quad M = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix},$$

c) $$\exists a,\ b \in \mathbb{R} \text{ tel que } a^2 + b^2 = 1,\ \exists \varepsilon \in \{-1,\ 1\}, \quad M = \begin{pmatrix} a & -\varepsilon b \\ b & \varepsilon a \end{pmatrix}.$$

Démonstration. Soit

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

une matrice de $\mathcal{M}_2(\mathbb{R})$. Si $M \in O_2(\mathbb{R})$ alors ses vecteurs lignes (ou colonnes) forment une base orthonormée, c'est-à-dire qu'on a : $a^2 + b^2 = 1,\ c^2 + d^2 = 1$ et $ac + bd = 0$.
Des deux premières équations on déduit qu'il existe $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}$ tels que $a = \cos(\alpha)$, $b = \sin(\alpha)$, $c = \sin(\beta)$ et $d = \cos(\beta)$. En injectant ceci dans la troisième égalité on trouve $\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) = 0$, c'est-à-dire $\sin(\alpha + \beta) = 0$, ou encore $\alpha + \beta = 0 \bmod \pi$. Ainsi si $\alpha = -\beta \bmod 2\pi$ alors

$$M = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

et si $\alpha = \pi - \beta \bmod 2\pi$,

$$M = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix},$$

donc $a)$ entraîne $b)$.

Soit maintenant M une matrice de la forme de celles données dans $b)$. En posant $a = \cos(\alpha)$ et $b = \sin(\alpha)$, on a $a^2 + b^2 = 1$ et en prenant $\varepsilon \in \{-1,\ 1\}$ suivant le cas, on retrouve $c)$.

Enfin si M est de la forme donnée dans $c)$ on a

$$M\,{}^tM = \begin{pmatrix} a & -\varepsilon b \\ b & \varepsilon a \end{pmatrix} \begin{pmatrix} a & b \\ -\varepsilon b & \varepsilon a \end{pmatrix} = \begin{pmatrix} a^2 + \varepsilon^2 b^2 & ab - \varepsilon^2 ab \\ ab - \varepsilon^2 ab & \varepsilon^2 a^2 + b^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

car $a^2 + b^2 = 1$ et $\varepsilon^2 = 1$. Donc $M \in O_2(\mathbb{R})$. □

Nous allons maintenant donner un résultat dans lequel on donne les propriétés des matrices de $O_2(\mathbb{R})$. Pour α un réel on définit les matrices de $O_2(\mathbb{R})$

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{et} \quad S_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Proposition 60. *Avec les notations définies ci-dessus on a : $\forall \alpha$ et $\beta \in \mathbb{R}$,*

$$\begin{array}{lcllcl} (R_\alpha)^{-1} & = & R_{-\alpha} & R_\alpha R_\beta & = & R_{\alpha+\beta} = R_\beta R_\alpha \\ (S_\alpha)^{-1} & = & S_\alpha & S_\alpha S_\beta & = & R_{\alpha-\beta} \\ R_\alpha S_\beta & = & S_{\alpha+\beta} & S_\beta R_\alpha & = & S_{\beta-\alpha} \end{array}$$

On peut remarquer que les matrices R_α sont dans $SO_2(\mathbb{R})$ et les matrices S_α sont dans $O_2^-(\mathbb{R}) = O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$. D'après la proposition ci-dessus les premières commutent entre elles mais pas les secondes.

DÉMONSTRATION. La démonstration se fait directement en utilisant le produit matriciel et les propriétés des fonctions trigonométriques. □

VI.2. Étude de $SO(E)$.

Soit E un espace euclidien de dimension 2 orienté par la base $(\vec{i}, \vec{j})$.

Proposition 61. *Soient $\mathcal{B}$ une base orthonormée directe de E et $f \in SO(E)$, alors il existe $\alpha \in \mathbb{R}$ tel que*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

De plus dans toute base orthonormée directe la matrice de f est exactement celle donnée ci-dessus et dans toute base orthonormée indirecte la matrice de f est

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

DÉMONSTRATION. Soit $\mathcal{B}$ une base orthonormée directe de E. Si f appartient à $SO(E)$ alors $M_{\mathcal{B}}(f)$ appartient à $SO_2(\mathbb{R})$ c'est-à-dire qu'il existe $\alpha \in \mathbb{R}$ tel que

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Soit maintenant $\mathcal{B}'$ une autre base orthonormée directe de E. Comme $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthonormées de E, la matrice de passage appartient à $O_2(\mathbb{R})$. De plus elle sont toutes les deux directes, donc le déterminant de la matrice de passage vaut 1, elle appartient donc à $SO_2(\mathbb{R})$. Elle est donc de la forme R_θ pour un certain $\theta \in \mathbb{R}$. La matrice de f dans la nouvelle base s'écrit donc

$$M_{\mathcal{B}'}(f) = (R_\theta)^{-1} M_{\mathcal{B}}(f) R_\theta = (R_\theta)^{-1} R_\alpha R_\theta = R_\alpha = M_{\mathcal{B}}(f),$$

d'après la Proposition 2.
Soit maintenant $\mathcal{B}''$ une base orthonormée indirecte de E. Donc la matrice de passage appartient à $O_2^-(\mathbb{R})$, c'est-à-dire qu'elle a la forme S_θ pour un certain $\theta \in \mathbb{R}$. La matrice de f dans la nouvelle base s'écrit donc

$$M_{\mathcal{B}''}(f) = (S_\theta)^{-1} M_{\mathcal{B}}(f) S_\theta = (S_\theta)^{-1} R_\alpha S_\theta = R_{-\alpha} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

d'après la Proposition 2. □

Définition 62. *Soit $f \in SO(E)$. Alors il existe $\alpha \in \mathbb{R}$ tel que sa matrice dans n'importe quelle base orthonormée directe de E est*

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

On dit que f est la rotation d'angle (orienté) α.

Proposition 63. *Soient u et v deux vecteurs unitaires de E. Alors, il existe une unique rotation $R_\alpha \in SO(E)$ telle que $R_\alpha(u) = v$. On dit alors que l'angle formé par u et v est l'angle α.*

DÉMONSTRATION. Si le vecteur u est fixé on peut alors trouver une base orthonormée directe (u, u_1) de E. Soit alors $v = au + bu_1$ écrit dans cette base. Si la rotation R existe on doit avoir $R(u) = v$. Donc la matrice de R dans cette base est

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Cette application convient : en effet, on a $R(u) = v$, et $1 = ||v||^2 = a^2 + b^2$. C'est-à-dire qu'il existe $\alpha \in \mathbb{R}$ tel que $a = \cos(\alpha)$ et $b = \sin(\alpha)$. Donc la matrice de R dans la base orthonormée directe est

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

qui est celle de la rotation d'angle α. □

Proposition 64. *Soit α l'angle formé par deux vecteurs unitaires u et v de E. Alors*

$$\begin{cases} \cos(\alpha) &= u \cdot v \\ \sin(\alpha) &= \det(u, v) \end{cases}$$

où $u \cdot v$ est le produit scalaire de u et v.

DÉMONSTRATION. D'après la proposition précédente v est l'image de u par une rotation d'angle α. De plus la démonstration montre que dans une base orthonormée directe de premier vecteur u, v a pour coordonnées

$$\begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix};$$

et u évidemment

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Ainsi lorsqu'on calcule le produit scalaire $u \cdot v$ on obtient $\cos(\alpha)$ et lorsqu'on calcule $\det(u, v)$ on obtient $\sin(\alpha)$. □

VI.3. Étude de $O^-(E) = O(E) \setminus SO(E)$.

Proposition 65. *Soient $\mathcal{B}$ une base orthonormée de E et $f \in O^-(E)$, alors il existe $\alpha \in \mathbb{R}$ tel que*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix},$$

et f est une symétrie orthogonale par rapport à la droite de ses vecteurs invariants. *Réciproquement, toute symétrie orthogonale appartient à $O^-(E)$.*

DÉMONSTRATION. Soit $f \in O^-(E)$. Donc la matrice de f dans la base $\mathcal{B}$ appartient à $O_2^-(\mathbb{R})$,c'est-à-dire qu'il existe $\alpha \in \mathbb{R}$ tel que

$$M_{\mathcal{B}}(f) = S_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Le polynôme caractéristique est donc égal à

$$\chi_f(X) = \det(S_\alpha - X\mathrm{I}_2) = X^2 - 1.$$

Il y a deux valeurs propres 1 et -1. Lorsqu'on détermine le sous-espace propre associé à $\lambda = 1$ on trouve :

$$\left\{\begin{array}{l}(\cos(\alpha)-1)x+\sin(\alpha)y=0\\ \sin(\alpha)x-(\cos(\alpha)+1)y=0\end{array}\right. \iff \left\{\begin{array}{l}-2\sin^2(\frac{\alpha}{2})x+2\sin(\frac{\alpha}{2})\cos(\frac{\alpha}{2})y=0\\ 2\sin(\frac{\alpha}{2})\cos(\frac{\alpha}{2})x-2\cos^2(\frac{\alpha}{2})y=0\end{array}\right.$$
$$\iff \left\{ -\sin(\tfrac{\alpha}{2})x+\cos(\tfrac{\alpha}{2})y = 0\right.$$

qui est l'équation *de la droite invariante*. Une base de E_1 est donc donnée par

$$u_1 = \begin{pmatrix}\cos(\frac{\alpha}{2})\\ \sin(\frac{\alpha}{2})\end{pmatrix}.$$

De même on montre que E_{-1}, le sous-espace propre associé à $\lambda = -1$ est déterminé par

$$\left\{ \cos(\tfrac{\alpha}{2})x+\sin)\tfrac{\alpha}{2})y = 0\right.$$

c'est-à-dire qu'une de ses bases est donnée par

$$u_2 = \begin{pmatrix}-\sin(\frac{\alpha}{2})\\ \cos(\frac{\alpha}{2})\end{pmatrix}.$$

La matrice de f dans (u_1, u_2) s'écrit alors

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

f est donc une symétrie orthogonale par rapport à E_1, c'est-à-dire par rapport à la droite d'équation $-\sin(\dfrac{\alpha}{2})x+\cos(\dfrac{\alpha}{2})y = 0$.
Réciproquement, si f est une symétrie orthogonale, dans une base orthonormée $\mathcal{B}$ convenable sa matrice s'écrit

$$M_{\mathcal{B}}(f) = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Donc $M_{\mathcal{B}}(f) \in O_2^{-1}(\mathbb{R})$ c'est-à-dire que f appartient à $O^{-1}(E)$. □

Proposition 66. *Toute rotation se décompose en produit de deux symétries orthogonales, la première par rapport à une droite quelconque et la seconde par rapport à une droite faisant un angle $\alpha/2$ avec la première droite.*

DÉMONSTRATION. Soit (u, v) une base orthonormée directe de E. La matrice de le rotation d'angle α s'écrit dans cette base

$$R_\alpha = \begin{pmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{pmatrix}.$$

On a alors

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

où la première matrice est celle d'une symétrie orthogonale par rapport à une droite faisant un angle $\alpha/2$ avec u et la seconde est celle de la symétrie orthogonale par rapport à u. □

Corollaire 67. *Les symétries orthogonales engendrent $O(E)$.*

VII. Espaces euclidiens de dimension 3

Soit E un espace euclidien de dimension 3, orienté par la base $(\vec{i}, \vec{j}, \vec{k})$.

VII.1. Produit mixte.

Définition 68. *Soit $\mathcal{B}$ une base orthonormée directe de E. Soient u, v et w trois vecteurs de E. Alors le déterminant de (u,v,w) dans $\mathcal{B}$, $\det_{\mathcal{B}}(u,v,w)$ ne dépend pas de la base orthonormée choisie. Il est appelé* produit mixte *u, v et w, et est noté $[u,v,w]$.*

Donnons quelques propriétés du produit mixte.

$$\begin{array}{rcl} E^3 & \to & \mathbb{R} \\ (u,v,w) & \mapsto & [u,v,w] \end{array}$$

est une application *trilinéaire alternée*, ce qui signifie que l'application est linéaire par rapport à chacune des variables, qu'elle est antisymétrique lorsqu'on échange deux variables, et qu'elle est nulle si et seulement si (u,v,w) sont liés.

Si on note

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \text{ et } \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}$$

les matrices colonnes des vecteurs u, v et w dans une base orthonormée directe $\mathcal{B}$ de E alors

$$[u,v,w] = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

VII.2. Produit vectoriel.

Si l'on développe le déterminant précédent par rapport à la troisième colonne on obtient

$$[u,v,w] = x_3 \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} - y_3 \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix},$$

ce qui correspond au produit scalaire dans la base orthonormée directe $\mathcal{B}$, des vecteurs de coordonnées respectives

$$\begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} \text{ et } \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ z_1 x_2 - z_2 x_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Définition 69. *Si u et v sont deux vecteurs de E, le vecteur défini comme ci-dessus est appelé* produit vectoriel *de u et v et est noté $u \wedge v$.*

D'après la construction ci-dessus on a $\forall w \in E$,

$$(u \wedge v, w) = (w, u \wedge v) = [u,v,w].$$

De plus l'application

$$\begin{array}{rcl} E^2 & \to & E \\ (u,v) & \mapsto & u \wedge v \end{array}$$

est bilinéaire alternée telle que $u \wedge v = 0$ si et seulement si u et v sont liés.

Proposition 70. *Si u et v sont libres alors $u \wedge v$ appartient à $\mathrm{Vect}(u,v)^{\perp}$ (l'orthogonal du sous-espace vectoriel engendré par u et v). De plus $(u, v, u \wedge v)$ forme une base directe de E.*

DÉMONSTRATION. On a par définition

$$\begin{array}{rcl} (u \wedge v, u) & = & [u,v,u] = 0 \\ (u \wedge v, v) & = & [u,v,v] = 0 \end{array}$$

ce qui montre le premier point de la proposition. Soit maintenant $\mathcal{B}$ une base orthonormée directe de E. Par définition du produit mixte on a, en notant $P_{\mathcal{B}\to\mathcal{B}'}$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}' = (u, v, u \wedge v)$,

$$\det_{\mathcal{B}}(P_{\mathcal{B}\to\mathcal{B}'}) = [u, v, u \wedge v] = (u \wedge v, u \wedge v) = ||u \wedge v||^2 \geq 0$$

c'est-à-dire que $\mathcal{B}'$ est une base directe de E. □

Proposition 71. *Si u et v sont libres alors l'angle α formé par les vecteurs u et v est mesuré dans le plan $\mathcal{P} = \mathrm{Vect}(u, v)$ et on a*

$$\begin{aligned} \cos(\alpha) &= \frac{(u, v)}{||u||\,||v||} \\ \sin(\alpha) &= \frac{[u, v, w]}{||u||\,||v||\,||w||} \end{aligned}$$

où w est un vecteur normal à $\mathcal{P}$ avec lequel on oriente $\mathcal{P}$.

DÉMONSTRATION. On pose $e_1 = u/||u||$, $e_3 = w/||w||$ et e_2 est un vecteur de $\mathcal{P}$ tel que (e_1, e_2, e_3) forme une base orthonormée directe de E, c'est-à-dire que $e_3 = e_1 \wedge e_2$. Comme v appartient à $\mathcal{P}$ et fait un angle α avec u et donc aussi avec e_1, on a $v/||v|| = \cos(\alpha)e_1 + \sin(\alpha)e_2$. Ainsi

$$\frac{(u, v)}{||u||\,||v||} = (e_1, \cos(\alpha)e_1 + \sin(\alpha)e_2) = \cos(\alpha)$$

et

$$\frac{[u, v, w]}{||u||\,||v||\,||w||} = \left[\frac{u}{||u||}, \frac{v}{||v||}, \frac{w}{||w||}\right] = [e_1, \cos(\alpha)e_1 + \sin(\alpha)e_2, e_3] = \sin(\alpha)$$

ce qui montre la proposition. □

Corollaire 72. *Si u et v sont libres alors $u \wedge v = ||u||\,||v|| \sin(\alpha) w$, où w est un vecteur unitaire normal à $\mathcal{P}$ avec lequel on oriente $\mathcal{P}$ et α est l'angle formé par les vecteurs u et v.*

DÉMONSTRATION. Cela provient de la deuxième relation de la proposition précédente en prenant $w = u \wedge v$. □

Proposition 73. FORMULE DU DOUBLE PRODUIT VECTORIEL.

$$\forall u,\ v, w \in E, \quad u \wedge (v \wedge w) = (u, w)v - (u, v)w.$$

DÉMONSTRATION. Cela ce montre directement en écrivant les formules en fonction des coordonnées dans une base orthonormée de E. □

VII.3. Étude de $O_3(\mathbb{R})$.

On rappelle que si $M \in O_3(\mathbb{R})$ alors les seules valeurs propres possibles pour M sont 1 et -1.

Théorème 74.

Soit M une matrice de $O_3(\mathbb{R})$. Alors M est semblable *à l'une des deux matrices suivantes*

$$\begin{pmatrix} a & c & 0 \\ b & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ou \quad \begin{pmatrix} a & c & 0 \\ b & d & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad avec \quad \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O_2(\mathbb{R}).$$

Démonstration. Soit E un espace euclidien de dimension 3 rapporté à une base $\mathcal{B}$ orthonormée. On considère $f \in O(E)$ dont la matrice dans $\mathcal{B}$ est la matrice M. Le polynôme caractéristique est de degré 3. Il a donc au moins une valeur propre qui est 1 ou -1. Soit e_3 un vecteur propre associé à cette valeur propre, de norme 1. On considère alors e_1 et $_2$ tels que $\mathcal{B}' = (e_1, e_2, e_3)$ forme une base orthonormée de E. Par conservation du produit scalaire, comme e_1 et e_3, et e_2 et e_3 sont orthogonaux, il en est de même pour $f(e_1)$ et $f(e_3)$, et $f(e_2)$ et $f(e_3)$. Or $f(e_3) = \pm e_3$, c'est-à-dire que $f(e_1) = ae_1 + be_2$ et $f(e_2) = ce_1 + de_2$. Donc la matrice de f dans $\mathcal{B}'$ s'écrit

$$\begin{pmatrix} a & c & 0 \\ b & d & 0 \\ 0 & 0 & \varepsilon \end{pmatrix} \qquad \text{avec} \qquad \varepsilon = \pm 1.$$

De plus comme $M \in O_3(\mathbb{R})$ on a $a^2 + b^2 = c^2 + d^2 = 1$ et $ac + bd = 0$ c'est-à-dire que

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O_2(\mathbb{R}).$$

□

Étude des différents cas.
Premier cas : M est semblable à

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

dans (e_1, e_2, e_3).
On dit que f est une *rotation vectorielle* autour de $\mathrm{Vect}(e_3)$. Si l'espace est orienté et si on oriente $\mathrm{Vect}(e_3)$ par e_3 on dira que f est une rotation d'angle α autour de $\mathrm{Vect}(e_3)$. f appartient à $SO(E)$ ou encore M appartient à $SO_3(\mathbb{R})$.

Deuxième cas : M est semblable à

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

dans (e_1, e_2, e_3).
La restriction de f au plan engendré par e_1 et e_2 est une symétrie orthogonale par rapport à une droite. Il existe donc une base orthonormée (e'_1, e'_2) dans laquelle la restriction de f a pour matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Dans la base orthonormée (e_3, e'_1, e'_2) la matrice de f s'écrit alors

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

f est donc une *symétrie orthogonale plane* par rapport au plan vectoriel $\mathrm{Vect}(e_3, e'_1)$, on dit encore que f est une *réflexion*. f appartient à $O^-(E)$ ou encore M appartient à $O_3^-(\mathbb{R})$.

Troisième cas : M est semblable à

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

dans (e_1, e_2, e_3).
La restriction de f au plan engendré par e_1 et e_2 est une symétrie orthogonale par rapport à une droite. Il existe donc une base orthonormée (e_1', e_2') dans laquelle la restriction de f a pour matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Dans la base orthonormée (e_1', e_2', e_3) la matrice de f s'écrit alors

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

f est donc une symétrie orthogonale par rapport à la droite dirigée par $\mathrm{Vect}(e_1)$, ou encore f est une rotation d'angle π autour de $\mathrm{Vect}(e_1)$, on dit que f est un *retournement d'axe* $\mathrm{Vect}(e_1)$. f appartient à $SO(E)$ ou encore M appartient à $SO_3(\mathbb{R})$.

Quatrième cas : M est semblable à

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

dans (e_1, e_2, e_3).
On a donc

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

f est la composée commutative d'une symétrie orthogonale par rapport au plan $\mathrm{Vect}(e_1, e_2)$ et d'une rotation d'angle α autour de e_3. f appartient à $O^-(E)$ ou encore M appartient à $O_3^-(\mathbb{R})$.
Si $\alpha \neq 0 \bmod 2\pi$, f n'a que le vecteur nul comme vecteur invariant, sinon $f = \mathrm{Id}_E$. Si $\alpha = \pi \bmod 2\pi$ alors $f = -\mathrm{Id}_E$, on dit que f est *une symétrie centrale* ou f est une *homothétie* de rapport -1

On obtient la conclusion suivante.

Proposition 75. *Soit $f \in O(E)$ alors, soit $f \in SO(E)$ et f est une rotation vectorielle (une droite invariante), soit $f \in O^-(E)$ et*

- *f est* une réflexion*, i.e. f est une symétrie plane (un plan invariant)*

ou

- *f est la composée des deux cas précédents (0 seul vecteur invariant).*

Remarquons que ce résultat signifie que les rotations et les réflexions engendrent $O(E)$.

VII.4. Caractérisation des rotations.

Soit E un espace euclidien de dimension 3 orienté. On note $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ une base orthonormée directe de E. Soit f un endomorphisme de E de matrice $M_{\mathcal{B}}(f)$ dans la base $\mathcal{B}$.

D'après la proposition précédente f est une rotation si et seulement si sa matrice vérifie

$${}^t M_{\mathcal{B}}(f) M_{\mathcal{B}}(f) = \mathrm{I}_3 \quad \text{et} \quad \det(M_{\mathcal{B}}(f)) = 1.$$

La première condition signifie que f est orthogonal, c'est-à-dire que f appartient à $O(E)$, et alors la deuxième précise ceci en disant que f appartient à $SO(E)$ ce qui caractérise bien les rotations de l'espace.

On suppose maintenant que f est une rotation de E et on va donner ses éléments caractéristiques.

a) L'axe de la rotation.
C'est l'ensemble des vecteurs invariants par f. Il est donc déterminé par l'équation

$$M_{\mathcal{B}}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

b) L'angle de la rotation.
Pour déterminer un angle il faut et il suffit de déterminer son cosinus et son sinus. Dans un premier temps on calcule la *trace* (= somme des coefficients diagonaux) de la matrice qui vérifie

$$\mathrm{tr}(M_{\mathcal{B}}(f))=1+2\cos(\alpha).$$

En effet $M_{\mathcal{B}}(f)$ est semblable à $\begin{pmatrix}\cos(\alpha) & -\sin(\alpha) & 0\\ \sin(\alpha) & \cos(\alpha) & 0\\ 0 & 0 & 1\end{pmatrix}$ et la trace de deux matrices semblables est identique.
Pour le sinus, on considère un vecteur e_3 qui est un vecteur directeur de l'axe de la rotation. On suppose que ce vecteur oriente l'axe. Soit alors e_1 un vecteur perpendiculaire à e_3, alors d'après les propriétés du produit mixte et du produit vectoriel on a

$$\sin(\alpha)=\frac{[e_1,f(e_1),e_3]}{||e_1||\,||f(e_1)||\,||e_3||}=\frac{[e_1,f(e_1),e_3]}{||e_1||^2\,||e_3||},$$

car une rotation préserve les normes.

Exemple 76. *On suppose que la matrice de f dans la base $\mathcal{B}$ est*

$$M=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}.$$

Comme ${}^tMM=\mathrm{I}_3$ et $\det(M_{\mathcal{B}}(f))=1$, f est bien une rotation. Donnons ses éléments caractéristiques. Son axe est déterminé par

$$M=\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\y\\z\end{pmatrix}\iff\left\{\begin{array}{rcl}y&=&x\\z&=&y\\x&=&z\end{array}\right.$$

C'est la droite de base

$$e_3=\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

On suppose que ce vecteur oriente l'axe.
L'angle α de la rotation, vérifie $0=\mathrm{tr}(M)=1+2\cos(\alpha)$, i.e. $\cos(\alpha)=-1/2$. D'autre part, soit

$$e_1=\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

un vecteur orthogonal à e_3. On a

$$f(e_1)=\begin{pmatrix}0\\1\\-1\end{pmatrix}$$

et

$$\sin(\alpha) = \frac{\begin{vmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix}}{(\sqrt{2})^2\sqrt{3}} = -\frac{\sqrt{3}}{2}.$$

Donc $\alpha = -2\pi/3 \bmod 2\pi$.

Soient alors $e'_3 = e_3/||e_3||$, $e'_1 = e_1/||e_1||$ *et* $e'_2 = e'_3 \wedge e'_1$. *Alors la matrice de* f *dans la base orthonormée directe* (e'_1, e'_2, e'_3) *est*

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

VII.5. Décomposition des rotations en produit de 2 réflexions.

Théorème 77.

> *Soit* E *un espace euclidien de dimension 3. On considère* R *la rotation de* E *d'angle* α *et d'axe* Δ *orienté par le vecteur unitaire* e_3. *Soit* $\mathcal{P}$ *un plan vectoriel contenant* e_3 *; on note* S *la réflexion par rapport à ce plan. Alors il existe une et une seule réflexion* S' *telle que* $R = S' \circ S$. *De plus* S' *est la réflexion par rapport au plan* $\mathcal{P}'$ *image de* $\mathcal{P}$ *par une rotation d'axe dirigé par* e_3 *et d'angle* $\alpha/2$.

DÉMONSTRATION. Soit e_1 un vecteur unitaire orthogonal à e_3. On considère alors la base orthonormée directe (e_1, e_2, e_3) où $e_2 = e_3 \wedge e_1$. Dans cette base la matrice de la rotation R est

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

De plus comme S est une symétrie orthogonale plane par rapport au plan $\mathcal{P}$ qui contient e_3, dans cette même base sa matrice est donnée par

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

avec $\beta \in \mathbb{R}$.

Si S' existe, on a $S' = R \circ S$, donc $S'(e_3) = e_3$. C'est-à-dire qu'il existe $\gamma \in \mathbb{R}$ tel que sa matrice dans (e_1, e_2, e_3) soit égale à

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

On obtient donc si S' existe

$$\begin{aligned} M(R) = M(S')M(S) &= \begin{pmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ \sin(\gamma) & -\cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\beta) & \sin(\beta) & 0 \\ \sin(\beta) & -\cos(\beta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\gamma-\beta) & -\sin(\gamma-\beta) & 0 \\ \sin(\gamma-\beta) & \cos(\gamma-\beta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

c'est-à-dire que $\alpha = \gamma - \beta \bmod 2\pi$, ou encore $\gamma = \alpha + \beta \bmod 2\pi$. De plus si

$$M(S') = \begin{pmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) & 0 \\ \sin(\alpha+\beta) & -\cos(\alpha+\beta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

S' est bien une symétrie plane qui vérifie $R = S' \circ S$.
Le plan invariant pour la réflexion S est déterminé par le vecteur e_3 et, d'après l'étude des symétries dans O_2 par le vecteur u_1 de coordonnées

$$\begin{pmatrix} \cos\beta/2 \\ \sin\beta/2 \\ 0 \end{pmatrix}$$

dans la base (e_1, e_2, e_3). De même, le plan invariant pour S' est déterminé par e_3 et le vecteur u_2 de coordonnées

$$\begin{pmatrix} \cos(\beta+\alpha)/2 \\ \sin(\beta+\alpha)/2 \\ 0 \end{pmatrix}.$$

Comme e_3 est l'image de e_3 par la rotation d'axe dirigé par e_3 et d'angle $\alpha/2$ et que u_2 est l'image de u_1 par cette même rotation, on en déduit que le plan invariant de S' est l'image par cette rotation du plan invariant de S. □

Corollaire 78. *Les réflexions engendrent $O(E)$.*

VII.6. Décomposition des rotations en produit de 2 retournements.

Théorème 79.

Soit E un espace euclidien de dimension 3. On considère R la rotation de E d'angle α et d'axe Δ orienté par le vecteur unitaire e_3. Soit Δ' une droite orthogonale à Δ ; on note $S_{\Delta'}$ le retournement par rapport à cette droite. Alors il existe un et un seul retournement $S_{\Delta''}$ telle que $R = S_{\Delta''} \circ S_{\Delta'}$. De plus $S_{\Delta''}$ est le retournement par rapport à la droite Δ'' image de Δ' par la rotation d'axe Δ et d'angle $\alpha/2$.

Démonstration. Soit e_1 un vecteur unitaire orthogonal à e_3 qui dirige Δ'. On considère alors la base orthonormée directe (e_1, e_2, e_3) où $e_2 = e_3 \wedge e_1$. Dans cette base la matrice de la rotation R est

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

De plus comme $S_{\Delta'}$ est un retournement par rapport à la droite Δ' dans cette même base sa matrice est donnée par $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Donc si $S_{\Delta''}$ existe, on a

$$M(S_{\Delta''}) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ \sin(\alpha) & -\cos(\alpha) & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

qui est bien la matrice d'un retournement. De plus si $M(S_{\Delta''})$ est définie comme ci-dessus La restriction de $S_{\Delta''}$ à (e_1, e_2) est une symétrie orthogonale de O_2. Donc la droite invariante de $S_{\Delta''}$ est dirigée par le vecteur u de coordonnées

$$\begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \\ 0 \end{pmatrix}$$

dans (e_1, e_2, e_3), c'est-à-dire par la droite Δ'' image de Δ' par la rotation d'axe Δ et d'angle $\alpha/2$. $\square$

Chapitre 5

ESPACES HERMITIENS

Nous allons définir sur un espace vectoriel complexe E de dimension n une structure analogue à celle d'espace vectoriel euclidien. Un premier point de vue consisterait à considérer, comme dans le chapitre précédent, des formes quadratiques sur E ; mais elles ne permettent pas d'obtenir la notion si utile de norme d'un vecteur. Ce chapitre sera donc consacré à une généralisation différente : les formes sesquilinéaires remplaceront les formes bilinéaires et les formes hermitiennes les formes quadratiques. Les formes hermitiennes définies positives conduisent à un produit scalaire hermitien et à la notion d'espace hermitien où le groupe unitaire joue un rôle analogue au groupe orthogonal de l'espace euclidien.

Les espaces vectoriels considérés seront supposés de dimension finie.

I. Définition et caractérisation des espaces hermitiens

I.1. Rappels sur les nombres complexes.

Le corps des nombres complexes, $\mathbb{C}$, peut être vu comme l'ensemble des objets de la forme $z = a + ib$, avec a et b deux nombres réels, et i vérifiant $i^2 = -1$. On peut alors se donner les règles de calcul suivantes :

$$\begin{aligned} &* \quad (a+ib)+(a'+ib') = (a+a')+i(b+b'),\\ &* \quad (a+ib)(a'+ib') = (aa'-bb')+i(ab'+a'b)\\ &* \quad \lambda(a+ib) = \lambda a + i\lambda b \end{aligned}$$

Enfin, si a ou b est non-nul, on voit que

$$\frac{1}{a+ib} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$$

est l'inverse de $a+ib$. La formule de l'inverse fait apparaître deux nombres intéressants :

* $a - ib$ est appelé le conjugué de $z = a+ib$ et est noté $\overline{z}$;
* $\sqrt{a^2+b^2}$ est appelé le module de $z = a+ib$ et est noté $|z|$(distance du point (a,b) à l'origine $(0,0)$).

On a les propriétés suivantes : pour tout $z = a+ib$ dans $\mathbb{C}$ et tout $z' = a - ib$ dans $\mathbb{C}$

$$\begin{aligned} &* \quad z\overline{z} = (a+ib)(a-ib) = a^2+b^2 = |z|^2,\\ &* \quad \overline{z+z'} = (a+a')-i(b+b') = a-ib+a'-ib' = \overline{z}+\overline{z'},\\ &* \quad \overline{z.z'} = (aa'-bb')-i(ab'+ba') = (a-ib)(a'-ib') = \overline{z}.\overline{z'} \end{aligned}$$

I.2. Formes sesquilinéaires.

Définition 1. *Soit E un $\mathbb{C}$-espace vectoriel. L'application $\varphi : E\times E \to \mathbb{C}$ définie par $(x,x') \mapsto \varphi(x,x')$, est* sesquilinéaire *si elle a les propriétés suivantes :*

- a) φ est linéaire par rapport à sa deuxième variable :

$$\forall x\in E,\quad \forall x'\in E,\quad \forall y'\in E,\quad \forall\lambda\in\mathbb{C},\quad \forall\mu\in\mathbb{C},$$
$$\varphi(x,\lambda x'+\mu y') = \lambda\varphi(x,x')+\mu\varphi(x,y');$$

- b) φ est additive par rapport à sa première variable :

$$\forall x \in E, \quad \forall y \in E, \ \forall x' \in E,$$
$$\varphi(x+y, x') = \varphi(x, x') + \varphi(y, x');$$

- c) Si λ est un nombre complexe et $\overline{\lambda}$ son conjugué on a

$$\forall x \in E, \quad \forall x' \in E,$$
$$\varphi(\lambda x, x') = \overline{\lambda}\varphi(x, x').$$

Si b) et c) sont vérifiés on dit que φ est antilinéaire *ou* semi-linéaire *par rapport à sa première variable.*

Remarquons qu'une forme sesquilinéaire n'est pas une forme bilinéaire.

Définition 2. *Soit φ une forme sesquilinéaire de $E \times E$ dans $\mathbb{C}$.*

— *On dit que φ est* hermitienne *si*

$$\forall x \in E, \ \forall x' \in E, \ \varphi(x, x') = \overline{\varphi(x', x)}.$$

— *On dit que φ est* définie *si*

$$\forall x \in E, \ \varphi(x, x) = 0 \iff x = 0.$$

— *On dit que φ est* positive *si*

$$\forall x \in E, \ \varphi(x, x) \geq 0.$$

On peut remarquer que la notion de forme sesquilinéaire hermitienne correspond à la notion de forme bilinéaire symétrique, même si elle sont évidemment différentes.

Définition 3. *A toute forme sesquilinéaire hermitienne φ sur $E \times E$ on peut associer l'application :*

$$\Phi : x \mapsto \Phi(x) = \varphi(x, x),$$

qui va de E dans $\mathbb{R}$ (car $\varphi(x, x) = \overline{\varphi(x, x)}$ entraîne que $\Phi(x) = \varphi(x, x) \in \mathbb{R}$). On dit que Φ est la forme quadratique hermitienne *associée à la forme sesquilinéaire hermitienne φ.*

Proposition 4. *Soit Φ une forme quadratique hermitienne associée à la forme sesquilinéaire hermitienne φ. On a : $\forall x \in E, \ \forall x' \in E$ et $\forall \lambda \in \mathbb{C}$*

$$\begin{array}{ll} i) & \Phi(\lambda x) = |\lambda|^2 \Phi(x) \\ ii) & \varphi(x, x') = \dfrac{1}{2}\left[\Phi(x+x') + i\Phi(x - ix') - (1+i)(\Phi(x) + \Phi(x'))\right] \\ iii) & \varphi(x, x') = \dfrac{1}{4}\left[\Phi(x+x') - \Phi(x-x') - i(\Phi(x+ix') - \Phi(x-ix'))\right] \end{array}$$

Démonstration. Pour le premier point on a pour tout x dans E,

$$\Phi(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda\overline{\lambda}\varphi(x, x) = |\lambda|^2 \Phi(x).$$

Pour le second point on utilise le fait que $\varphi(x, x')$ étant dans $\mathbb{C}$, on peut l'écrire sous la forme $A + iB$ avec A et B deux réels, de sorte que

$$\varphi(x', x) = \overline{\varphi(x, x')} = A - iB.$$

On obtient alors

$$\left\{ \begin{array}{ll} a) \ \Phi(x+x') = \Phi(x) + \Phi(x') + 2A & b) \ \Phi(x-x') = \Phi(x) + \Phi(x') - 2A \\ c) \ \Phi(x+ix') = \Phi(x) + \Phi(x') - 2B & d) \ \Phi(x-ix') = \Phi(x) + \Phi(x') + 2B \end{array} \right.$$

En ajoutant a) et d) multiplié par i, on obtient ii). Puis en soustrayant a) et b) que l'on ajoute à la différence de d) et c) multipliée par i, on trouve iii). □

Cette proposition montre que si deux formes sesquilinéaires hermitiennes sont associées à une même forme quadratique hermitienne alors elles sont égales.

I.3. Produit scalaire hermitien.

Définition 5. *On appelle* produit scalaire hermitien *sur un espace vectoriel complexe E, une forme sesquilinéaire hermitienne définie positive, c'est-à-dire une application $\varphi : E \times E \to \mathbb{C}$ qui vérifie*

a) $x \mapsto \varphi(x, x')$ *est antilinéaire ou semi-linéaire (pour x' fixé) ;*

b) $x' \mapsto \varphi(x, x')$ *est linéaire (pour x fixé) ;*

c) $\varphi(x, x') = \overline{\varphi(x', x)}$ *pour tout x et x' dans E ;*

d) $\varphi(x, x) \geq 0$ *pour tout x dans E ;*

e) $\varphi(x, x) = 0 \iff x = 0$.

On notera le produit scalaire hermitien $\varphi(x, y) = (x|y)$, pour le distinguer du produit scalaire usuel.

Définition 6. *Soit φ un produit scalaire hermitien sur un espace vectoriel complexe E, dont on note Φ sa forme quadratique hermitienne associée. L'application de $N : E \to \mathbb{R}^+$ qui à x associe*

$$N(x) = \sqrt{\Phi(x)} = \sqrt{\varphi(x, x)}$$

est appelée norme hermitienne *associée à φ. On la note $||x|| = N(x)$.*

Définition 7. *Un espace vectoriel complexe E sur lequel on a défini un produit scalaire hermitien (et donc une norme hermitienne associée) s'appelle un* espace hermitien.

Théorème 8.

Dans l'espace vectoriel hermitien E le produit scalaire hermitien et la norme hermitienne vérifient les conditions suivantes :

a) $||x|| = 0 \iff x = 0$ *;*

b) $\forall x \in E,\ \forall \lambda \in \mathbb{C},\quad ||\lambda x|| = |\lambda|\,||x||$ *;*

c) $\forall x \in E,\ \forall y \in E,\ |(x|y)| \leq ||x||\,||y||$ *(inégalité de Cauchy-Schwarz) ;*

d) $\forall x \in E,\ \forall y \in E,\ ||x + y|| \leq ||x|| + ||y||$ *(inégalité triangulaire ou de Minkowsky).*

Démonstration.

a) On a

$$||x|| = 0 \iff 0 = ||x||^2 = (x|x) \iff x = 0,$$

car le produit scalaire hermitien est défini.

b) On a

$$||\lambda x||^2 = (\lambda x|\lambda x) = \lambda\overline{\lambda}(x|x) = |\lambda|^2||x||^2,$$

et donc en prenant la racine on obtient le résultat.

c) Si $y = 0$ l'inégalité est vérifié, c'est même un cas où il y a égalité. Supposons $y \neq 0$. Pour tout $\lambda \in \mathbb{C}$, on a $(x - \lambda y|x - \lambda y) \geq 0$, c'est-à-dire en développant :

$$(x|x) - \overline{\lambda}(y|x) - \lambda(x|y) + \lambda\overline{\lambda}(y|y) \geq 0.$$

Si on prend $\lambda = (y|x)/(y|y)$, compte tenu de $(y|x) = \overline{(x|y)}$, on obtient

$$(x|x) - \frac{|(x|y)|^2}{(y|y)} - \frac{|(x|y)|^2}{(y|y)} + \frac{|(x|y)|^2}{(y|y)} \geq 0,$$

et donc c) en multipliant par $(y|y)$. De plus l'égalité dans c) entraîne pour, la valeur $\lambda = (y|x)/(y|y)$, $(x - \lambda y|x - \lambda y) = 0$, donc $x - \lambda y = 0$ et les deux vecteurs sont liés. Ils le sont aussi si $y = 0$.

Réciproquement, si x et y sont liés et $y \neq 0$ alors il existe λ (il suffit de prendre $\lambda = (y|x)/(y|y)$) tel que $x - \lambda y = 0$, et donc $|(x|y)| = ||x||\,||y||$.
Pour qu'on ait égalité dans Cauchy-Schwarz, il faut et il suffit que les deux vecteurs soient liés.
d) On a

$$||x+y||^2 = (x+y|x+y) = ||x||^2 + ||y||^2 + (x|y) + (y|x),$$

et d'après c)

$$|(x|y) + (y|x)| \leq |(x|y)| + |(y|x)| \leq 2||x||\,||y||$$

et donc

$$||x+y||^2 \leq ||x||^2 + ||y||^2 + 2||x||\,||y|| = (||x|| + ||y||)^2,$$

et l'inégalité d). D'après ce qui précède, pour qu'on ait égalité il faut qu'il ait égalité dans Cauchy-Schwarz. Supposons donc que $y = \lambda x$. On a

$$||x+y||^2 = ||x||^2 + |\lambda|^2||x||^2 + \overline{\lambda}||x||^2 + \lambda||x||^2$$

et

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

En égalisant ces deux expressions on trouve $\overline{\lambda} + \lambda = 2|\lambda|$ ce qui n'est possible que si λ est un réel positif. Réciproquement, si x et y sont liés par un coefficient positif on vérifie que l'on a bien égalité.
Pour qu'on ait égalité dans Minkowsky il faut et il suffit que les deux vecteurs soient liés par un coefficient positif. □

Exemple 9. *Soit E un espace vectoriel de dimension finie n sur $\mathbb{C}$, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. On note*

$$x = \sum_{j=1}^{n} x_j e_j \text{ et } y = \sum_{j=1}^{n} y_j e_j$$

deux vecteurs quelconques de E. L'application $\varphi : E \times E \to \mathbb{C}$, définie par

$$\varphi(x,y) = \sum_{j=1}^{n} \overline{x_j} y_j,$$

est une forme sesquilinéaire hermitienne définie positive. C'est donc un produit scalaire de E. De plus on a

$$(e_j|e_k) = \varphi(e_j, e_k) = \delta_{jk} = \begin{cases} 0 & \text{si } j \neq k \\ 1 & \text{si } j = k \end{cases}.$$

C'est-à-dire que la base $\mathcal{B}$ est orthonormée pour φ.

I.4. Orthogonalité et base orthonormée.

Les démonstrations de ce paragraphe sont identiques à celles du paragraphe correspondant dans le chapitre des espaces euclidiens. Elles sont donc laissées au lecteur.

Définition 10. *On dit que deux vecteurs x et y d'un espace hermitien E sont* orthogonaux *si leur produit scalaire hermitien $(x|y)$ est nul. On note $x \perp y$.*
Une base $(e_1, \ldots, e_n)$ de E est dite orthonormée *si l'on a $(e_j|e_k) = \delta_{jk}$.*

Théorème 11.

> *Pour tout espace vectoriel hermitien E de dimension n il existe une base orthonormée.*

Pour construire une base orthonormée, on peut encore utiliser le procédé d'orthonormalisation de Schmidt.

Théorème 12.

> *Si F est un sous-espace vectoriel d'un espace hermitien E, toute base orthonormée de F peut être prolongée en une base orthonormée de E.*

Définition 13. *On dit que deux sous-espaces vectoriels, F et G, de E sont* orthogonaux *si pour tout x dans F et tout y dans G : $(x|y) = 0$.*
Si F est un sous-espace vectoriel de E, $F^{\perp} = \{y \in E,\ \forall x \in F,\ (x|y) = 0\}$ est un sous-espace vectoriel de E appelé orthogonal *de F.*

Proposition 14. *Soit F est un sous-espace vectoriel d'un espace hermitien E. On a les propriétés suivantes :*
a) $E = F \oplus F^{\perp}$;
b) $(F^{\perp})^{\perp} = F$.

Proposition 15. *Soit E un espace euclidien et $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée de E. On note $x = \displaystyle\sum_{j=1}^{n} x_j e_j$ et $y = \displaystyle\sum_{j=1}^{n} y_j e_j$ deux vecteurs quelconques de E ; alors l'expression du produit scalaire est donné par*

$$(x|y) = \sum_{j=1}^{n} y_j \overline{x_j},$$

et celle de la norme par

$$||x|| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}.$$

Du point de vue matriciel, si on note X et Y les matrices colonnes de x et y dans la base $\mathcal{B}$, on a

$$(x|y) = \ X^* Y \quad \text{et} \quad ||X|| = \sqrt{X^* X},$$

où $X^* = {}^t\big(\overline{X}\big) = \overline{{}^tX}$.

II. Groupe unitaire

II.1. Automorphismes unitaires.

Définition 16. *Soient E un espace hermitien et f un endomorphisme de E. On dit que f est* unitaire *si*

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

Proposition 17. *Si f est unitaire alors f est bijectif, c'est donc un automorphisme.*

Démonstration. Soit x tel que $f(x) = 0$. Alors en appliquant la propriété ci-dessus pour $y = x$ on obtient $0 = ||f(x)||^2 = ||x||^2$, c'est-à-dire que $x = 0$ et donc f est injective. De plus comme E est de dimension finie, f est bijective.

□

Définition 18. *Les automorphismes unitaires sur un espace hermitien E forment un groupe pour la composition des applications. Ce groupe s'appelle* le groupe unitaire *de $\mathcal{L}(E)$ et est noté $U(E)$.*

Proposition 19. *f appartient à $U(E)$ si et seulement si pour tout x dans E $||f(x)|| = ||x||$.*

Démonstration. Si f est unitaire on a pour tout x et y dans E, $(f(x)|f(y)) = (x|y)$, et donc en particulier si $y = x$, $||f(x)||^2 = (f(x)|f(x)) = (x|x) = ||x||^2$, et donc $||f(x)|| = ||x||$. Réciproquement si $||f(x)|| = ||x||$ pour tout x par la formule de polarisation (proposition 4 ii) ou iii)) on obtient que $(f(x)|f(y)) = (x|y)$ pour tout x et y dans E. □

Théorème 20.

> *Soient E un espace hermitien et f un endomorphisme de E. f appartient à $U(E)$ si et seulement l'image d'une base orthonormée quelconque de E par f est encore une base orthonormée de E.*

Démonstration. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E. On note $\mathcal{B}' = (f(e_1), \dots, f(e_n))$ son image par f. Si f est unitaire alors

$$(f(e_i)|f(e_j)) = (e_i|e_j) = \delta_{ij},$$

c'est-à-dire que $\mathcal{B}'$ est une base orthonormée de E. Réciproquement, soit $(e_1, \dots, e_n)$ une base orthonormée de E, dont l'image $(e'_1, \dots, e'_n)$ par f est encore une base orthonormée de E. On pose

$$x = x_1e_1 + \cdots + x_ne_n \text{ et } y = y_1e_1 + \cdots + y_ne_n.$$

On a alors

$$\big(f(x)|f(y)\big) = \Big(\sum_{j=1}^{n} x_je'_j\Big|\sum_{k=1}^{n} y_ke'_k\Big) = \sum_{j=1}^{n}\sum_{k=1}^{n} \overline{x_j}y_k(e'_j, e_k) = \sum_{j=1}^{n} \overline{x_j}y_j = (x|y),$$

c'est-à-dire que f est unitaire. □

Proposition 21. *Si f appartient à $U(E)$, alors les valeurs propres de f sont de module* 1.

Démonstration. Soit λ une valeur propre de $f \in U(E)$. Il existe donc $x \neq 0$ appartenant à E tel que $f(x) = \lambda x$. Ainsi $||x|| = ||f(x)|| = ||\lambda x|| = |\lambda|\,||x||$, ce qui entraîne que $|\lambda| = 1$ car $||x|| \neq 0$. □

II.2. Matrices unitaires.

Définition 22. *Soit M une matrice de $\mathcal{M}_n(\mathbb{C})$. On appelle* matrice adjointe *de M la matrice*

$$M^* = {}^t\big(\overline{M}\big) = \overline{{}^tM},$$

dont les termes sont les conjugués de ceux de la transposée de M, autrement dit $m^*_{jk} = \overline{m_{kj}}$ pour $1 \le j \le n$ et $1 \le k \le n$.

Notons que l'on a les propriétés suivantes.

Proposition 23. *Soient M_1 et M_2 deux matrices de $\mathcal{M}_n(\mathbb{C})$ et λ un nombre complexe. On a :*
a) $(M_1^*)^* = M_1$;
b) $(M_1 + M_2)^* = M_1^* + M_2^*$;
c) $(M_1M_2)^* = M_2^*\,M_1^*$;
d) $(\lambda M_1)^* = \overline{\lambda}M_1^*$.

Ces résultats se déduisent directement de la définition et sont donc laissés au lecteurs.

Définition 24. *Soit M une matrice de $\mathcal{M}_n(\mathbb{C})$. On dit que M est* unitaire *si ${}^t\big(\overline{M}\big)M = \mathrm{I}_n$, c'est-à-dire si $M^*M = \mathrm{I}_n$. L'ensemble des matrices unitaires de $\mathcal{M}_n(\mathbb{C})$ s'appelle le* groupe unitaire *de $\mathcal{M}_n(\mathbb{C})$ et est noté $U(n)$.*

Remarque 25. *On a $M^* = M^{-1}$ et donc $MM^* = MM^{-1} = \mathrm{I}_n$. De plus comme $(M^*)^* = M$, si M est unitaire alors $M^* = M^{-1}$ est unitaire. Enfin, $U(n)$ est un groupe pour la multiplication des matrices.*

Exemple 26. *La matrice*

$$M = \begin{pmatrix} 0 & 0 & i \\ 1 & 0 & 0 \\ 0 & -i & 0 \end{pmatrix}$$

est unitaire car

$$M^* = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & i \\ -i & 0 & 0 \end{pmatrix} \text{ et } MM^* = \mathrm{I}_3.$$

Proposition 27. *Soit M une matrice de $\mathcal{M}_n(\mathbb{C})$. Les assertions suivantes sont équivalentes :*
a) M est unitaire ;
b) Les vecteurs lignes de M forment une base orthonormée de $\mathbb{C}^n$;
c) Les vecteurs colonnes de M forment une base orthonormée de $\mathbb{C}^n$.

Démonstration. Soit M une matrice de $\mathcal{M}_n(\mathbb{C})$. On note $e_1, \ldots, e_n$ ses n vecteurs lignes. Alors les vecteurs colonnes de M^* sont exactement $\overline{e_1}, \ldots, \overline{e_n}$. Soit $A = MM^*$. Les coefficients de la matrice A sont égaux à $a_{jk} = (e_k, e_j)$ pour $1 \leq j \leq n$ et $1 \leq k \leq n$. On obtient

$$\begin{aligned} M \text{ est unitaire} \quad &\Longleftrightarrow \quad A = MM^* = \mathrm{Id}_n \\ &\Longleftrightarrow \quad \forall 1 \leq j \leq n,\ \forall 1 \leq k \leq n, \quad (e_k, e_j) = \delta_{kj} \end{aligned}$$

ce qui est équivalent à $(e_1, \ldots, e_n)$ est une base orthonormée de $\mathbb{C}^n$.
On obtient la deuxième équivalence par la même méthode en utilisant le fait que $M^*M = \mathrm{Id}_n$. □

Proposition 28. *Soient E un espace hermitien et $\mathcal{B}$ une base orthonormée de E. On note f un endomorphisme de E et M sa matrice dans la base $\mathcal{B}$. On a alors : f est un automorphisme unitaire si et seulement si M est une matrice unitaire.*

Remarque 29. *La base orthonormée $\mathcal{B}$ étant quelconque, le résultat reste vrai dans toutes les bases orthonormées, c'est-à-dire que si f est unitaire sa matrice dans toute base orthonormée est unitaire.*
Ce résultat permet d'identifier les groupes $U(E)$ et $U(n)$ (ils sont isomorphes). Cela signifie que la structure de $U(E)$ ne dépend ni de E, ni du produit scalaire choisi, mais uniquement de la dimension n de E.

Démonstration. Soit M la matrice de f dans une base $\mathcal{B} = (e_1, \ldots, e_n)$ orthonormée de E. Si M est unitaire, alors ceci est équivalent à : les vecteurs colonnes $(f(e_1), \ldots, f(e_n))$ forment une base orthonormée de E ou encore à f est un automorphisme unitaire. □

Proposition 30. *Le déterminant d'une matrice unitaire a pour module* 1.

Attention la réciproque est fausse.

Démonstration. Soit M une matrice unitaire. On a donc $MM^* = \mathrm{I}_n$. On en déduit que $\det(MM^*) = \det(M)\det(M^*) = \det(\mathrm{I}_n) = 1$. De plus comme $\det(M^*) = \overline{\det(M)}$ on obtient $|\det(M)|^2 = 1$ c'est-à-dire $|\det(M)| = 1$. □

Définition 31. *On appelle* groupe spécial unitaire *de $U(n)$ le sous-groupe de $U(n)$ constitué des matrices de déterminant 1, et on le note $SU(n)$.*
Si E est un espace hermitien, $SU(E)$, le groupe spécial unitaire *de $U(E)$, est l'ensemble des automorphismes unitaires de déterminant 1.*

Il est intéressant de remarquer que le groupe orthogonal est un sous-groupe du groupe unitaire et donc, que le groupe spécial orthogonal est un sous-groupe du groupe spécial unitaire. En effet, une matrice orthogonale est une matrice réelle telle que ${}^tMM = \mathrm{I}_n$ et donc $M^*M = \mathrm{I}_n$ car pour une matrice réelle $M^* = {}^tM$.

II.3. Changement de bases orthonormées. .

Comme dans le cas d'un espace euclidien, il existe une formule plus pratique pour les espaces hermitiens. Soient f un endomorphisme d'un espace hermitien E et $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormées de E. On note M la matrice de f dans $\mathcal{B}$ et M' la matrice de f dans $\mathcal{B}'$. Rappelons que l'on a :

$$M' = P_{\mathcal{B}\to\mathcal{B}'}^{-1} M P_{\mathcal{B}\to\mathcal{B}'}.$$

Or la matrice $P_{\mathcal{B}\to\mathcal{B}'}$ est celle d'un endomorphisme de E qui transforme une base orthonormée en une autre. On en déduit que c'est une matrice unitaire. Elle vérifie donc $P_{\mathcal{B}\to\mathcal{B}'}^{*} = P_{\mathcal{B}\to\mathcal{B}'}^{-1}$. La formule de changement de base devient donc

$$M' = P_{\mathcal{B}\to\mathcal{B}'}^{*} M P_{\mathcal{B}\to\mathcal{B}'}.$$

III. Endomorphismes hermitiens ou auto-adjoints

Définition 32. *On dit qu'un endomorphisme f d'un espace hermitien E est* hermitien *ou* auto-adjoint *si* $\forall x \in E,\ \forall y \in E, \quad (f(x)|y) = (x|f(y))$.
On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{C})$ est hermitienne *si et seulement si $M^* = M$.*

Exemple 33. *On peut vérifier aisément que les matrices symétriques réelles sont hermitiennes. De plus si M est antisymétrique réelle alors iM est hermitienne.*

Proposition 34. *Si f est un endomorphisme hermitien de E et si F est un sous-espace vectoriel de E stable par f alors $F^\perp$ est stable par f.*

DÉMONSTRATION. Soit $x \in F^\perp$. Pour tout $y \in F$ on a $(f(x)|y) = (x|f(y))$. Or F est stable par f donc $f(y)$ appartient à F et on en déduit que $(f(x)|y) = (x|f(y)) = 0$, c'est-à-dire que $f(x) \in F^\perp$. □

Proposition 35. *Soit $f \in \mathcal{L}(E)$. f est hermitien si et seulement si sa matrice dans une base orthonormée de E est hermitienne.*

DÉMONSTRATION. Soit M la matrice de f dans une base orthonormée de E. Pour tout x et y dans E on a :

$$(f(x)|y) = (MX)^*Y = X^*M^*Y \text{ et } (x, f(y)) = X^*(MY) = X^*MY.$$

On en déduit que f est hermitien si et seulement si $M^* = M$, c'est-à-dire que M est une hermitienne. □

Proposition 36. *Si λ est une valeur propre d'un endomorphisme hermitien f (respectivement d'une matrice hermitienne M) alors $\lambda \in \mathbb{R}$.*

DÉMONSTRATION. Soit λ une valeur propre d'un endomorphisme hermitien f. Il existe donc un vecteur $x \neq 0$ de E tel que $f(x) = \lambda x$. On a

$$\begin{aligned}(f(x)|x) &= (\lambda x|x) = \overline{\lambda}||x||^2 \\ &= (x|\lambda x) = \lambda||x||^2,\end{aligned}$$

c'est-à-dire $\overline{\lambda} = \lambda$ et donc $\lambda \in \mathbb{R}$. □

Théorème 37.

> *Si f est un endomorphisme hermitien il existe une base orthonormée formée de vecteurs propres de f. Autrement dit f est diagonalisable dans une base orthonormée.*

Démonstration. On va procéder par récurrence sur la dimension de l'espace E. Si $\dim(E) = 1$ le résultat est trivial. On suppose que le résultat a été établi si $\dim(E) = n$. On considère donc f un endomorphisme hermitien sur un espace hermitien de dimension $n+1$. Soit λ une valeur propre de f (f est un endomorphisme sur un espace vectoriel complexe, il est donc scindé) et E_λ le sous-espace propre associé. E_λ est stable par f et donc $E_\lambda^\perp$ aussi. De plus on a et $\dim(E_\lambda^\perp) \leq n$. La restriction de f à $E_\lambda^\perp$ est donc un endomorphisme hermitien sur un espace de dimension p inférieure à n. D'après l'hypothèse de récurrence il existe une base orthonormée $(e_{n-p+1}, \ldots, e_n)$ de $E_\lambda^\perp$ qui diagonalise cette restriction. Alors $(e_1, \ldots, e_{n-p}, e_{n-p+1}, \ldots, e_n)$ est une base orthonormée de E qui diagonalise f. □

Corollaire 38. *Si M est une matrice hermitienne, il existe $P \in U(n)$ telle que $M' = P^{-1}MP = P^*MP$ soit diagonale. De plus M' est réelle.*

Remarque 39. *Si x et y sont deux vecteurs propres associées à des valeurs propres λ et μ différentes alors x et y sont orthogonaux. En effet, si λ est une valeur propre, alors $\lambda \in \mathbb{R}$ et donc $\lambda(x|y) = \overline{\lambda}(x|y) = (\lambda x|y)$. Ainsi $\lambda(x|y) = (f(x)|y) = (x|f(y)) = \mu(x|y)$. et donc $(\lambda - \mu)(x|y) = 0$ c'est-à-dire $(x|y) = 0$ car $\lambda \neq \mu$.*

IV. Adjoint d'un endomorphisme

Proposition 40. *Si f est un endomorphisme de E, il existe un unique endomorphisme g de E tel que*

$$\forall x \in E,\ \forall y \in E, \quad (f(x)|y) = (x|g(y)).$$

Cet endomorphisme est appelé adjoint *de f et est noté f^*.*

L'adjointe d'une matrice a été défini dans la Définition 19 par $M^* = {}^t\left(\overline{M}\right) = \overline{{}^tM}$.

Démonstration. Soit $(e_1, \ldots, e_n)$ une base orthonormée de E. On note M la matrice de f dans cette base et M' celle de g, si g existe. On a pour tout $x,\ y \in E$

$$\begin{aligned} (f(x)|y) &= (MX)^*Y = X^*M^*Y \\ &= (x|g(y)) = X^*(M'Y) = X^*M'Y, \end{aligned}$$

donc $M' = M^*$ et g est unique. Soit maintenant g l'endomorphisme de matrice M^* dans $(e_1, \ldots, e_n)$. On a pour tout $x,\ y \in E$

$$(x|g(y)) = X^*M^*Y = (MX)^*Y = (f(x)|y);$$

G est donc une solution. □

Corollaire 41. *Si M est la matrice d'un endomorphisme f dans une base orthonormée alors la matrice de f^* est M^*.*

Corollaire 42.

f est unitaire si et seulement si $f^* = f^{-1}$.
f est hermitien si et seulement si $f^* = f$.

Démonstration. f est unitaire $\iff M^*M = \mathrm{I}_n \iff M^* = M^{-1} \iff f^* = f^{-1}$.
f est hermitien $\iff M^* = M \iff f^* = f$. □

De la Proposition 20 sur les matrices adjointes on déduit facilement la proposition suivante.

Proposition 43. *Soient f et g deux endomorphismes de E et λ un nombre complexe.*

a) $(f^*)^* = f$;
b) $(f+g)^* = f^* + g^*$.
c) $(f \circ g)^* = g^* \circ f^*$;
d) $(\lambda f)^* = \overline{\lambda} f^*$.

V. Endomorphismes normaux

Définition 44.
On dit qu'un endomorphisme f de E est normal *s'il vérifie*

$$f \circ f^* = f^* \circ f.$$

On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{C})$ est normale *si et seulement si $M^*M = MM^*$.*

D'après la définition et les propriétés d'un adjoint d'un endomorphisme, on a le résultat suivant.

Proposition 45. *Un endomorphisme f de E est normal si et seulement si matrice dans une base orthonormée est normale.*

On déduit facilement de ce résultat qu'un endomorphisme unitaire est normal et qu'un endomorphisme hermitien aussi. Cette proposition signifie également que si une matrice dans une base orthonormée est normale sa matrice dans n'importe quelle base orthonormée est normale ; autrement dit :

Corollaire 46. *Si A est une matrice normale et P une matrice unitaire, alors*

$$B = P^{-1}AP = P^*AP$$

est une matrice normale.

DÉMONSTRATION. Comme P est unitaire, $P^* = P^{-1}$ et on a :

$$\begin{aligned} BB^* &= (P^{-1}AP)(P^{-1}AP)^* = P^{-1}APP^*A^*(P^{-1})^* = P^{-1}AA^*P \\ B^*B &= (P^{-1}AP)^*(P^{-1}AP) = P^*A^*(P^{-1})^*P^{-1}AP = P^{-1}A^*AP; \end{aligned}$$

donc $AA^* = A^*A$ entraîne $BB^* = B^*B$. □

On peut vérifier aisément qu'une matrice diagonale est normale.

Théorème 47.

Pour qu'un endomorphisme f soit normal, il faut et il suffit que l'on ait pour tous les vecteurs x et y de E :

$$(f(x)|f(y)) = (f^*(x)|f^*(y)).$$

DÉMONSTRATION. Par définition de l'adjoint de f, on a

$$(f(x)|f(y)) = (x|f^* \circ f(y)) \quad \text{et} \quad (f^*(x)|f^*(y)) = (x|f \circ f^*(y))$$

et donc le résultat. □

Théorème 48.

Pour qu'un endomorphisme f soit normal, il faut et il suffit qu'il y ait une base orthonormée dans laquelle f soit représenté par une matrice diagonale.
Pour les matrices cela est équivalent à : pour qu'une matrice M soit normale il faut et il suffit qu'il existe une matrice P unitaire telle que $P^{-1}MP = P^*MP$ soit diagonale.

Démonstration. Raisonnons par récurrence sur la dimension n. Pour $n = 1$ le résultat est trivial. On suppose que le résultat a été démontré jusqu'en dimension n, et on considère un endomorphisme normal f sur un espace de dimension $n + 1$. Soit λ une valeur propre de f. On considère $(e_1, \dots, e_p)$ une base orthonormée du sous-espace propre E_λ associé à λ. $E_\lambda^\perp$ est stable par f. En effet soit $y \in E_\lambda^\perp$. On a donc $(y|e_j) = 0$ pour tout $1 \leq j \leq p$. Par définition de l'adjoint on a encore $(f(y)|e_j) = (y|f^*(e_j))$. De plus comme f est normal on a :

$$f \circ f^*(e_j) = f^* \circ f(e_j) = \lambda f^*(e_j),$$

c'est-à-dire que $f^*(e_j) \in E_\lambda$ et donc $(y|f^*(e_j)) = 0$. Ainsi $(f(y)|e_j) = 0$ pour $1 \leq j \leq p$, autrement dit $f(y) \in E_\lambda^\perp$.
De la stabilité de $E_\lambda^\perp$ par f on déduit que la restriction g de f à $E_\lambda^\perp$ est encore normale, et comme cet espace est de dimension $n - p$, l'hypothèse de récurrence affirme qu'il existe une base $(e_{p+1}, \dots, e_n)$ qui diagonalise g. Donc la base $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ diagonalise f. □

De ce résultat on déduit que :

Théorème 49.

Pour qu'une matrice M soit unitaire il faut et il suffit qu'il existe une matrice P unitaire telle que $P^{-1}MP = P^*MP$ soit diagonale avec des valeurs propres de module 1.
Pour qu'une matrice M soit hermitienne il faut et il suffit qu'il existe une matrice P unitaire telle que $P^{-1}MP = P^*MP$ soit diagonale réelle.