Algèbre et géométrie I
Licence 3

Yann Chaubet

7 janvier 2026

Sommaire

Chapitre 1

Groupe symétrique et déterminant

Sommaire

Dans tout ce chapitre, n désigne un entier supérieur ou égal à 1.

1.1 Le groupe symétrique

1.1.1 Premières définitions

Définition 1.1.1 (Groupe symétrique). Le *groupe symétrique* ou *groupe des permutations* de $\{1, \dots, n\}$ est l'ensemble des bijections de $\{1, \dots, n\}$. Cet ensemble est noté $\mathfrak{S}_n$ et est un groupe pour la la loi de composition.

On utilisera parfois la notation

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

EXEMPLE 1.1.2. La permutation

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \tag{1.1}$$

est la permutation vérifiant $\tau(1) = 2$, $\tau(2) = 1$ et $\tau(3) = 3$.

DÉFINITION 1.1.3 (Support d'une permutation). Le *support* d'une permutation $\sigma \in \mathfrak{S}_n$ est l'ensemble $\operatorname{supp} \sigma \subset \{1, \ldots, n\}$ défini par

$$\operatorname{supp} \sigma = \big\{ i \in \{1, \ldots, n\} \ : \ \sigma(i) \neq i \big\}.$$

Notons qu'on a toujours $\sigma(\operatorname{supp} \sigma) = \operatorname{supp} \sigma$.

EXEMPLE 1.1.4. Le support de la permutation τ donnée par (1.1) est

$$\operatorname{supp} \tau = \{1, 2\}.$$

PROPOSITION 1.1.5. *Si $\sigma, \tau \in \mathfrak{S}_n$ sont deux permutations à supports disjoints, alors $\sigma\tau = \tau\sigma$, i.e. σ et τ commutent.*

Démonstration. Soit $i \in \{1, \ldots, n\}$. Si $i \notin (\operatorname{supp} \sigma \cup \operatorname{supp} \tau)$, alors $\sigma(i) = i$ et $\tau(i) = i$ de sorte que $(\sigma\tau)(i) = i = (\tau\sigma)(i)$. Si $i \in \operatorname{supp} \sigma$, alors $i \notin \operatorname{supp} \tau$ car σ et τ sont à supports disjoints. Puisque $\sigma(i) \in \operatorname{supp} \sigma$ on a aussi $\sigma(i) \notin \operatorname{supp} \tau$. Ainsi $\tau(i) = i$ et $\tau(\sigma(i)) = \sigma(i)$. On obtient donc

$$(\tau\sigma)(i) = \tau(\sigma(i)) = \sigma(i) = \sigma(\tau(i)) = (\sigma\tau)(i),$$

et on montre de même que si $i \in \operatorname{supp} \tau$ alors $(\tau\sigma)(i) = (\sigma\tau)(i)$. Ceci achève la démonstration. □

DÉFINITION 1.1.6 (Cycles). Soit $r \geqslant 2$. Une permutation $\sigma \in \mathfrak{S}_n$ est appelé *cycle de longueur r* s'il existe r entiers distincts $i_1, \ldots, i_r$ de $\{1, \ldots, n\}$ tels que

$$\sigma(i_1) = i_2, \quad \sigma(i_2) = i_3, \quad \ldots, \quad \sigma(i_r) = i_1,$$

et $\sigma(i) = i$ si i est distinct des i_m. Un tel cycle sera noté

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_r \end{pmatrix}. \tag{1.2}$$

Le support du cycle σ est $\{i_1, \ldots, i_r\}$. Un cycle de longueur 2 est appelé *transposition.*

Remarque 1.1.7. L'écriture (1.2) n'est pas unique : elle l'est à permutation circulaire des facteurs $i_1, \dots, i_k$ près. En effet pour tout $2 \leqslant q \leqslant k$, on a

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_r \end{pmatrix} = \begin{pmatrix} i_q & \cdots & i_k & i_1 & \cdots & i_{q-1} \end{pmatrix}.$$

Exemple 1.1.8. La permutation τ donnée par (1.1) est une transposition, et on a

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

Si $\sigma_1, \dots, \sigma_s$ sont des permutations de $\mathfrak{S}_n$, on notera

$$\sigma_1 \cdots \sigma_s = \sigma_1 \circ \cdots \circ \sigma_s.$$

Exemple 1.1.9. La permutation $\begin{pmatrix} i_1 & i_2 & \cdots & i_r \end{pmatrix} \begin{pmatrix} j_1 & i_2 & \cdots & i_p \end{pmatrix}$ est la composition $\sigma_1 \circ \sigma_2$ où $\sigma_1 = \begin{pmatrix} i_1 & i_2 & \cdots & i_r \end{pmatrix}$ et $\sigma_2 = \begin{pmatrix} j_1 & i_2 & \cdots & i_p \end{pmatrix}$.

Exercice 1.1.10. Montrer que pour tous $i_1, \dots, i_k$ distincts, on a

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix} = \begin{pmatrix} i_1 & i_k \end{pmatrix} \begin{pmatrix} i_1 & i_{k-1} \end{pmatrix} \cdots \begin{pmatrix} i_1 & i_2 \end{pmatrix}. \tag{1.3}$$

Pour tout $\sigma \in \mathfrak{S}_n$ et $\ell \in \mathbb{N}$, on définit $\sigma^\ell \in \mathfrak{S}_n$ par

$$\sigma^0 = \mathrm{id} \quad \text{et} \quad \sigma^{\ell+1} = \sigma\sigma^\ell = \sigma^\ell\sigma, \quad \ell \geqslant 0.$$

Autrement dit, on a

$$\sigma^\ell = \underbrace{\sigma \circ \cdots \circ \sigma}_{\ell \text{ fois}}.$$

Si $\ell \geqslant 0$, on note $\sigma^{-\ell} = (\sigma^{-1})^{-\ell}$, de sorte que $\sigma^{-\ell} = (\sigma^\ell)^{-1}$ et

$$\sigma^{\ell_1+\ell_2} = \sigma^{\ell_1}\sigma^{\ell_2}, \quad \ell_1, \ell_2 \in \mathbb{Z}. \tag{1.4}$$

1.1.2 Décomposition en produits de cycles et de transpositions

Le résultat suivant est crucial, et permet de décomposer toute permutation en des cycles à supports disjoints

Théorème 1.1.11 (Décomposition en produit de cycles à supports disjoints). *Toute permutation σ peut être décomposée comme un produit de cycles à support disjoints. Cette décomposition est unique, à l'ordre des cycles près.*

Avant de démontrer ce résultat, énonçons une conséquence importante.

Corollaire 1.1.12 (Décomposition en produit de transpositions). *Toute permutation σ peut être décomposée (de manière non unique !) comme un produit de transpositions.*

Démonstration du corollaire 1.1.12. D'après le théorème précédent il suffit de montrer que tout cycle $\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}$ s'exprime comme un produit de transpositions, ce qui découle de l'équation (1.3). Ceci conclut la démonstration. □

EXEMPLE 1.1.13. On suppose $n = 4$ et on regarde le cycle $\sigma = \begin{pmatrix} 2 & 1 & 4 \end{pmatrix}$, de sorte que

$$\sigma(1) = 4, \quad \sigma(2) = 1, \quad \sigma(3) = 3 \quad \text{et} \quad \sigma(4) = 2.$$

Alors on a $\sigma = \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \end{pmatrix}$.

Avant de démontrer le Théorème 1.1.11, introduisons quelques notions. Si $\sigma \in \mathfrak{S}_n$ et $i \in \{1, \ldots, n\}$, on note

$$\mathcal{O}_\sigma(i) = \left\{ \sigma^\ell(i) \ : \ \ell \in \mathbb{N} \right\}$$

l'*orbite* de i sous σ. Notons que $\sigma(\mathcal{O}_\sigma(i)) \subset \mathcal{O}_\sigma(i)$.

LEMME 1.1.14. *Pour tous $\sigma \in \mathfrak{S}_n$ et $i \in \{1, \ldots, n\}$, la suite $(\sigma^\ell(i))_{\ell \in \mathbb{N}}$ est périodique. On note*

$$\ell_\sigma(i) = \inf\{\ell \in \mathbb{N} \ : \ \sigma^\ell(i) = i\}$$

la période minimale de i sous σ.

Démonstration. Pour tout $\ell \in \mathbb{N}$ on a $\sigma^\ell(i) \in \{1, \ldots, n\}$. Ainsi, les $n+1$ nombres $i, \sigma(i), \ldots, \sigma^n(i)$ prennent au plus n valeurs distinctes, dont deux de ces nombres sont égaux, et il existe $\ell_1, \ell_2 \in \{0, \ldots, n\}$ tels que $\ell_1 < \ell_2$ et $\sigma^{\ell_1}(i) = \sigma^{\ell_2}(i)$. Mais alors en appliquant $\sigma^{-\ell_1}$ à cette égalité on obtient par (1.4)

$$i = \sigma^{\ell_0}(i)$$

où $\ell_0 = \ell_2 - \ell_1 > 0$. On obtient ainsi par (1.4)

$$\sigma^{\ell + \ell_0}(i) = \sigma^\ell(\sigma^{\ell_0}(i)) = \sigma^\ell(i)$$

donc la suite $(\sigma^\ell(i))$ est périodique de période ℓ_0. Par conséquence $\ell_\sigma(i)$ est bien défini. □

LEMME 1.1.15. *Si σ est un cycle et $i \in \operatorname{supp} \sigma$, on a*

$$\sigma = \begin{pmatrix} i & \sigma(i) & \cdots & \sigma^{\ell_\sigma(i)-1}(i) \end{pmatrix}.$$

Démonstration. On écrit $\sigma = \begin{pmatrix} i_1 & \cdots & i_\ell \end{pmatrix}$ où ℓ est la longueur de σ. Si $i \in \operatorname{supp} \sigma$, il existe q tel que $i_q = i$. Mais alors

$$\sigma = \begin{pmatrix} i & i_{q+1} & \cdots & i_\ell & i_1 & \cdots & i_{q-1} \end{pmatrix},$$

et on a bien $i_{q+m} = \sigma^m(i)$ si $0 \leqslant m \leqslant \ell - q$ et $\sigma^m(i) = i_{m-\ell+q}$ si $\ell - q + 1 \leqslant m < \ell$. Par ailleurs comme les i_m sont distincts et que $\sigma(i_{q-1}) = i_q = i$ par définition de σ, on a $\ell_\sigma(i) = \ell$, ce qui conclut la démonstration. □

Notons que le Lemme 1.1.14 implique que pour tous $i \in \{1, \ldots, n\}$ et tout $\sigma \in \mathfrak{S}_n$, on a

$$\mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j) \quad \text{pour tout } j \in \mathcal{O}_\sigma(i). \tag{1.5}$$

Ceci implique facilement que

$$j \in \mathcal{O}_\sigma(i) \quad \Longleftrightarrow \quad \mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j) \quad \Longleftrightarrow \quad i \in \mathcal{O}_\sigma(j). \tag{1.6}$$

Ainsi la relation $\sim$ définie par

$$i \sim j \quad \Longleftrightarrow \quad \mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j)$$

est une relation d'équivalence sur $\{1, \ldots, n\}$. D'autre part, notons que

$$i \notin \operatorname{supp} \sigma \quad \Longleftrightarrow \quad \mathcal{O}_\sigma(i) = \{i\}.$$

En utilisant (1.6) on en déduit alors

$$i \in \operatorname{supp} \sigma \quad \Longleftrightarrow \quad \mathcal{O}_\sigma(i) \subset \operatorname{supp} \sigma. \tag{1.7}$$

On peut maintenant procéder à la démonstration de la décomposition en cycles à supports disjoints.

Démonstration du Théorème 1.1.11. Soit $\sigma \in \mathfrak{S}_n$. Si $\operatorname{supp} \sigma = \{1, \ldots, n\}$ alors $\sigma = \mathrm{id}$ et il n'y a rien à démontrer. Sinon, on se donne $i_1 \in \operatorname{supp} \sigma$. Alors par (1.7) on a que $\mathcal{O}_1 = \mathcal{O}_\sigma(i_1)$ est contenu dans $\operatorname{supp} \sigma$, et on a

$$\mathcal{O}_1 = \left\{i_1, \sigma(i_1), \ldots, \sigma^{\ell_1 - 1}(i_1)\right\} \quad \text{où} \quad \ell_1 = \ell_\sigma(i_1) > 1.$$

Si $\mathcal{O}_1 = \operatorname{supp} \sigma$, on s'arrête. Sinon, on choisit $i_2 \in \operatorname{supp} \sigma \setminus \mathcal{O}_1$ et on pose $\mathcal{O}_2 = \mathcal{O}_\sigma(i_2)$ et $\ell_2 = \ell_\sigma(i_2)$. Alors on a $\mathcal{O}_2 \cap \mathcal{O}_1 = \emptyset$ puisque $i_2 \notin \mathcal{O}_1$. Par récurrence, on construit s éléments $i_1, \ldots, i_s$ deux à deux distincts, et des sous-ensembles

$$\mathcal{O}_k = \left\{i_k, \sigma(i_k), \ldots, \sigma^{\ell_k - 1}(i_k)\right\}, \quad \ell_k = \ell_\sigma(i_k), \quad k = 1, \ldots, s,$$

tels que $\operatorname{supp} \sigma = \mathcal{O}_1 \sqcup \cdots \sqcup \mathcal{O}_s$. On affirme alors que

$$\sigma = \sigma_1 \cdots \sigma_s \quad \text{où} \quad \sigma_k = \begin{pmatrix} i_k & \sigma(i_k) & \cdots & \sigma^{\ell_k - 1}(i_k) \end{pmatrix} \quad \text{pour tout} \quad k = 1, \ldots, s. \tag{1.8}$$

En effet, soit $i \in \{1, \ldots, n\}$. Si $i \notin \operatorname{supp} \sigma$, on a $i \notin \mathcal{O}_k$ pour tout k, donc $\sigma_k(i) = i$ pour tout k, de sorte que $\sigma_1 \cdots \sigma_s(i) = i = \sigma(i)$. Si $i \in \operatorname{supp} \sigma$, il existe un unique $k \in \{1, \ldots, s\}$ tel que $k \in \mathcal{O}_k = \operatorname{supp} \sigma_k$. Mais alors pour $m \neq k$ on a $i \notin \operatorname{supp} \sigma_m = \mathcal{O}_m$, ce qui donne $\sigma_m(i) = i$ pour $m \neq k$. Puisque les σ_m commutent par la Proposition 1.1.5, on obtient

$$\sigma_1 \cdots \sigma_s(i) = \sigma_k(i).$$

Or $i = \sigma^\ell(i_k)$ pour un certain $\ell \in \{0, \dots, \ell_k - 1\}$ car $i \in \mathcal{O}_k$, donc

$$\sigma(i) = \sigma(\sigma^\ell(i_k)) = \sigma^{\ell+1}(i_k) = \sigma_k(i_k)$$

(si $\ell < \ell_k - 1$, la dernière inégalité est claire ; si $\ell = \ell_k - 1$ on utilise que $\sigma_k(\sigma^{\ell_k - 1}(i_k)) = i_k = \sigma^{\ell_k}(i_k) = i_k$ par définition de ℓ_k).

Il reste à démontrer l'unicité de la décomposition. Soit $\sigma = \tau_1 \cdots \tau_q$ une autre décomposition en cycles à supports disjoints. Soit $k \in \{1, \dots, s\}$ et $i \in \operatorname{supp} \sigma_k$. Alors il existe un unique m tels que $i \in \operatorname{supp} \tau_m$ puisque $i \in \operatorname{supp} \sigma$. On a alors

$$\tau_m^\ell(i) = \sigma^\ell(i) = \sigma_k^\ell(i)$$

pour tout $\ell \in \mathbb{N}$. On obtient que $\ell_{\sigma_k}(i) = \ell_\sigma(i) = \ell_{\tau_m}(i)$, et comme σ_k et τ_m sont des cycles, on a par le Lemme 1.1.15,

$$\begin{aligned} \sigma_k &= \begin{pmatrix} i & \sigma_k(i) & \cdots & \sigma_k^{\ell_{\sigma_k}(i)-1}(i) \end{pmatrix} \\ &= \begin{pmatrix} i & \sigma(i) & \cdots & \sigma^{\ell_\sigma(i)-1}(i) \end{pmatrix} \\ &= \begin{pmatrix} i & \tau_m(i) & \cdots & \sigma^{\ell_{\tau_m}(i)-1}(i) \end{pmatrix} = \tau_m. \end{aligned}$$

On a montré qu'il existe une application $\bar{m} : \{1, \dots, k\} \to \{1, \dots, q\}$ telle que

$$\sigma_k = \tau_{\bar{m}(k)}, \quad k = 1, \dots, s.$$

En inversant les rôles des σ_k et des τ_m, on construit de la même manière une fonction $\bar{k} : \{1, \dots, q\} \to \{1, \dots, s\}$ telle que $\tau_m = \sigma_{\bar{k}(m)}$ pour tout m. Par construction on a la relation $\bar{k} \circ \bar{m} = \mathrm{id}$. Ceci implique $q = s$ et aussi que $(\tau_1, \dots, \tau_s)$ est un ré-ordonnement des $(\sigma_1, \dots, \sigma_s)$. Ceci conclut la démonstration. □

1.1.3 Signature

Dans toute la suite on note

$$\mathrm{P}_n = \{\{i, j\} \ : \ i \neq j\}$$

l'ensemble des paires (non ordonnées !) de $\{1, \dots, n\}$. On a en particulier

$$\operatorname{card} \mathrm{P}_n = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Toute permutation $\sigma \in \mathfrak{S}_n$ induit une application

$$\hat{\sigma} : \mathrm{P}_n \to \mathrm{P}_n, \quad \{i, j\} \mapsto \{\sigma(i), \sigma(j)\}.$$

On dit qu'une paire $\{i, j\} \in \mathrm{P}_n$ est *inversée* par σ si l'ordre de $\sigma(i), \sigma(j)$ est inversé par rapport à celui de i, j. Autrement dit, la paire $\{i, j\} \in \mathrm{P}_n$ est inversée si

$$D_\sigma(\{i, j\}) < 0, \quad \text{où} \quad D_\sigma(\{i, j\}) = \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Une paire est *non inversée* si elle n'est pas inversée, c'est-à-dire si $D_\sigma(\{i, j\}) > 0$.

Définition 1.1.16 (Nombre d'inversions et signature d'une permutation). Le *nombre d'inversions* $N(\sigma)$ de σ est, comme son nom l'indique, le nombre de paires non ordonnées $\{i, j\} \in \mathrm{P}_n$ qui sont inversées par σ. La *signature* $\varepsilon(\sigma)$ d'une permutation σ est donnée par

$$\varepsilon(\sigma) = (-1)^{N(\sigma)}.$$

On a la formule suivante pour $\varepsilon(\sigma)$.

Proposition 1.1.17. *Pour $\sigma \in \mathfrak{S}_n$ on a*

$$\varepsilon(\sigma) = \prod_{\{i,j\} \in \mathrm{P}_n} D_\sigma(\{i, j\}) = \prod_{\{i,j\} \in \mathrm{P}_n} \frac{\sigma(j) - \sigma(i)}{j - i}. \tag{1.9}$$

Démonstration. On définit $\tilde{\varepsilon}(\sigma)$ par le terme de droite de (1.9). Par définition de $N(\sigma)$, on voit tout de suite que le signe de $\tilde{\varepsilon}(\sigma)$ est le même que celui de $(-1)^{N(\sigma)}$. En outre, le changement de variables $\{k, \ell\} = \hat{\sigma}(\{i, j\}) = \{\sigma(i), \sigma(j)\}$ donne

$$|\tilde{\varepsilon}(\sigma)| = \frac{\prod_{\{i,j\} \in \mathrm{P}_n} |\sigma(j) - \sigma(i)|}{\prod_{\{i,j\} \in \mathrm{P}_n} |j - i|} = \frac{\prod_{\{k,\ell\} \in \mathrm{P}_n} |k - \ell|}{\prod_{\{i,j\} \in \mathrm{P}_n} |j - i|} = 1.$$

Mais comme $\tilde{\varepsilon}(\sigma)$ a même signe que $(-1)^{N(\sigma)}$ on obtient $\tilde{\varepsilon}(\sigma) = (-1)^{N(\sigma)} = \varepsilon(\sigma)$, ce qui conclut la démonstration. □

Théorème 1.1.18. *Pour toutes permutations $\sigma, \tau \in \mathfrak{S}_n$, on a*

$$\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau).$$

Autrement dit, la signature est un morphisme de groupes $\mathfrak{S}_n \to \{-1, 1\}$.

Démonstration. Soient $\sigma, \tau \in \mathfrak{S}_n$. Notons que pour $\{i, j\} \in \mathrm{P}_n$, on a

$$\begin{aligned} D_{\sigma\tau}(\{i, j\}) &= \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \frac{\tau(j) - \tau(i)}{j - i} \\ &= D_\sigma(\{\tau(i), \tau(j)\}) D_\tau(\{i, j\}). \end{aligned}$$

On en déduit que

$$\varepsilon(\sigma\tau) = \prod_{\{i,j\}\in \mathrm{P}_n} D_{\sigma\tau}(\{i,j\}) = \left(\prod_{\{i,j\}\in \mathrm{P}_n} D_\sigma(\hat{\tau}\{i,j\})\right)\left(\prod_{\{i,j\}\in \mathrm{P}_n} D_\tau(\{i,j\})\right).$$

Le dernier produit dans le membre de droite de l'égalité ci-dessus est $\varepsilon(\tau)$. D'autre part en effectuant le changement de variable $\{k,\ell\} = \hat{\tau}\{i,j\}$, on obtient

$$\prod_{\{i,j\}\in \mathrm{P}_n} D_\sigma(\hat{\tau}\{i,j\}) = \prod_{\{k,\ell\}\in \mathrm{P}_n} D_\sigma(\{k,\ell\}) = \varepsilon(\sigma).$$

Finalement on a bien obtenu $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. □

Théorème 1.1.19 (Propriétés de la signature). *La signature* $\varepsilon : \mathfrak{S}_n \to \{-1,1\}$ *vérifie les propriétés suivantes :*

(i) *si* $\sigma = \mathrm{id}$ *est la permutation identité, alors* $\varepsilon(\sigma) = 1$ *;*

(ii) *si* σ^{-1} *est l'inverse de* σ*, on a* $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1} = \varepsilon(\sigma)$ *;*

(iii) *si* σ *est un cycle de longueur* k*, alors* $\varepsilon(\sigma) = (-1)^{k-1}$ *;*

(iv) *si* τ *est une transposition, alors* $\varepsilon(\tau) = -1$.

Démonstration. Les points (i) est clair par le Théorème 1.1.18, puisque $\sigma(\mathrm{id}) = \sigma(\mathrm{id}\circ\mathrm{id}) = \sigma(\mathrm{id})^2 = 1$ puisque $\varepsilon(\mathrm{id}) \in \{-1,1\}$. Pour le point (ii), on remarque simplement que

$$1 = \varepsilon(\mathrm{id}) = \varepsilon(\sigma\circ\sigma^{-1}) = \varepsilon(\sigma)\varepsilon(\sigma^{-1})$$

ce qui donne $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1}$. Puisqu'un cycle de longueur k peut s'écrire comme un produit de $k-1$ transpositions (cf. (1.3)), le point (iii) se déduit du point (iv) facilement en utilisant le Théorème 1.1.18.

Montrons à présent le point (iv). Soit $\tau = \begin{pmatrix} i & j \end{pmatrix}$ une transposition. On peut supposer sans restriction de généralité que $i < j$, puisque $\tau = \begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} j & i \end{pmatrix}$. Une paire $\{k,\ell\} \in \mathrm{P}_n$ telle que $k,\ell \notin \operatorname{supp}\tau$ est stabilisée par τ (en effet $\tau(k) = k$ et $\tau(\ell) = \ell$) de sorte qu'elle n'est pas inversée par τ. Si $\{k,\ell\} = \{i,j\}$, alors $\{k,\ell\}$ est inversée par τ, puisque

$$i < j \quad \text{mais} \quad \tau(i) = j > i = \tau(j).$$

Il reste maintenant deux cas possibles : $k = i$ et $\ell \neq j$, ou $k = j$ et $\ell \neq i$. Si $k = i$ et $\ell \neq j$, on a que $\tau(\ell) = \ell$. En particulier, si $\ell < i$, alors la paire $\{k,\ell\}$ n'est pas inversée par τ. Si $i < \ell < j$, elle est inversée par τ puisqu'alors $\tau(\ell) = \ell < j = \tau(i)$. Enfin si $\ell > j$, la paire n'est pas inversée. Finalement, il y a exactement $j - i - 1$ paires $\{k,\ell\}$ qui sont inversées et telles que $k = i$ et $\ell \neq j$. De même, on montre qu'il y a exactement $j-i-1$ paires $\{k,\ell\}$ qui sont inversées et telles que $k=j$ et $\ell \neq i$. Ainsi on a obtenu que le nombre d'inversions

$$N(\tau) = 1 + 2(j-i-1)$$

est impair, donc $\varepsilon(\tau) = -1$. □

Une conséquence immédiate des propriétés précédentes est le résultat suivant, qui peut être utile en pratique.

COROLLAIRE 1.1.20. *Si $\sigma = \sigma_1 \cdots \sigma_s$, où les σ_m sont des cycles à supports disjoints, alors*

$$\varepsilon(\sigma) = (-1)^{\ell_1+\cdots+\ell_s-s}$$

où ℓ_m est la longueur du cycle σ_m.

Si $\sigma = \tau_1 \cdots \tau_s$ est un produit de s transpositions, alors $\varepsilon(\sigma) = (-1)^s$.

La dernière inégalité nous dit que, même si la décomposition en produit de transpositions n'est pas unique, la parité du nombre de transpositions d'une telle décomposition doit être toujours la même.

1.2 Déterminant

1.2.1 Formes n-linéaires alternées

Dans toute la suite, on note $E = \mathrm{K}^n$.

DÉFINITION 1.2.1. Une application $\mu : E^n \to \mathrm{K}$ est une *forme n-linéaire alternée* sur E si elle vérifie les propriétés suivantes :

(i) μ est linéaire en chacune de ses variables ;

(ii) pour tous $v_1, \ldots, v_n \in \mathrm{K}^n$, et tous $1 \leqslant i < j \leqslant n$,

$$\mu(v_1, \ldots, v_i, \ldots, v_j, \ldots v_n) = -\mu(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n).$$

On notera $\wedge^n E$ l'ensemble des formes n-linéaires alternées sur K^n. Étant donné $\mathbf{v} = (v_1, \ldots, v_n) \in E^n$ et $\sigma \in \mathfrak{S}_n$, on note

$$\sigma \cdot \mathbf{v} = (v_{\sigma(1)}, \ldots, v_{\sigma(n)}).$$

LEMME 1.2.2. *Pour des permutations $\tau, \rho \in \mathfrak{S}_n$, on a*

$$(\tau\rho) \cdot \mathbf{v} = \rho \cdot (\tau \cdot \mathbf{v}), \quad \mathbf{v} \in E^n. \tag{1.10}$$

Démonstration. En effet, si $(w_1, \ldots, w_n) = \tau \cdot (v_1, \ldots, v_n)$, on a pour tout i, $w_i = v_{\tau(i)}$ d'où $w_{\tau^{-1}(i)} = v_i$, de sorte que

$$w_{\rho(i)} = w_{\tau^{-1}(\tau(\rho(i)))} = v_{(\tau\rho)(i)}.$$

Ainsi

$$\rho \cdot (\tau \cdot \mathbf{v}) = \rho \cdot (w_1, \ldots, w_n) = (w_{\rho(1)}, \ldots, w_{\rho(n)}) = (v_{(\tau\rho)(1)}, \ldots, v_{(\tau\rho)(n)}) = (\tau\rho) \cdot \mathbf{v},$$

d'où l'on tire (1.10). $\square$

PROPOSITION 1.2.3. *Soit μ une forme n-linéaire alternée sur K^n. Alors pour toute permutation $\sigma \in \mathfrak{S}_n$, et tous $v_1, \ldots, v_n \in \mathrm{K}^n$, on a*

$$\mu(v_{\sigma(1)}, \ldots v_{\sigma(n)}) = \varepsilon(\sigma)\mu(v_1, \ldots, v_n).$$

Démonstration. Pour tout $\tau \in \mathfrak{S}_n$, on note $\tau \cdot (v_1, \ldots, v_n) = (v_{\tau(1)}, \ldots, v_{\tau(n)})$. Le point (i) de la Définition 1.2.1 implique que si $\tau = \begin{pmatrix} i & j \end{pmatrix}$ est une transposition avec $i < j$, on a

$$\begin{aligned} \mu(\tau \cdot (v_1, \ldots, v_n)) &= \mu(\tau \cdot (v_1, \ldots v_i, \ldots, v_j, \ldots v_n)) \\ &= \mu(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) \\ &= -\mu(v_1, \ldots, v_n). \end{aligned}$$

Décomposons à présent σ comme un produit de transpositions $\tau_1 \cdots \tau_s$. Par (1.10) on a, si $\mathbf{v} = (v_1, \ldots, v_n) \in (\mathrm{K}^n)^n$,

$$\begin{aligned} \mu(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) &= \mu(\sigma \cdot \mathbf{v}) \\ &= \mu\left((\tau_1 \cdots \tau_s) \cdot \mathbf{v}\right) \\ &= \mu\left(\tau_s \cdot (\tau_2 \cdots \tau_{s-1}) \cdot \mathbf{v}\right) \\ &= -\mu\left((\tau_2 \cdots \tau_{s-1}) \cdot \mathbf{v}\right). \end{aligned}$$

Par récurrence immédiate on obtient

$$\mu(\sigma \cdot \mathbf{v}) = (-1)^s \mu(\mathbf{v}) = \varepsilon(\sigma)\mu(\mathbf{v})$$

puisque $\varepsilon(\sigma) = (-1)^s$ par le corollaire 1.1.20. Ceci conclut la démonstration. $\square$

REMARQUE 1.2.4. Si $\mathbf{v} = (v_1, \ldots, v_n) \in (\mathrm{K}^n)^n$ est tel que $v_i = v_j$ avec $i \neq j$, alors $\mu(\mathbf{v}) = 0$. En effet, si $\tau = \begin{pmatrix} i & j \end{pmatrix}$ est la transposition qui intervertit i et j, on a $\tau \cdot \mathbf{v} = \mathbf{v}$, donc par la proposition précédente, on obtient

$$\mu(\mathbf{v}) = \mu(\tau \cdot \mathbf{v}) = -\mu(\mathbf{v})$$

donc $\mu(\mathbf{v}) = 0$.

1.2.2 Le déterminant comme forme n-linéaire alternée

Théorème–Définition 1.2.5. *Il existe une unique forme n-linéaire alternée, notée* $\det$ *et appelée déterminant, qui vaut* 1 *sur la base canonique de* K^n*. L'ensemble* $\wedge^n(\mathrm{K}^n)$ *est un* K *espace vectoriel de dimension* 1 *généré par* $\det$.

Démonstration. Soit $\mathbf{v} = (v_1, \dots, v_n) \in (\mathrm{K}^n)^n$ et $\mu \in A_n(\mathrm{K})$. On décompose chaque v_i dans la base canonique $\mathbf{e} = (e_j)_{1 \leqslant j \leqslant n}$ de K^n, en écrivant

$$v_i = \sum_{j=1}^{n} v_{i,j} e_j, \quad 1 \leqslant i \leqslant n.$$

Par multilinéarité, on obtient

$$\begin{aligned} \mu(\mathbf{v}) &= \mu\left(\sum_{j_1=1}^{n} v_{i,j_1} e_{j_1}, \dots, \sum_{j_n=1}^{n} v_{i,j_n} e_{j_n}\right) \\ &= \sum_{1 \leqslant j_1, \dots, j_n \leqslant n} v_{1,j_1} \cdots v_{n,j_n} \, \mu(e_{j_1}, \dots, e_{j_n}). \end{aligned}$$

S'il existe k, ℓ tels que $j_k = j_\ell$, alors $\mu(e_{j_1}, \dots, e_{j_n}) = 0$ par la remarque 1.2.4. Ainsi, dans la deuxième somme de l'égalité ci-dessus, seuls les termes pour lesquels $j_1, \dots, j_n$ sont deux à deux distincts peuvent être non nuls. En notant

$$Q_n = \{(j_1, \dots, j_n) \in \{1, \dots, n\}^n \ : \ j_k \neq j_\ell \text{ pour tous } k \neq \ell\},$$

on remarque qu'on a une bijection $\mathfrak{S}_n \to Q_n$ donnée par

$$\sigma \mapsto (\sigma(1), \dots \sigma(n)).$$

Ainsi on obtient

$$\begin{aligned} \mu(\mathbf{v}) &= \sum_{(j_1, \dots, j_n) \in Q_n} v_{1,j_1} \cdots v_{n,j_n} \, \mu(e_{j_1}, \dots, e_{j_n}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} v_{1,\sigma(1)} \cdots v_{n,\sigma(n)} \, \mu(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\ &= \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) v_{1,\sigma(1)} \cdots v_{n,\sigma(n)}\right) \mu(e_1, \dots, e_n), \end{aligned}$$

où dans la dernière inéaglité on a utilisé la Proposition 1.2.3, qui implique que $\mu(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma)\mu(e_1, \dots, e_n)$. Ainsi, la forme μ est uniquement déterminée par la valeur de $\mu(\mathbf{e})$. Définissons à présent

$$\det(\mathbf{v}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) v_{1,\sigma(1)} \cdots v_{n,\sigma(n)}.$$

Il est aisé de voir que det est une forme multilinéaire alternée. En effet, la linéarité par rapport à chacune des variables est claire. Si $\tau = \begin{pmatrix} k & \ell \end{pmatrix}$, notons que

$$v_{k,\sigma(\ell)} = v_{k,\sigma(\tau(k))} \quad \text{et} \quad v_{\ell,\sigma(k)} = v_{\ell,\sigma(\tau(\ell))}$$

et $v_{i,\sigma(i)} = v_{i,\sigma(\tau(i))}$ si $i \neq k, \ell$. On obtient

$$\begin{aligned} \det(v_1, \ldots, v_\ell, \ldots, v_k, \ldots, v_n) &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) v_{1,\sigma(1)} \cdots v_{\ell,\sigma(k)} \cdots v_{k,\sigma(\ell)} \cdots v_{n,\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) v_{1,\sigma(\tau(1))} \cdots v_{\ell,\sigma(\tau(\ell))} \cdots v_{k,\sigma(\tau(k))} \cdots v_{n,\sigma(\tau(n))}. \end{aligned}$$

En faisant le changement de variable $\tilde{\sigma} = \sigma\tau$, de sorte que $\varepsilon(\tilde{\sigma}) = \varepsilon(\sigma\tau) = -\varepsilon(\sigma)$, on obtient

$$\det(v_1, \ldots, v_\ell, \ldots, v_k, \ldots, v_n) = -\sum_{\tilde{\sigma} \in \mathfrak{S}_n} \varepsilon(\tilde{\sigma}) v_{1,\tilde{\sigma}(n)} \cdots v_{n,\tilde{\sigma}(n)} = -\det(\mathbf{v}).$$

Ainsi det est alternée, et par ce qui précède tout $\mu \in \wedge^n(\mathrm{K}^n)$ s'écrit

$$\mu = \mu(\mathbf{e}) \det .$$

Ceci achève la démonstration. $\square$

1.2.3 Déterminant d'une matrice

Si $A = (a_{i,j}) \in \mathrm{M}_n(\mathrm{K})$ est une matrice carrée de taille n, on pose

$$\det(A) = \det(v_1(A), \ldots, v_n(A))$$

où $v_j(A) \in \mathrm{K}^n$ contient les coefficients de la j^e colonne de A, de sorte que

$$v_j(A) = (a_{1,j}, \ldots, a_{n,j}), \ldots j = 1, \ldots, n.$$

En particulier, on a

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j}. \tag{1.11}$$

Proposition 1.2.6. *Si ${}^tA = (a_{j,i})_{1 \leqslant i,j \leqslant n}$ est la transposée de A, on a*

$$\det({}^tA) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)} = \det(A).$$

Démonstration. La première égalité est claire. La deuxième découle du changement de variable $k = \sigma(j)$ dans le produit de (1.11), qui donne

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} a_{k,\sigma^{-1}(j)}.$$

En faisant le changement de variable $\tilde{\sigma} = \sigma^{-1}$, on obtient $\varepsilon(\tilde{\sigma}) = \varepsilon(\sigma)$, d'où

$$\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} a_{k,\sigma^{-1}(j)} = \sum_{\tilde{\sigma} \in \mathfrak{S}_n} \varepsilon(\tilde{\sigma}) \prod_{k=1}^{n} a_{k,\tilde{\sigma}(j)}$$

ce qui achève la démonstration. □

PROPOSITION 1.2.7. *Soit $A \in \mathrm{M}_n(\mathrm{K})$ une matrice carrée. Le déterminant vérifie les propriétés suivantes.*

(i) *Si deux colonnes de A sont égales, alors* $\det A = 0$.

(ii) *Si deux colonnes de A sont interverties, son déterminant change de signe.*

(iii) *Si on multiplie par $\lambda \in \mathrm{K}$ une des colonnes de A, son déterminant est multiplié par λ.*

(iv) *Si une colonne est combinaison linéaire des autres colonnes, alors* $\det A = 0$.

REMARQUE 1.2.8. La proposition précédente reste vraie si on remplace le mot “colonne” par “ligne”, ce qui découle de la Proposition 1.2.6.

Démonstration. Ces propriétés découlent immédiatement du fait que det est une forme n-linéaire alternée. □

Soit $E = \mathrm{K}^n$. Si $u \in \mathcal{L}(E)$ est un endomorphisme de E, on note $u^{\oplus n}$ l'endomorphisme $u^{\oplus n} : E^n \to E^n$ défini par

$$u^{\oplus n}(v_1, \dots, v_n) = (u(v_1), \dots, u(v_n)), \quad v_1, \dots, v_n \in E.$$

THÉORÈME 1.2.9. *Si $A = (a_{i,j})_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathrm{K})$ est la matrice de u dans la base canonique de E, on a*

$$\det(u(v_1), \dots, u(v_n)) = \det(A) \det(v_1, \dots, v_n), \quad v_1, \dots, v_n \in E.$$

Démonstration. On définit $\mu : E^n \to E^n$ par $(v_1, \dots, v_n) \mapsto \det(u(v_1), \dots, u(v_n))$. Alors μ est une forme multilinéaire alternée, ce qui découle de la linéarité de u et du caractère multilinéaire alterné de det. Par le Théorème-Définition 1.2.5, il existe un scalaire $\lambda \in \mathrm{K}$ tel que

$$\mu = \lambda \det : E^n \to \mathrm{K}.$$

En appliquant ceci à $(v_1, \ldots, v_n) = (e_1, \ldots, e_n)$, on obtient

$$\mu(e_1, \ldots, e_n) = \det(u(e_1), \ldots, u(e_n)) = \lambda.$$

Cependant on a, par définition de A,

$$u(e_j) = (a_{1,j}, \ldots, a_{n,j}), \quad j = 1, \ldots, n.$$

Ceci implique immédiatement que $\lambda = \det(u(e_1), \ldots, u(e_n)) = \det A$. On a donc obtenu $\mu = \det(A)\det$, ce qui conclut la démonstration. □

Théorème 1.2.10. *Soient A et B deux matrices de* $\mathrm{M}_n(\mathrm{K})$. *Alors*

$$\det(AB) = \det(A)\det(B).$$

Démonstration. Soient u, v les endomorphismes de E canoniquement associés à A et B, respectivement. Alors $u \circ v$ est canoniquement associé à $u \circ v$. Soient $v_1, \ldots, v_n \in E$. On applique le Théorème 1.2.9 à l'endomorphisme $u \circ v$, ce qui donne

$$\det((u \circ v)(v_1), \ldots, (u \circ v)(v_n)) = \det(AB)\det(v_1, \ldots, v_n).$$

On applique maintenant le Théorème 1.2.9 à l'endomorphisme u, ce qui donne, si $w_i = v(v_i)$ pour $i = 1, \ldots, n$,

$$\det(u(w_1), \ldots, u(w_n)) = \det(A)\det(w_1, \ldots, w_n).$$

Enfin le Théorème 1.2.9 appliqué à l'endomorphisme v donne

$$\det(w_1, \ldots, w_n) = \det(v(v_1), \ldots, v(v_n)) = \det(B)\det(v_1, \ldots, v_n).$$

En prenant $v_i = e_i$, on obtient $\det(v_1, \ldots, v_n) = \det(e_1, \ldots, e_n)$, de sorte que les trois égalités précédentes donnent

$$\begin{aligned}
\det(AB) &= \det((u \circ v)(e_1), \ldots, (u \circ v)(e_n)) \\
&= \det(u(w_1), \ldots, u(w_n)) \\
&= \det(A)\det(w_1, \ldots, w_n) \\
&= \det(A)\det(B),
\end{aligned}$$

ce qu'on souhaitait démontrer. □

On déduit immédiatement du Théorème 1.2.10 le résultat suivant.

Corollaire 1.2.11 (Invariance du déterminant par similitude). *Soit $P \in \mathrm{GL}_n(\mathrm{K})$ une matrice inversible et $A \in \mathrm{M}_n(\mathrm{K})$. Alors*

$$\det(P^{-1}AP) = \det(A).$$

Démonstration. Il suffit d'appliquer le Théorème 1.2.10 avec A remplacée par AP et B remplacée par P^{-1}. □

Une deuxième conséquence porte sur l'inversibilité d'une matrice.

COROLLAIRE 1.2.12. *Une matrice $A \in \mathrm{M}_n(\mathrm{K})$ est inversible si et seulement si* $\det A \neq 0$.

Démonstration. Si A est inversible, alors

$$1 = \det(\mathrm{I}_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

donc $\det A \neq 0$. Si A n'est pas inversible, alors il existe une combinaison linéaire non triviale de ses colonnes qui est nulle. Par la Proposition 1.2.7, on obtient $\det A = 0$. Ceci conclut la démonstration. □

1.2.4 Développement par rapport aux lignes et aux colonnes

THÉORÈME 1.2.13 (Développement du déterminant par rapport à une ligne ou une colonne). *Soit $A = (a_{i,j}) \in \mathrm{M}_n(\mathrm{K})$ une matrice. Alors pour tout $i = 1, \ldots, n$ on a*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \Delta_{i,j} = \sum_{j=1}^{n} (-1)^{i+j} a_{j,i} \Delta_{j,i} \tag{1.12}$$

où $\Delta_{i,j}$ est le déterminant de la sous-matrice de A obtenue en retirant la i-eme ligne et la j-eme colonne, soit

$$\Delta_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{n,n} \end{vmatrix}. \tag{1.13}$$

Démonstration. On note $\ell_i(A) = (a_{i,1}, \ldots, a_{i,n}) \in \mathrm{K}^n$ la i-ième ligne de A. En écrivant $\ell_i(A) = \sum_{j=1}^{n} a_{i,j} e_j$ où (e_j) est la base canonique de K^n, on obtient via la linéarité de det par rapport à son i-ième facteur,

$$\det(A) = \det(\ell_1(A), \ldots, \ell_n(A)) = \sum_{j=1}^{n} a_{i,j} D_{i,j} \tag{1.14}$$

où pour tous i, j, avec $D_{i,j} = \det(M_{i,j})$, où

$$M_{i,j} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Pour tout $k \neq i$, on retire $a_{k,j}$ fois la i-ième ligne de $M_{i,j}$ à la k-ième ligne de $M_{i,j}$, ce qui ne modifie pas son déterminant, de sorte que

$$D_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

Après $i-1$ échanges de lignes et $j-1$ échanges de colonnes, on obtient

$$D_{i,j} = (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \vdots & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

Nous aurons besoin du résultat intermédiaire suivant.

Lemme 1.2.14. *Soient* $B = (b_{i,j})_{1 \leqslant i,j \leqslant n-1}$, $c_1, \ldots, c_{n-1} \in \mathrm{K}$ *et*

$$A = \begin{pmatrix} 1 & c_1 & \cdots & c_{n-1} \\ 0 & b_{1,1} & \cdots & b_{1,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n-1,1} & \cdots & b_{n-1,n-1} \end{pmatrix}.$$

Alors $\det A = \det B$.

Démonstration du Lemme 1.2.14. On calcule

$$\det A = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1}\cdots a_{\sigma(n),n}.$$

On a $a_{\sigma(1),1} = 0$ si $\sigma(1) \neq 1$. Si $\sigma(1) = 1$ alors $\sigma(j) \neq 1$ pour tout $j = 2, \ldots, n$, de sorte que

$$a_{\sigma(1),1}\cdots a_{\sigma(n),n} = b_{\sigma(2)-1,2-1}\cdots b_{\sigma(n)-1,n-1}$$

puisque $a_{\sigma(1),1} = a_{1,1} = 1$, et $a_{i,j} = b_{i-1,j-1}$ si $i, j > 1$. Dès lors

$$\det(A) = \sum_{\substack{\sigma\in\mathfrak{S}_n\\ \sigma(1)=1}} \varepsilon(\sigma) b_{\sigma(2)-1,2-1}\cdots b_{\sigma(n)-1,n-1}$$

où la somme porte sur les permutations σ vérifiant $\sigma(1) = 1$. Notons qu'on a une bijection

$$\Psi : \{\sigma \in \mathfrak{S}_n \ : \ \sigma(1) = 1\} \to \mathfrak{S}_{n-1}, \quad \sigma \mapsto \hat{\sigma},$$

où $\hat{\sigma} \in \mathfrak{S}_{n-1}$ est définie par

$$\hat{\sigma}(j) = \sigma(j+1) - 1, \quad j = 1, \ldots, n-1.$$

Alors on a $\varepsilon(\hat{\sigma}) = \varepsilon(\sigma)$, ce qui se voit aisément en décomposant σ en produit de cycles à supports disjoints. En faisant le changement de variables $\hat{\sigma} = \Psi(\sigma)$, on obtient donc

$$\det(A) = \sum_{\hat{\sigma}\in\mathfrak{S}_{n-1}} \varepsilon(\hat{\sigma}) \prod_{j=1}^{n-1} b_{\hat{\sigma}(j),j} = \det(B),$$

ce qui achève la démonstration du lemme. □

Le lemme donne alors

$$D_{i,j} = (-1)^{i+j-2} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{n,n} \end{vmatrix} = (-1)^{i+j}\Delta_{i,j},$$

et avec (1.14) on obtient bien la première égalité de (1.12). La deuxième égalité se déduit immédiatement de la première égalité appliquée à ${}^tA$. La démonstration du Théorème 1.2.13 est complète. □

Nous donnons une application cruciale du résultat précédent.

Corollaire 1.2.15. *Le déterminant d'une matrice triangulaire supérieure est égal au produit de ses éléments diagonaux.*

Démonstration. On montre le résultat par récurrence sur la dimension. Pour $n = 1$, le résultat est trivial. Soit maintenant $n \geqslant 2$ et

$$A = \begin{pmatrix} a_{1,1} & \cdots & \cdots & a_{1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix}$$

une matrice triangulaire supérieure. En développant le déterminant par rapport à la première colonne de A, on obtient

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \Delta_{i,1}.$$

Puisque $a_{i,1} = 0$ pour $i \geqslant 2$, il vient

$$\det(A) = a_{1,1}\Delta_{1,1} \quad \text{avec} \quad \Delta_{1,1} = \begin{vmatrix} a_{2,2} & \cdots & \cdots & a_{2,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{vmatrix}.$$

Par hypothèse de récurrence, on a $\Delta_{1,1} = a_{2,2} \cdots a_{n,n}$ d'où $\det(A) = a_{1,1} \cdots a_{n,n}$. La récurrence est établie et le corollaire est démontré. □

1.2.5 Déterminants par blocs

Théorème 1.2.16 (Déterminant d'une matrice triangulaire supérieure par blocs). *Soient $n \geqslant 1$ et $A_k \in \mathrm{M}_{n_k}(\mathrm{K})$, $k = 1, \ldots, r$ des matrices avec $n_1 + \cdots + n_r = n$. Soit $A \in \mathrm{M}_n(\mathrm{K})$ une matrice par blocs de la forme*

$$A = \begin{pmatrix} \boxed{A_1} & \star & \star & \star \\ 0 & \ddots & \star & \star \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \boxed{A_r} \end{pmatrix}$$

où les $\star$ représentent des matrices quelconques. Alors on a

$$\det A = \det(A_1) \cdots \det(A_r).$$

Démonstration. Par récurrence immédiate, il suffit de prouver le théorème pour $r = 2$, de sorte qu'on supposer que A est de la forme

$$\begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$$

avec $A_j \in \mathrm{M}_{n_k}(\mathrm{K})$ pour $k = 1, 2$ et $B \in \mathrm{M}_{n_1,n_2}(\mathrm{K})$. Soit $E = \mathrm{K}^{n_1}$. On considère l'application $\mu_1 : E^{n_1} \to \mathrm{K}$ donnée par

$$\mu_1(\mathbf{v}) = \det \begin{pmatrix} A_1(\mathbf{v}) & B \\ 0 & A_2 \end{pmatrix}, \quad \mathbf{v} = (v_1, \ldots, v_{n_1}) \in E^{n_1},$$

où $A_1(\mathbf{v})$ est la matrice dont la i-ième colonne a pour coefficients ceux de v_i, pour $i = 1, \ldots, n_1$. Alors par n-linéarité du déterminant, μ_1 est une forme n_1-linéaire alternée sur E. Par conséquent, le Théorème 1.2.5 implique qu'il existe une constante λ qui dépend *a priori* de B et de A_2, telle que

$$\mu_1(\mathbf{v}) = \lambda \det(\mathbf{v}), \quad \mathbf{v} \in E_1^{n_1}. \tag{1.15}$$

En évaluant sur la base canonique $\mathbf{e}$ de E, on obtient

$$\lambda = \mu_1(\mathbf{e}) = \det \begin{pmatrix} \mathrm{I}_{n_1} & B \\ 0 & A_2 \end{pmatrix}.$$

Une récurrence immédiate combinée au Lemme 1.2.14 donne que le déterminant de la matrice ci-dessus coincide avec $\det A_2$, de sorte que

$$\lambda = \det A_2.$$

On note maintenant $c_i(A) \in \mathrm{K}^{n_1}$ le vecteur dont les coefficients sont ceux de la i-ième colonne de A. En notant $\mathbf{c} = (c_1(A), \ldots, c_{n_1}(A))$ on a $A_1(\mathbf{c}) = A_1$ et donc par (1.15) on obtient

$$\mu_1(\mathbf{c}) = \det \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix} = \lambda \det(\mathbf{c}) = \det(A_2) \det(\mathbf{c}).$$

Or $\det(\mathbf{c}) = \det A_1$ par définition du déterminant de A_1, ce qui conclut la démonstration. $\square$

1.2.6 Formule de la comatrice

Définition 1.2.17 (Comatrice). Pour toute matrice $A \in \mathrm{M}_n(\mathrm{K})$, on note $\mathrm{com}(A)$ la matrice des co-facteurs de A, c'est-à-dire la matrice dont le coefficient en place (i, j) est $(-1)^{i+j}\Delta_{i,j}$, où $\Delta_{i,j}$ est défini dans (1.13).

Théorème 1.2.18 (Formule de la comatrice). *Pour toute matrice* $A = (a_{i,j}) \in \mathrm{M}_n(\mathrm{K})$,

$$A\,{}^t\operatorname{com}(A) = \det(A)\mathrm{I}_n.$$

En particulier si A est inversible alors

$$A^{-1} = \frac{{}^t\operatorname{com}(A)}{\det A}.$$

Démonstration. On pose $B = {}^t\operatorname{com}(A)$ et on note $b_{i,j}$ les coefficients de B, de sorte que $b_{i,j} = (-1)^{j+i}\Delta_{j,i}$. Alors le coefficient en place (i,j) de AB est donné par

$$(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} = \sum_{k=1}^{n}(-1)^{k+j}a_{i,k}\Delta_{j,k}. \tag{1.16}$$

En particulier, si $i = j$, le Théorème 1.2.13 donne

$$(AB)_{i,i} = \sum_{k=1}^{n}(-1)^{k+i}a_{i,k}\Delta_{i,k} = \det(A). \tag{1.17}$$

Fixons à présent $i \neq j$. On considère la matrice $\tilde{A} = (\tilde{a}_{k,\ell})$ obtenue en remplaçant la j-ième ligne de A par la i-ième ligne de A. Alors $\tilde{A}$ a deux lignes égales, donc $\det \tilde{A} = 0$. Mais en développement le déterminant par rapport à la j-ième ligne,

$$0 = \det \tilde{A} = \sum_{k=1}^{n}(-1)^{j+k}\tilde{a}_{j,k}\tilde{\Delta}_{j,k}$$

où $\tilde{\Delta}_{j,k}$ est le déterminant de la matrice réduite de $\tilde{A}$ en enlevant la j-ième ligne et la k-ième colonne. Puisque la ℓ-ième ligne de $\tilde{A}$ est la même que la ℓ-ième ligne de A pour tout $\ell \neq j$, on a $\tilde{\Delta}_{j,k} = \Delta_{j,k}$. D'autre part on a $\tilde{a}_{j,k} = a_{i,k}$ puisque la j-ième ligne de $\tilde{A}$ coincide avec la i-ième ligne de A. En combinant ceci avec (1.16) on obtient, si $i \neq j$,

$$(AB)_{i,j} = \sum_{k=1}^{n}(-1)^{j+k}a_{i,k}\Delta_{j,k} = \sum_{k=1}^{n}(-1)^{j+k}\tilde{a}_{j,k}\tilde{\Delta}_{j,k} = 0.$$

En se rappelant de (1.17), on obtient bien $A\,{}^t\operatorname{com}(A) = \det(A)\mathrm{I}_n$. $\square$

Chapitre 2

Réduction des endomorphismes

Sommaire

2.1 Rappels

Si $f \in \mathscr{L}(E)$ et $k \in \mathbb{N}$, on note $f^k = \mathrm{id}$ si $k = 0$ et $f^k = f \circ \cdots \circ f$ (k termes) si $k \geqslant 1$.

2.1.1 Valeurs propres et vecteurs propres

Définition 2.1.1. Soit f un endomorphisme de E.

(i) Un scalaire $\lambda \in K$ est une *valeur propre* de f s'il existe $x \in E \setminus \{0\}$ tel que $f(x) = \lambda x$.

(ii) Un tel vecteur x est appelé *vecteur propre* de f pour la valeur propre λ.

(iii) *L'espace propre* de f associé à la valeur propre λ est le sous-espace vectoriel

$$E_\lambda = \ker(f - \lambda\,\mathrm{id}) = \{x \in E \ : \ f(x) = \lambda x\}.$$

(iv) Le spectre $\mathrm{sp}(f)$ de f est l'ensemble de ses valeurs propres,

$$\mathrm{sp}(f) = \{\lambda \in K \ : \ \exists x \in E \setminus \{0\},\ f(x) = \lambda x\}.$$

(v) Si $\lambda \in \mathrm{sp}(f)$, le nombre $m_f^{\mathrm{geom}}(\lambda) = \dim E_\lambda$ est appelé sa *multiplicité algébrique.*

Un résultat fondamental est que les espaces propres associés à des valeurs propres distinctes sont en somme directe.

Lemme 2.1.2. *Soient $\lambda_1, \ldots, \lambda_r$ des valeurs propres de f deux à deux distinctes, et $x_1, \ldots, x_r$ des vecteurs propres associés. Alors la famille $(x_1, \ldots, x_r)$ est libre.*

Démonstration. On raisonne par récurrence sur r. Le lemme est évident vrai pour $r = 1$ car tout vecteur propre est non nul. Supposons le lemme vrai pour un $r \geqslant 1$, et donnons nous $x_1, \ldots, x_{r+1}$ une famille de vecteurs propres de f associés à des valeurs propres deux à deux distinctes $\lambda_1, \ldots, \lambda_{r+1}$. Montrons que la famille $(x_1, \ldots, x_{r+1})$ est libre ; soient $\alpha_1, \ldots, \alpha_{r+1} \in K$ tels que

$$\sum_{j=1}^{r+1} \alpha_j x_j = 0. \tag{2.1}$$

En appliquant f à l'égalité précédente, on obtient

$$\sum_{j=1}^{r+1} \lambda_j \alpha_j x_j = 0. \tag{2.2}$$

En multipliant (2.1) par λ_{r+1} et en retranchant (2.2), on obtient

$$\sum_{j=1}^{r}(\lambda_j - \lambda_{r+1})\alpha_j x_j = 0.$$

Par hypothèse de récurrence, la famille $(x_1, \ldots, x_r)$ est libre ; ainsi, puisqu'on a $\lambda_j - \lambda_{r+1} \neq 0$ pour tout $j = 1, \ldots, r$ on obtient $\alpha_1 = \cdots = \alpha_r = 0$ et (2.1) donne $\alpha_{r+1} = 0$. Il suit que $(x_1, \ldots, x_{r+1})$ est libre et la récurrence est établie. $\square$

Remarque 2.1.3. Le résultat précédent implique directement que le spectre de f est de cardinal au plus n.

Définition 2.1.4. Soit f un endomorphisme de E. On dit que f est *diagonalisable* (resp. *trigonalisable*) s'il existe une base de E constituée de vecteurs propres pour f, i.e. dans laquelle la matrice de f est diagonale (resp. triangulaire supérieure).

Ces notions existent aussi pour les matrices.

Définition 2.1.5. Soit $A \in \mathrm{M}_n(\mathrm{K})$ une matrice. On dit que A est *diagonalisable* (resp. *trigonalisable*) s'il existe $P \in \mathrm{GL}_n(\mathrm{K})$ telle que $P^{-1}AP$ est diagonale (resp. triangulaire supérieure).

Remarque 2.1.6. Un endomorphisme $f \in \mathscr{L}(E)$ est diagonalisable (resp. trigonalisable) si, et seulement si, sa matrice dans n'importe quelle base est diagonalisable (resp. trigonalisable).

Proposition 2.1.7 (Critère de diagonalisation, I). *Un endomorphisme $f \in \mathscr{L}(E)$ est diagonalisable si, et seulement si, on a la somme directe*

$$E = \bigoplus_{\lambda \in \mathrm{sp}(f)} E_\lambda$$

Démonstration. Supposons f diagonalisable. Soit $\beta = (e_1, \ldots, e_n)$ une base de E formée de vecteurs propres de f. Notons $\mathrm{sp}(f) = \{\lambda_1, \ldots, \lambda_r\}$, et pour tout $j = 1, \ldots, r$ on pose

$$n_j = \sharp\{1 \leqslant \ell \leqslant n \ : \ e_\ell \in E_{\lambda_j}\}.$$

Alors on a $n_j \leqslant \dim E_{\lambda_j}$ pour tout j et

$$n = n_1 + \cdots + n_r \leqslant \sum_{j=1}^{r} \dim E_{\lambda_j} = \dim\left(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}\right),$$

où on a utilisé le Lemme 2.1.2 pour la dernière égalité. Il suit que

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

Réciproquement, supposons qu'on a la somme directe précédente, et prenons β une base adaptée à la décomposition. Alors la matrice de f est diagonale dans cette base puisque tout vecteur de β appartient à un des espaces propres E_{λ_j}. $\square$

2.1.2 Polynôme caractéristique

Définition 2.1.8. Le polynôme caractéristique d'une matrice carrée $A \in \mathrm{M}_n(K)$ est l'application $\chi_A : K \to K$ définie par

$$\chi_A(\lambda) = \det(\lambda \mathrm{I}_n - A), \quad \lambda \in K.$$

Proposition 2.1.9. *Le polynôme caractéristique d'une matrice est une application polynomiale. Plus précisément, pour toute matrice* $A \in \mathrm{M}_n(\mathrm{K})$*, il existe* $a_0, \dots, a_n \in \mathrm{K}$ *tels que*

$$\chi_A(\lambda) = \sum_{k=0}^{n} a_k \lambda^k, \quad \lambda \in \mathrm{K}.$$

De plus, on a les expressions

$$a_0 = (-1)^n \det A, \quad a_{n-1} = -\operatorname{tr} A \quad \textit{et} \quad a_n = 1.$$

Démonstration. Pour $A = (a_{i,j}) \in \mathrm{M}_n(\mathrm{K})$ on utilise la formule du déterminant

$$\det(\lambda \mathrm{I}_n - A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} \left(\lambda \delta_{j,\sigma(j)} - a_{j,\sigma(j)}\right).$$

Ici $\mathfrak{S}_n$ est l'ensemble des permutations de $\{1, \dots, n\}$ et $\delta_{i,j}$ est le symbole de Kronecker donné par $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ sinon. L'expression de droite est manifestement polynomiale en la variable λ. En écrivant $\chi_A(\lambda) = a_0 + \cdots + a_n \lambda^n$, on obtient

$$a_0 = \det(-A) = (-1)^n \det A.$$

D'autre part, si $\sigma \in \mathfrak{S}_n$ est une permutation différente de la permutation triviale, le produit

$$P_\sigma(\lambda) = \prod_{j=1}^{n} \left(\lambda \delta_{j,\sigma(j)} - a_{j,\sigma(j)}\right)$$

est un polynôme en λ de degré au plus $n - 2$ puisque toute permutation non triviale contient un cycle de longueur au moins 2. Si $\sigma = e$ est la permutation triviale, un calcul immédiat donne

$$P_e(\lambda) = \prod_{j=1}^{n} (\lambda - a_{j,j}) = \lambda^n - \lambda^{n-1} \sum_{j=1}^{n} a_{j,j} + Q_e(\lambda)$$

où Q_e est un polynôme de degré au plus $n - 2$. Par conséquent on obtient que

$$\chi_A(\lambda) = \lambda^n - \lambda^{n-1} \operatorname{tr} A + Q_e(\lambda) + \sum_{\sigma \in \mathfrak{S}_n \setminus \{e\}} \varepsilon(\sigma) P_\sigma(\lambda).$$

Par ce qui précède, $Q_e(\lambda) + \sum_{\sigma \in \mathfrak{S}_n \setminus \{e\}} \varepsilon(\sigma) P_\sigma(\lambda)$ est un polynôme en λ de degré au plus $n-2$ et on en déduit immédiatement les expressions annoncées pour les coefficients a_n et a_{n-1}. □

Corollaire 2.1.10. *Si* $A \in \mathrm{M}_2(\mathrm{K})$ *on a*

$$\chi_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A).$$

Le polynôme caractéristique est un invariant de similitude, comme le montre la proposition suivante.

Proposition 2.1.11. *Si* $A \in \mathrm{M}_n(\mathrm{K})$ *et* $P \in \mathrm{GL}_n(\mathrm{K})$ *on a* $\chi_{P^{-1}AP} = \chi_A$.

Démonstration. En effet, le déterminant est un invariant de similitude, d'où

$$\det(\lambda \mathrm{I}_n - P^{-1}AP) = \det\left(P^{-1}(\lambda \mathrm{I}_n - A)P\right) = \det(\lambda \mathrm{I}_n - A),$$

ce qui montre la proposition. □

Ainsi, le polynôme caractéristique χ_A ne dépend que de la classe de conjugaison de A, ce qui suggère la définition suivante.

Définition 2.1.12. Le polynôme caractéristique d'un endomorphisme $f \in \mathscr{L}(E)$ est le polynôme caractéristique de la matrice représentant f dans n'importe quelle base de E.

Proposition 2.1.13. *Pour tout* $f \in \mathscr{L}(E)$, *on a* $\lambda \in \mathrm{sp}(f)$ *si, et seulement si,* λ *est une racine de* χ_f.

Démonstration. En effet, on a $\lambda \in \mathrm{sp}(f)$ si, et seulement si $\lambda\,\mathrm{id} - f$ n'est pas inversible ce qui équivaut à $\det(\lambda\,\mathrm{id} - f) = 0$ ou encore $\chi_f(\lambda) = 0$. □

Définition 2.1.14. Soit $f \in \mathscr{L}(E)$. La *multiplicité algébrique* $m_f^{\mathrm{alg}}(\lambda)$ d'une valeur propre $\lambda \in \mathrm{sp}(f)$ est la multiplicité de λ comme racine de χ_f, c'est-à-dire

$$m_f^{\mathrm{alg}}(\lambda) = \max\left\{k \in \mathbb{N} \;:\; (X-\lambda)^k \text{ divise } \chi_f(X)\right\}.$$

Proposition 2.1.15. *Pour tout* $f \in \mathscr{L}(E)$ *et* $\lambda \in \mathrm{sp}(f)$, *on a*

$$1 \leqslant m_f^{\mathrm{alg}}(\lambda) \leqslant \dim E_\lambda.$$

Démonstration. Soit $m = \dim E_\lambda$. On a bien sûr $1 \leqslant m$ puisque $\lambda \in \mathrm{sp}(f)$. Soit $(e_1, \ldots, e_m)$ une base de E_λ, que l'on complète en une base β de E. La matrice de f dans la base β est alors de la forme

$$[f]_\beta = \begin{pmatrix} \mathrm{I}_m & A \\ 0 & B \end{pmatrix}.$$

Par conséquent on a

$$[X\,\mathrm{id} - f]_\beta = \begin{pmatrix} (X-\lambda)\mathrm{I}_m & -A \\ 0 & X\mathrm{I}_{n-m} - B \end{pmatrix}$$

et il suit que

$$\chi_f(X) = \det[X\,\mathrm{id} - f]_\beta = (X-\lambda)^m \det(X\mathrm{I}_{n-m} - B) = (X-\lambda)^m \chi_B(X).$$

En particulier $(X-\lambda)^m$ divise χ_f et donc $m \leqslant m_f^{\mathrm{alg}}(\lambda)$ par définition de $m_f^{\mathrm{alg}}(\lambda)$. □

Théorème 2.1.16 (Critère de diagonalisabilité, II). *Un endomorphisme $f \in \mathscr{L}(E)$ est diagonalisable si, et seulement si, les deux propriétés suivantes sont vérifiées :*

(i) *χ_f est scindé sur* K *;*

(ii) *pour toute valeur propre $\lambda \in \mathrm{sp}(f)$, on a $m_f^{\mathrm{alg}}(\lambda) = m_f^{\mathrm{geom}}(\lambda)$.*

Démonstration. Soit $f \in \mathscr{L}(E)$ et notons $\mathrm{sp}(f) = \{\lambda_1, \dots, \lambda_r\}$ où les λ_j sont deux à deux distincts. Supposons $f \in \mathscr{L}(E)$ diagonalisable. Alors $E = \bigoplus_{j=1}^r E_{\lambda_j}$ par la Proposition 2.1.7. En choisissant une base β adaptée à cette décomposition, la matrice $[f]_\beta$ est diagonale ; plus précisément on a

$$[f]_\beta = \begin{pmatrix} \lambda_1 \mathrm{I}_{m_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_r \mathrm{I}_{m_r} \end{pmatrix}$$

où $m_j = \dim E_{\lambda_j}$ pour $j = 1, \dots, r$. Il suit que $\chi_f = \prod_{j=1}^r (X-\lambda_j)^{m_j}$ est scindé, et comme les λ_j sont deux à deux distincts, on a immédiatement $m_j = m_f^{\mathrm{alg}}(\lambda_j)$.

Réciproquement, supposons (i) et (ii). Alors on a

$$\begin{aligned} \dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}) &= \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_r} \\ &= m_f^{\mathrm{alg}}(\lambda_1) + \cdots + m_f^{\mathrm{alg}}(\lambda_r) \\ &= n. \end{aligned}$$

La première égalité vient du fait que les E_{λ_j} sont en somme directe, la deuxième découle du point (ii) et enfin la troisième résulte du point (i). Ainsi on a montré que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$ et donc f est diagonalisable par la Proposition 2.1.7. □

Définition 2.1.17 (Matrice compagnon d'un polynôme unitaire). Soit $P \in \mathrm{K}[X]$ un polynôme unitaire de degré n, que l'on écrit

$$P = X^n + \sum_{k=0}^{n-1} a_k X^k.$$

La *matrice compagnon* de P est la matrice

$$C(P)=\begin{pmatrix}0&\cdots&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&\ddots&\ddots&\vdots&\vdots\\\vdots&\ddots&\ddots&0&\vdots\\0&\cdots&0&1&-a_{n-1}\end{pmatrix}\in \mathrm{M}_n(\mathrm{K}).$$

PROPOSITION 2.1.18. *Pour tout polynôme $P\in\mathrm{K}[X]$ de degré n, on a $\chi_{C(P)}=P$.*

Démonstration. On a

$$\chi_{C(P)}=\begin{vmatrix}X&\cdots&\cdots&0&a_0\\-1&X&\cdots&0&a_1\\0&\ddots&\ddots&\vdots&\vdots\\\vdots&\ddots&\ddots&X&\vdots\\0&\cdots&&-1&X+a_{n-1}\end{vmatrix}.$$

Pour tout $j=2,\dots,n$, on ajoute la j^{e} ligne multipliée par X^{j-1} à la première ; on obtient

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

En développant par rapport à la première ligne, il vient

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

ce qui est bien le résultat voulu. □

2.1.3 Diagonalisation, trigonalisation

THÉORÈME 2.1.19 (Critère de trigonalisabilité). *Un endormorphisme $f\in\mathscr{L}(E)$ est* trigonalisable *sur* K *si, et seulement si, χ_f est scindé sur* K.

On rappelle qu'un polynôme $P \in \mathrm{K}[X]$ est scindé sur K si, et seulement si, il existe $a, \alpha_1, \ldots, \alpha_n \in \mathrm{K}$ tels que

$$P = a \prod_{k=1}^{n} (X - \alpha_k).$$

Puisque tout polynôme est scindé sur $\mathbb{C}$ par le théorème fondamental de l'algèbre, on obtient immédiatement le corollaire suivant.

Corollaire 2.1.20. *Soit E est un $\mathbb{C}$-espace vectoriel de dimension finie. Alors tout endomorphisme de E est trigonalisable.*

Démonstration du Théorème 2.1.19. En vertu de la remarque 2.1.6, il suffit de montrer le résultat pour les matrices. Supposons que $A \in \mathrm{M}_n(\mathrm{K})$ soit trigonalisable sur K. Alors il existe $P \in \mathrm{GL}_n(\mathrm{K})$ et $\lambda_1, \ldots, \lambda_n \in \mathrm{K}$ tels que

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Ainsi on obtient que

$$\chi_A = \chi_{P^{-1}AP} = \det(X\mathrm{I}_n - P^{-1}AP) = (X - \lambda_1) \cdots (X - \lambda_n)$$

est scindé sur K.

On montre la réciproque par récurrence sur n. Elle est claire pour $n = 1$. Supposons-la vraie pour un certain $n \geqslant 1$, et soit $A \in \mathrm{M}_{n+1}(\mathrm{K})$ telle que χ_A est scindé sur K. En particulier, χ_A admet une racine λ_1. Ainsi λ_1 est valeur propre de A par la Proposition 2.1.13. Soit $x \in \mathrm{M}_{n+1,1}(\mathrm{K})$ un vecteur propre associé, que l'on complète en une base $\beta = (x, e_2, \ldots, e_{n+1})$ de $\mathrm{M}_{n+1,1}(\mathrm{K})$. Si P est la matrice de passage de la base canonique de $\mathrm{M}_{n+1,1}(\mathrm{K})$ à β, on a que $P^{-1}AP$ est de la forme

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & \ell \\ 0 & B \end{pmatrix}$$

où $\ell \in \mathrm{M}_{1,n}(\mathrm{K})$ et $B \in \mathrm{M}_n(\mathrm{K})$. En particulier on obtient

$$\chi_A = \chi_{P^{-1}AP} = (X - \lambda_1)\chi_B.$$

Comme χ_A est scindé sur K, il en est de même pour χ_B et par hypothèse de récurrence, il existe $Q \in \mathrm{GL}_n(\mathrm{K})$ telle que $Q^{-1}BQ$ est triangulaire supérieure. En notant $\widetilde{Q} = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$, on a $\widetilde{Q}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix}$ et la matrice

$$\widetilde{Q}^{-1}P^{-1}AP\widetilde{Q} = \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & \ell \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} 1 & \ell Q \\ 0 & Q^{-1}BQ \end{pmatrix}$$

est triangulaire supérieure. La récurrence est établie. □

2.2 Polynômes d'endomorphismes

2.2.1 Opérations sur les endomorphismes

On commence par introduire la notion de polynômes d'endomorphismes. On rappelle qu'on a les opérations suivantes pour $\lambda \in \mathrm{K}$ et $f, g \in \mathscr{L}(E)$:

- le produit extérieur $\lambda f \in \mathscr{L}(E)$, donné par
$$(\lambda f)(x) = \lambda f(x), \quad x \in E \ ;$$
- la somme $f + g \in \mathscr{L}(E)$, donnée par
$$(f+g)(x) = f(x) + g(x), \quad x \in E \ ;$$
- la composition $f \circ g \in \mathscr{L}(E)$, donnée par
$$(f \circ g)(x) = f(g(x)), \quad x \in E.$$

Pour $f \in \mathscr{L}(E)$ et $k \in \mathbb{N}$, l'endomorphisme $f^k \in \mathscr{L}(E)$ est défini par les relations de récurrence
$$f^0 = \mathrm{id}_E \quad \text{et} \quad f^{k+1} = f \circ f^k, \quad k \in \mathbb{N}.$$
Autrement dit, pour tout $k \geqslant 1$ on a
$$f^k = \underbrace{f \circ \cdots \circ f}_{k \text{ fois}}.$$

Notons qu'on a
$$f^{k+\ell} = f^k \circ f^\ell = f^\ell \circ f^k, \quad k, \ell \in \mathbb{N}.$$

Définition 2.2.1 (Polynôme d'endomorphisme). Soit $P = \sum_{k=0}^N a_k X^k \in \mathrm{K}[X]$ et $f \in \mathscr{L}(E)$. Alors $P(f)$ est défini par
$$P(f) = \sum_{k=0}^N a_k f^k \in \mathscr{L}(E).$$

On dira qu'un endomorphisme $g \in \mathscr{L}(E)$ est un *polynôme en f* s'il existe un polynôme $Q \in \mathrm{K}[X]$ tel que $g = Q(f)$.

Remarque 2.2.2. Cette notion est en adéquation avec celle des polynômes de matrices. En effet si $f \in \mathscr{L}(E)$, $P = \sum_k a_k X^k \in \mathrm{K}[X]$ et β est une base de E, on a
$$[P(f)]_\beta = P([f]_\beta),$$

où pour toute matrice A, la matrice $P(A) = \sum_k a_k A^k$ est définie à l'aide du produit matriciel. En effet, cela résulte de ce qu'on a

$$[f \circ g]_\beta = [f]_\beta [g]_\beta$$

pour tous $f, g \in \mathscr{L}(E)$.

Les proposition suivantes seront utiles. Leurs démonstrations sont de simples vérifications et sont laissées en exercice.

Proposition 2.2.3. *Soient $P, Q \in \mathrm{K}[X]$, $\lambda \in \mathrm{K}$ et $f \in \mathscr{L}(E)$. Alors*

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

(ii) $(PQ)(f) = P(f) \circ Q(f)$;

(iii) $(\lambda P)(f) = \lambda P(f)$.

Proposition 2.2.4. *Soient $f, g \in \mathscr{L}(E)$ tels que $f \circ g = g \circ f$. Alors pour tous $P, Q \in \mathrm{K}[X]$ on a*

$$P(f) \circ Q(g) = Q(g) \circ P(f).$$

2.2.2 Lemme des noyaux

Un résultat très utile pour la réduction des endomorphismes est le suivant.

Théorème 2.2.5 (Lemme des noyaux). *Soit $f \in \mathscr{L}(E)$, $P_1, \ldots, P_r \in \mathrm{K}[X]$ des polynômes deux à deux premiers entre eux et $P = P_1 \cdots P_k$. Alors on a*

$$\ker P(f) = \ker P_1(f) \oplus \cdots \oplus \ker P_r(f).$$

Démonstration. On montre d'abord le résultat pour $r = 2$, le cas général s'obtenant par récurrence. Soient donc $P_1, P_2 \in \mathrm{K}[X]$ deux polynômes premiers entre eux et $P = P_1 P_2$. On note $F_j = \ker P_j(f)$ pour $j = 1, 2$ et $F = \ker P(f)$. Remarquons d'abord que $F_1, F_2 \subset F$ d'où $F_1 + F_2 \subset F$. Par ailleurs, soient $x_j \in F_j$ pour $j = 1, 2$, tels que

$$x_1 + x_2 = 0. \tag{2.3}$$

Comme les P_j sont premiers entre eux, il existe des polynômes $Q_1, Q_2 \in \mathrm{K}[X]$ tels que $P_1 Q_1 + P_2 Q_2 = 1$, ce qui implique

$$(P_1 Q_1)(f) + (P_2 Q_2)(f) = \mathrm{id}_E\,. \tag{2.4}$$

Notons que $x_1 \in F_1$ et donc $(P_1 Q_1)(f)(x_1) = 0$; ainsi (2.3) et (2.4) impliquent

$$0 = (P_1 Q_1)(f)(x_1 + x_2) = x_2 - (P_2 Q_2)(f)(x_2) = x_2.$$

Ainsi $x_2 = 0$ et $x_1 = 0$: on a montré que F_1 et F_2 sont en somme directe. Il reste à montrer que $F \subset F_1 + F_2$. Pour $x \in F$, on utilise (2.4) pour écrire

$$x = (P_1Q_1)(f)(x) + (P_2Q_2)(f)(x) = x_2 + x_1.$$

On a $x_2 \in F_2$ car $P_2(f)(x_2) = (P_2P_1Q_1)(f)(x) = 0$ puisque $x \in F$. De même $x_1 \in F_1$ et le résultat est démontré pour $r = 2$.

On suppose maintenant le résultat vrai pour un certain $r \geqslant 2$ et on se donne $P_1, \dots, P_{r+1}$ des polynômes deux à deux premiers entre eux. On pose

$$\tilde{P}_1 = P_1 \quad \text{et} \quad \tilde{P}_2 = P_2 \cdots P_{r+1}.$$

Alors $\tilde{P}_1$ et $\tilde{P}_2$ sont premiers entre eux, donc par le résultat montré pour $r = 2$ on obtient $\ker P(f) = \ker(\tilde{P}_1\tilde{P}_2)(f) = \ker \tilde{P}_1(f) \oplus \ker \tilde{P}_2(f)$. D'autre part, les polynômes $P_2, \dots, P_{r+1}$ sont deux à deux premiers entre eux, donc par hypothèse de récurrence,

$$\ker \tilde{P}_2(f) = \ker(P_2 \dots P_{r+1})(f) = \ker P_2(f) \oplus \cdots \ker P_{r+1}(f).$$

La récurrence est établie. □

Remarque 2.2.6. Soit $f \in \mathscr{L}(E)$ et $\lambda_1, \dots, \lambda_r \in \mathrm{K}$ des valeurs propres de f, deux à deux distinctes. Alors le lemme des noyaux appliqué à $P = (X - \lambda_1) \cdots (X - \lambda_r)$ permet de montrer que les espaces propres $E_{\lambda_1}, \dots, E_{\lambda_r}$ sont en somme directe : on a retrouvé la Proposition 2.1.2.

Corollaire 2.2.7. *Soit $f \in \mathscr{L}(E)$. Si $P \in \mathrm{K}[X]$ est scindé à racines simples et vérifie $P(f) = 0$ alors f est diagonalisable.*

Démonstration. Comme P est scindé à racines simples, on peut écrire

$$P = \prod_{j=1}^{r} (X - \lambda_j)$$

pour des $\lambda_1, \dots, \lambda_r \in \mathrm{K}$ deux à deux distincts. En appliquant le lemme des noyaux aux polynômes $X - \lambda_j$, on obtient

$$E = \ker P(f) = \bigoplus_{j=1}^{r} \ker(f - \lambda_j \,\mathrm{id})$$

donc E est somme des sous-espaces propres de f, donc f est diagonalisable. □

2.2.3 Polynôme minimal

Nous introduisons à présent la notion de polynôme annulateur.

Définition 2.2.8 (Polynôme annulateur). Soit $P \in \mathrm{K}[X]$ et $f \in \mathscr{L}(E)$. On dit que P est un *polynôme annulateur* de f (ou plus simplement que P annule f) si on a

$$P(f) = 0.$$

On notera $\mathcal{I}(f) \subset \mathscr{L}(E)$ l'ensemble des polynômes annulateurs de f.

Remarque 2.2.9 (Existence d'un polynôme annulateur). Il existe toujours un polynôme annulateur non nul pour un endomorphisme $f \in \mathscr{L}(E)$. En effet, $\mathscr{L}(E)$ est un espace vectoriel de dimension n^2 et par conséquent la famille $(\mathrm{id}_E, f, \cdots, f^{n^2})$ est liée ; par suite il existe $a_0, \ldots, a_{n^2} \in \mathrm{K}$ non tous nuls tels que

$$a_0 \,\mathrm{id}_E + \cdots + a_{n^2} f^{n^2} = 0.$$

Autrement dit, le polynôme $\sum_{k=0}^{n^2} a_k X^k$ annule f.

Proposition 2.2.10 (Propriétés de $\mathcal{I}(f)$). *Soit $f \in \mathscr{L}(E)$. Alors*

(i) *$\mathcal{I}(f)$ est sous-un espace vectoriel de $\mathrm{K}[X]$;*

(ii) *pour tout $P \in \mathcal{I}(f)$ et tout $Q \in \mathrm{K}[X]$, $PQ \in \mathcal{I}(f)$ — on dit que $\mathcal{I}(f)$ est un idéal ;*

(iii) *si $P \in \mathcal{I}(f)$ et $\lambda \in \mathrm{sp}(f)$ alors $P(\lambda) = 0$.*

Démonstration. Pour le point (i), on remarque que pour tout $\lambda \in \mathrm{K}$ et tous $P, Q \in \mathcal{I}(f)$, on a $(\lambda P + Q)(f) = \lambda P(f) + Q(f) = 0$ donc $\lambda P + Q \in \mathcal{I}(f)$. Pour le point (ii), on remarque que $(PQ)(f) = P(f) \circ Q(f) = 0$. Enfin pour le dernier point, on prend $P = \sum_{k=0}^{d} a_k X^k \in \mathcal{I}(f)$ et $\lambda \in \mathrm{sp}(f)$. Si $x \in E$ vérifie $f(x) = \lambda x$, avec x non nul, on a $f^k(x) = \lambda^k x$ pour tout k, d'où

$$0 = P(f)(x) = \sum_{k=0}^{d} a_k f^k(x) = \sum_{k=0}^{d} a_k \lambda^k x = P(\lambda) x$$

ce qui implique $P(\lambda) = 0$. □

Le théorème suivant caractérise l'idéal $\mathcal{I}(f)$.

Théorème–Définition 2.2.11. *Soit $f \in \mathscr{L}(E)$. Il existe un unique polynôme unitaire $\mu_f \in \mathrm{K}[X]$ tel que*

$$\mathcal{I}(f) = \mu_f \cdot \mathrm{K}[X] = \{\mu_f \cdot P \ : \ P \in \mathrm{K}[X]\}.$$

Le polynôme μ_f est le polynôme minimal *de f.*

Autrement dit, le polynôme minimal μ_f vérifie la propriété suivante :

Un polynôme $P \in \mathrm{K}[X]$ annule f si, et seulement si, c'est un multiple de μ_f.

Démonstration. On commence par l'existence de μ_f. On vient de voir dans la remarque 2.2.9 que $\mathcal{I}(f) \setminus \{0\}$ était non vide. Ainsi, on peut définir

$$m = \min\{\deg(P) \ : \ P \in \mathcal{I}(f) \setminus \{0\}\},$$

et on choisit un polynôme $P_0 \in \mathcal{I}(f)$ unitaire avec $\deg(P_0) = m$. Soit maintenant $P \in \mathcal{I}(f)$. On effectue la division euclidienne de P par P_0, en écrivant $P = QP_0 + R$ avec $\deg R < m$. On a alors $0 = P(f) = Q(f) \circ P_0(f) + R(f) = R(f)$ puisque $P_0 \in \mathcal{I}(f)$. Ainsi R annule f, et comme $\deg R < m$ on a nécessairement $R = 0$ par définition de m. Ainsi $P = QP_0$, et on a montré que $\mathcal{I}(f) = P_0 \cdot \mathrm{K}[X]$. Montrons à présent que P_0 est unique : si $P_1 \in \mathcal{I}(f)$ vérifie $\mathcal{I}(f) = P_1 \cdot \mathrm{K}[X]$, alors P_1 est un multiple de P_0, et réciproquement ; comme les deux polynômes sont unitaires, on en déduit qu'ils sont égaux. Ceci conclut la démonstration. □

Proposition 2.2.12 (Critère de diagonalisabilité, III). *Soit $f \in \mathscr{L}(E)$. Alors les propriétés suivantes sont équivalentes :*

(i) *f est diagonalisable ;*

(ii) *il existe un polynôme $P \in \mathrm{K}[X]$ scindé à racines simples qui annule f ;*

(iii) *μ_f est scindé à racines simples.*

Démonstration. Si f est diagonalisable, il existe une base β, des scalaires $\lambda_1, \ldots, \lambda_r \in \mathrm{K}$ deux à deux distincts et $n_1, \ldots, n_r \geqslant 1$ des entiers tels que

$$[f]_\beta = \begin{pmatrix} \boxed{\lambda_1 \mathrm{I}_{n_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_r \mathrm{I}_{n_r}} \end{pmatrix}.$$

Dès lors, si $P = \prod_{j=1}^r (X - \lambda_j)$, on a

$$[P(f)]_\beta = P([f]_\beta) = \begin{pmatrix} \boxed{P(\lambda_1) \mathrm{I}_{n_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{P(\lambda_r) \mathrm{I}_{n_r}} \end{pmatrix} = 0.$$

Ainsi P annule f et est scindé à racines simples. Supposons maintenant qu'il existe un polynôme P scindé à racines simples qui annule f. Alors μ_f divise P donc μ_f

est aussi scindé à racines simples. Enfin, supposons $\mu_f = \prod_{j=1}^{r}(X - \lambda_j)$ scindé à racines simples. Alors $\mu_f(f) = 0$ donc le lemme des noyaux donne

$$E = \ker \mu_f(f) = \bigoplus_{j=1}^{r} \ker(f - \lambda_j \,\mathrm{id}) = \bigoplus_{j=1}^{r} E_{\lambda_j}.$$

Ainsi E est somme directe des espaces propres de f, donc f est diagonalisable. On a bien montré que (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i). Ceci conclut la démonstration. $\square$

On conclut ce paragraphe par une conséquence de ce critère qui est très utile en pratique.

Lemme 2.2.13. *Soit $f \in \mathscr{L}(E)$ un endomorphisme diagonalisable et $F \subset E$ un sous-espace de E qui est stable par f, i.e. $f(F) \subset F$. Alors la restriction $g = f|_F \in \mathscr{L}(F)$ est diagonalisable.*

Démonstration. Puisque f est diagonalisable, μ_f est scindé à racines simples. Comme μ_f annule f, il annule aussi g, donc g admet un polynôme annulateur scindé à racines simples. Par le corollaire 2.2.7, g est diagonalisable. $\square$

2.2.4 Théorème de Cayley–Hamilton

On termine cette section en énonçant un théorème crucial, celui de Cayley–Hamilton.

Théorème 2.2.14 (Cayley–Hamilton). *Pour tout $f \in \mathscr{L}(E)$, on a $\chi_f(f) = 0$.*

Ce théorème nous dit que $\chi_f \in \mathcal{I}(f)$, ce qui équivaut à dire que χ_f est un multiple de μ_f par le théorème-Définition 2.2.11.

Démonstration. Soit $x \in E$ non nul : on veut montrer que $\chi_f(f)(x) = 0$. On pose $\nu = \min\{k \in \mathbb{N} \ : \ (x, f(x) \dots, f^k(x)) \text{ est liée}\}$. Alors $(x, f(x), \dots, f^{\nu-1}(x))$ est libre, et $f^\nu(x)$ est combinaison linéaire des $f^k(x)$ avec $0 \leqslant k \leqslant \nu - 1$, de sorte qu'il existe $a_0, \dots, a_{\nu-1} \in \mathrm{K}$ tels que

$$a_0 x + \cdots + a_{\nu-1} f^{\nu-1}(x) + f^\nu(x) = 0. \tag{2.5}$$

On complète la famille $(x, \dots, f^{\nu-1}(x))$ en une base β de E. En utilisant (2.5) on obtient alors que la matrice de f dans la base β est de la forme

$$[f]_\beta = \begin{pmatrix} C(P) & \star \\ 0 & B \end{pmatrix}$$

avec $B \in \mathrm{M}_{n-\nu}(\mathrm{K})$ et $P = X^\nu + \sum_{k=0}^{\nu-1} a_k X^k$. On en déduit que

$$\chi_f = \chi_{C(P)} \cdot \chi_B = P \cdot \chi_B$$

où on a utilisé l'égalité $\chi_{C(P)} = P$ qui provient de la Proposition 2.1.18. Notons que (2.5) implique que $P(f)(x) = 0$. On en déduit que

$$\chi_f(f)(x) = (P\chi_B)(f)(x) = \big(\chi_B(f) \circ P(f)\big)(x) = \chi_B(f)\big(P(f)(x)\big) = 0,$$

ce qui achève la démonstration. □

Un corollaire du théorème de Cayley-Hamilton est que les racines de μ_f sont exactement les valeurs propres de f.

COROLLAIRE 2.2.15. *Soit $f \in \mathscr{L}(E)$. Alors $\lambda \in \mathrm{K}$ est racine de μ_f si et seulement si $\lambda \in \mathrm{sp}(f)$.*

Démonstration. Puisque $\chi_f(f) = 0$, μ_f divise χ_f. Ainsi les racines de μ_f sont aussi des racines χ_f. Par suite les racines de μ_f sont des valeurs propres de f. Réciproquement, soit $\lambda \in \mathrm{K}$ une valeur propre de f. Alors $\mu_f(\lambda) = 0$ par le dernier point de la Proposition 2.2.10. □

Une deuxième conséquence est la caractérisation suivante de la trigonalisabilité.

PROPOSITION 2.2.16 (Critère de trigonalisabilité, II). *Soit $f \in \mathscr{L}(E)$. Alors les trois propriétés suivantes sont équivalentes :*

(i) *f est trigonalisable ;*

(ii) *il existe un polynôme $P \in \mathrm{K}[X]$ scindé tel que $P(f) = 0$;*

(iii) *μ_f est scindé.*

REMARQUE 2.2.17. Cette proposition combinée au Théorème 2.1.19 implique que pour tout $f \in \mathscr{L}(E)$, μ_f est scindé ssi χ_f l'est.

Démonstration. Supposons (i), i.e. f est trigonalisable. Alors χ_f est scindé et $\chi_f(f) = 0$ par le théorème de Cayley-Hamilton, donc (ii) est vérifié. Supposons (ii) : il existe $P \in \mathrm{K}[X]$ scindé tel que $P(f) = 0$. Comme μ_f divise P, on en déduit que μ_f est scindé donc (iii) est vérifié.

Il reste à montrer que (iii) implique (i). On raisonne par récurrence sur la dimension, et on suppose le résultat vrai pour les espaces vectoriels de dimension $n-1$. Soit E un espace de dimension n et $f \in \mathscr{L}(E)$ tel que μ_f est scindé. Alors μ_f a une racine $\lambda \in \mathrm{K}$, qui est une valeur propre de f par le corollaire 2.2.15. Soit $x \in E$ non nul un vecteur propre associé, que l'on complète en une base $\beta = (x, e_2, \ldots, e_n)$ de E. Alors $[f]_\beta$ est de la forme

$$\begin{pmatrix} \lambda & \star \\ 0 & B \end{pmatrix}$$

avec $B \in \mathrm{M}_n(\mathrm{K})$, si bien que pour tout polynôme $P \in \mathrm{K}[X]$, la matrice $P([f]_\beta)$ est de la forme

$$\begin{pmatrix} P(\lambda) & \star \\ 0 & P(B) \end{pmatrix}.$$

Dans le cas particulier où $P = \mu_f$, on a $\mu_f([f]_\beta) = [\mu_f(f)]_\beta = 0$, ce qui implique $\mu_f(B) = 0$. Ainsi $\mu_f(g) = 0$ où g est l'endomorphisme de K^{n-1} canoniquement associé à B. Ainsi μ_g divise μ_f, qui est scindé ; par conséquent μ_g l'est aussi. Par hypothèse de récurrence, on obtient que g, et donc B, sont trigonalisables. On peut alors procéder exactement comme dans la preuve du Théorème 2.1.19 pour conclure que f est trigonalisable. La récurrence est établie. □

On conclut ce paragraphe avec une deuxième conséquence du théorème de Cayley–Hamilton.

Proposition 2.2.18 (Polynôme minimal d'une matrice compagnon). *Soit* $P = a_0 + \cdots + a_{n-1}X^{n-1} + X^n \in \mathrm{K}[X]$ *un polynôme unitaire. Alors*

$$\chi_{C(P)} = \mu_{C(P)} = P.$$

Démonstration. Comme $\mu_{C(P)}$ divise $\chi_{C(P)}$ par le théorème de Cayley–Hamilton et que les deux polynômes sont unitaires, il suffit de montrer que $\deg \mu_{C(P)} = n$. Pour cela on va montrer qu'il n'existe aucun polynôme non nul de degré strictement inférieur à n qui annule $C(P)$. En effet, soit $f \in \mathscr{L}(\mathrm{K}^n)$ l'endomorphisme canoniquement associé à $C(P)$. Soit $Q = \sum_{k=0}^d b_k X^k$ un polynôme annulateur de f de degré $d < n$. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ la base canonique de K^n. Alors $f^k(e_1) = e_{k+1}$ pour tout $k < n$, si bien que

$$0 = Q(f)(e_1) = \sum_{k=0}^{n-1} b_k f^k(e_1) = \sum_{k=0}^{n-1} b_k e_{k+1}.$$

Comme la famille $(e_1, \ldots, e_n)$ est libre, on en déduit $b_k = 0$ pour tout k, donc $Q = 0$. Il suit que $\deg \mu_f = \deg \mu_{C(P)} = n$, ce qui conclut la démonstration. □

2.3 Réduction de Dunford et de Jordan

Dans toute la suite, si F est un K-espace vectoriel de dimension finie, on note $\mathscr{N}(F)$ l'ensemble des endomorphismes nilpotents de F, c'est-à-dire

$$\mathscr{N}(F) = \{f \in \mathscr{L}(F) \ : \ \exists m \in \mathbb{N}^*, \ f^m = 0\}.$$

Si $f \in \mathscr{N}(F)$, l'*indice de nilpotence* de f est l'entier

$$\nu = \inf\{m \in \mathbb{N}^* \ : \ f^m = 0\}.$$

Comme application de la Proposition 2.2.16, on a le résultat suivant.

Proposition 2.3.1 (Polynôme caractéristique d'un endomorphisme nilpotent). *Si $f \in \mathscr{N}(F)$, on a $\chi_f = X^{\dim F}$ et $\mu_f = X^\nu$ où ν est l'indice de nilpotence de f.*

Démonstration. On a $f^m = 0$ avec $m \geqslant 1$, donc μ_f divise X^m. Ainsi $\mu_f = X^p$ pour un certain $p \geqslant 1$. En particulier, μ_f est scindé, donc χ_f est scindé par la remarque 2.2.17. Or $\mathrm{sp}(f) = \{0\}$ par le corollaire 2.2.15, donc 0 est la seule racine de χ_f. Ceci implique $\chi_f = X^{\dim F}$. Montrons que $\mu_f = X^\nu$. Comme X^ν annule f (car $f^\nu = 0$), on a que μ_f divise X^ν donc $p \leqslant \nu$. Mais μ_f doit annuler f, donc $p = \nu$ par minimalité de ν. □

2.3.1 Sous-espaces caractéristiques

Définition 2.3.2. Soit $f \in \mathscr{L}(E)$ et $\lambda \in \mathrm{sp}(f)$. L'*espace caractéristique* C_λ associé à λ est défini par

$$C_\lambda = \ker\Big((f - \lambda\,\mathrm{id})^{m_f^{\mathrm{alg}}(\lambda)}\Big).$$

Remarque 2.3.3. Les C_λ sont stables par f, au sens où $f(C_\lambda) \subset C_\lambda$.

Proposition 2.3.4. *Pour tout $\lambda \in \mathrm{sp}(f)$, on a*

$$f_\lambda = \lambda\,\mathrm{id}_{C_\lambda} + h_\lambda, \quad \textit{où} \quad f_\lambda = f|_{C_\lambda} \quad \textit{et} \quad h_\lambda \in \mathscr{N}(C_\lambda).$$

Démonstration. On pose $h_\lambda = f_\lambda - \lambda\mathrm{id}_{C_\lambda}$. Alors $f_\lambda = \lambda\mathrm{id}_{C_\lambda} + h_\lambda$. Soit $x \in C_\lambda$. En notant $m = m_f^{\mathrm{alg}}(\lambda)$, on a, puisque C_λ est préservé par $f - \lambda\,\mathrm{id}_E$,

$$h_\lambda^m(x) = (f_\lambda - \lambda\mathrm{id}_{C_\lambda})^m(x) = (f - \lambda\,\mathrm{id}_E)^m(x) = 0$$

par définition de $C_\lambda = \ker\big((f - \lambda\,\mathrm{id}_E)^m\big)$. □

Proposition 2.3.5 (Décomposition en sous-espaces caractéristiques). *Soit $f \in \mathscr{L}(E)$ un endomorphisme trigonalisable. Alors*

$$E = \bigoplus_{\lambda \in \mathrm{sp}(f)} C_\lambda.$$

En particulier, si $\mathrm{K} = \mathbb{C}$, alors E est la somme directe des espaces caractéristiques associés à n'importe quel endomorphisme.

Démonstration. Comme f est trigonalisable, son polynôme caractéristique

$$\chi_f(\lambda) = \prod_{j=1}^{r} (\lambda - \lambda_j)^{m_j}$$

est scindé. Ici les λ_j sont distincts et $m_j = m_f^{\mathrm{alg}}(\lambda_j)$. Le théorème de Cayley–Hamilton (Théorème 2.2.14) nous dit que

$$E = \ker \chi_f(f).$$

Par ailleurs, le lemme des noyaux (Théorème 2.2.5) donne

$$\ker \chi_f(f) = \bigoplus_{j=1}^{r} \ker\big((f - \lambda_j \,\mathrm{id})^{m_j}\big).$$

Avec les deux dernières inégalités on obtient $E = \bigoplus_{j=1}^{r} C_{\lambda_j}$, ce qu'on voulait démontrer. $\square$

Proposition 2.3.6 (Propriétés des sous-espaces caractéristiques). *Soit $f \in \mathscr{L}(E)$ tel que χ_f est scindé. Soit $\lambda \in \mathrm{sp}(f)$ et $f_\lambda = f|_{C_\lambda}$. On note m_λ et ν_λ les ordres de λ en tant que racine de χ_f et de μ_f, respectivement. Alors*

(i) $\dim C_\lambda = m_\lambda$;

(ii) $C_\lambda = \ker(f - \lambda \,\mathrm{id})^{\nu_\lambda}$.

Démonstration. En choisissant une base adaptée à la somme directe $\bigoplus_{\lambda\in\mathrm{sp}(f)} C_\lambda$, on voit que $\chi_f = \prod_{\lambda\in\mathrm{sp}(f)} \chi_{f_\lambda}$. Puisque $f_\lambda = \lambda\,\mathrm{id} + h_\lambda$ avec h_λ nilpotente par la Proposition 2.3.4, on a $\chi_{f_\lambda}(X) = \chi_{h_\lambda}(X-\lambda) = (X-\lambda)^{\dim C_\lambda}$ par la Proposition 2.3.1. Par conséquent, on obtient $\chi_f = \prod_{\lambda\in\mathrm{sp}(f)} (X-\lambda)^{\dim C_\lambda}$, ce qui signifie que $\dim C_\lambda = m_\lambda$. Pour le deuxième point, on remarque que le lemme des noyaux appliqué à $\mu_f = \prod_{\lambda\in\mathrm{sp}(f)} (X-\lambda)^{\nu_f}$ implique que E coïncide avec la somme directe des $\widetilde{C}_\lambda$ où $\widetilde{C}_\lambda = \ker(f-\lambda\,\mathrm{id})^{\nu_\lambda}$. Or $\nu_\lambda \leqslant m_\lambda$ donc $\widetilde{C}_\lambda \subset C_\lambda$. Comme E est aussi somme directe des C_λ, on conclut que $\widetilde{C}_\lambda = C_\lambda$. $\square$

2.3.2 Réduction de Dunford

Théorème 2.3.7 (Réduction de Dunford). *Soit $f \in \mathscr{L}(E)$ un endomorphisme trigonalisable. Alors il existe $\delta \in \mathscr{L}(E)$ diagonalisable et $h \in \mathscr{N}(E)$ tels que*

$$f = \delta + h \quad \textit{et} \quad \delta \circ h = h \circ \delta.$$

De plus, le couple (δ, h) est unique.

Démonstration. On procède par analyse-synthèse. Soient δ et h tels que $f = \delta + h$ avec δ diagonalisable, h nilpotent et $\delta \circ h = h \circ \delta$. Comme h et δ commutent, on a $f \circ \delta = (\delta + h) \circ \delta = \delta \circ (\delta + h) = \delta \circ f$, de sorte que δ commute avec f. Ainsi δ laisse stable les sous-espaces caractéristiques de f et on peut considérer

$\delta_\lambda = \delta|_{C_\lambda} \in \mathscr{L}(C_\lambda)$. De même h commute avec f et on note $h_\lambda = h|_{C_\lambda}$. Par la Proposition 2.3.4, on peut écrire $f_\lambda = \lambda \operatorname{id}_{C_\lambda} + \nu_\lambda$, de sorte que

$$\delta_\lambda = \lambda \operatorname{id}_{C_\lambda} + \nu_\lambda - h_\lambda.$$

Notons que h_λ commute avec f_λ et donc avec $\nu_\lambda = f_\lambda - \lambda \operatorname{id}_{C_\lambda}$. Par suite $\nu_\lambda - h_\lambda$ est nilpotente. En effet, si $m \in \mathbb{N}^*$ vérifie $\nu_\lambda^m = h_\lambda^m = 0$, on a

$$(\nu_\lambda - h_\lambda)^{2m} = \sum_{\ell=0}^{2m} (-1)^\ell \nu_\lambda^\ell h_\lambda^{2m-\ell} = 0,$$

puisqu'on a $\ell \geqslant m$ ou $2m - \ell \geqslant m$ pour tout $\ell = 0, \dots, 2m$. Par la Proposition 2.3.1, on obtient

$$\chi_{\delta_\lambda}(X) = \chi_{\nu_\lambda - h_\lambda}(X - \lambda) = (X - \lambda)^{m_\lambda}.$$

Ainsi λ est la seule valeur propre de δ_λ. Mais δ_λ est diagonalisable par le Lemme 2.2.13, donc $\delta_\lambda = \lambda \operatorname{id}_{C_\lambda}$, et $h_\lambda = \nu_\lambda = f_\lambda - \lambda \operatorname{id}_{C_\lambda}$.

Ainsi, on a montré l'unicité : si δ et h existent, alors nécessairement leurs restrictions à C_λ sont données respectivement par $\lambda \operatorname{id}_{C_\lambda}$ et $f_\lambda - \lambda \operatorname{id}_{C_\lambda}$ pour tout $\lambda \in \operatorname{sp}(f)$. Autrement dit, si $x = \sum_{\lambda \in \operatorname{sp}(f)} x_\lambda$ avec $x_\lambda \in C_\lambda$, on a

$$\delta(x) = \sum_{\lambda \in \operatorname{sp}(f)} \lambda x_\lambda \quad \text{et} \quad h(x) = \sum_{\lambda \in \operatorname{sp}(f)} (f(x_\lambda) - \lambda x_\lambda). \tag{2.6}$$

Réciproquement, on vérifie facilement en utilisant la Proposition 2.3.4 que les endomorphismes δ et h définis par la formule ci-dessus vérifient les conclusions du théorème. □

Soit $f \in \mathscr{L}(E)$ tel que χ_f est scindé. Pour tout $\lambda \in \operatorname{sp}(f)$, on définit l'application $\pi_\lambda : E \to C_\lambda$ comme étant la projection sur C_λ parallèlement à $\bigoplus_{\mu \neq \lambda} C_\mu$. Autrement dit,

$$\pi_\lambda \left(\sum_{\mu \in \operatorname{sp}(f)} x_\mu \right) = x_\lambda$$

pour tout $x = \sum_{\mu \in \operatorname{sp}(f)} x_\mu$ appartenant à $E = \bigoplus_{\mu \in \operatorname{sp}(f)} C_\mu$. Les applications π_λ sont appelés *projecteurs spectraux* et sont caractérisées par

$$\operatorname{im} \pi_\lambda = C_\lambda, \quad \pi_\lambda^2 = \pi_\lambda \quad \text{et} \quad \pi_\lambda \circ \pi_\mu = \pi_\mu \circ \pi_\lambda = 0 \tag{2.7}$$

pour toutes valeurs propres $\lambda, \mu \in \operatorname{sp}(f)$ telles que $\lambda \neq \mu$. Avec ces notations, l'équation (2.6) se ré-écrit

$$\delta = \sum_{\lambda \in \operatorname{sp}(f)} \lambda \pi_\lambda \quad \text{et} \quad h = f - \delta. \tag{2.8}$$

Proposition 2.3.8. *Soit $f \in \mathscr{L}(E)$ tel que χ_f est scindé. Alors les projecteurs π_λ, $\lambda \in \mathrm{sp}(f)$, sont des polynômes en f.*

Démonstration. On note $\chi_f = \prod_{\lambda \in \mathrm{sp}(f)} (X-\lambda)^{m_\lambda}$. Soit $\lambda \in \mathrm{sp}(f)$. On note

$$P = (X-\lambda)^{m_\lambda} \quad \text{et} \quad Q = \prod_{\mu \neq \lambda} (X-\mu)^{m_\mu}.$$

Alors P et Q sont premiers entre eux et PQ annule f, si bien que le lemme des noyaux nous donne

$$E = \ker(PQ)(f) = \ker P(f) \oplus \ker Q(f).$$

D'autre part, il existe $R, S \in \mathrm{K}[X]$ tels que $PR + QS = 1$. Notons $p = (QS)(f)$. Montrons que $p = \pi_\lambda$. D'abord, si $x \in \ker P(f) = C_\lambda$, on a $(PR)(f)(x) = 0$ d'où l'on tire

$$p(x) = (QS)(f)(x) = (1 - PR)(f)(x) = x.$$

D'autre part, si $x \in \ker Q(f)$, on a $p(x) = (QS)(f)(x) = 0$. Ceci montre que p est la projection sur C_λ parallèlement à $\ker Q(f)$. Mais par le lemme des noyaux on a

$$\ker Q(f) = \bigoplus_{\mu \neq \lambda} C_\mu,$$

donc $p = \pi_\lambda$. Ceci conclut puisque $p = (QS)(f)$ est un polynôme en f. □

Corollaire 2.3.9. *Les endomorphismes δ et h donnés par le Théorème 2.3.7 sont des polynômes en f.*

Démonstration. C'est une conséquence directe du résultat précédent et de l'identité (2.8). □

2.3.3 Réduction des endomorphismes nilpotents

Le but de ce paragraphe est de donner une forme normale pour les endomorphismes nilpotents.

Théorème 2.3.10 (Réduction des endomorphismes nilpotents). *Soit $f \in \mathscr{N}(E)$. Alors il existe une base β de E et des entiers $q_1, \ldots, q_s \in \mathbb{N}^*$ tels que*

$$[f]_\beta = \begin{pmatrix} \boxed{\mathrm{J}_{q_1}} & & (0) \\ & \ddots & \\ (0) & & \boxed{\mathrm{J}_{q_s}} \end{pmatrix}.$$

Ici $\mathrm{J}_q \in \mathrm{M}_q(\mathrm{K})$ *est la matrice dont tous les coefficients sont nuls, sauf ceux de la sur-diagonale qui valent* 1*, soit*

$$\mathrm{J}_q = \begin{bmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{bmatrix} \in \mathrm{M}_q(\mathrm{K}).$$

Commençons par un résultat intermédiaire.

Proposition 2.3.11 (Suite des noyaux). *Soit* $f \in \mathscr{L}(E)$*. Alors il existe un entier* $0 \leqslant \nu \leqslant n$ *tel que pour tout* $k \geqslant \nu$ *on a*

$$\{0\} = \ker \mathrm{id}_E \subsetneq \ker f \subsetneq \cdots \subsetneq \ker f^\nu = \ker f^{\nu+1} = \cdots = \ker f^k. \tag{2.9}$$

Démonstration. Pour tout $k \in \mathbb{N}$ on note $F_k = \ker f^k$. Alors $F_k \subset F_{k+1}$ pour tout k. Notons $d_k = \dim F_k$. La suite (d_k) est une suite croissante d'entiers inférieurs ou égaux à n, donc elle est stationnaire à partir d'un certain rang. Notons

$$\nu = \inf\{k \in \mathbb{N} \ : \ F_k = F_{k+1}\} = \inf\{k \in \mathbb{N} \ : \ d_k = d_{k+1}\}.$$

Montrons que $F_k \subsetneq F_{k+1}$ si et seulement si $k < \nu$, ce qui impliquera (2.9). D'abord, notons que $d_k \neq d_{k+1}$ pour tout $k < \nu$ par minimalité de ν, ce qui implique $F_k \subsetneq F_{k+1}$. Réciproquement, montrons que $k \leqslant \nu$ implique $F_k = F_{k+1}$. On procède par récurrence. D'abord, c'est vrai au rang ν par définition de ν. Supposons maintenant que $F_k = F_{k+1}$, et donnons-nous $x \in F_{k+2}$. On a $0 = f^{k+2}(x) = f^{k+1}(f(x))$, d'où $f(x) \in F_{k+1}$. Mais $F_{k+1} = F_k$, donc $f(x) \in F_k$ et $0 = f^k(f(x)) = f^{k+1}(x)$. Ainsi $x \in F_{k+1}$. Ainsi $F_{k+2} \subset F_{k+1}$ donc $F_{k+1} = F_{k+2}$. La récurrence est établie. □

Démonstration du Théorème 2.3.10. Soit $f \in \mathscr{N}(E)$. Nous allons montrer qu'il existe des sous-espaces $G_k \subset F_k$, $k = 1, \ldots, \nu$ tels que $F_k = F_{k-1} \oplus G_k$, avec $G_1 = F_1$ et $f|_{G_k}$ est injective $G_k \to G_{k-1}$ pour $k > 1$.

On raisonne par récurrence descendante. Pour $k = \nu$, on se donne un supplémentaire G_ν de $F_{\nu-1}$ dans F_ν. Alors $\ker f \cap G_\nu = F_1 \cap G_\nu \subset F_{\nu-1} \cap G_\nu = \{0\}$ donc $f|_{G_\nu}$ est injective. On suppose maintenant qu'on a construit $G_\nu, \ldots, G_k$ avec $k \geqslant 2$ tels que $G_\ell \oplus F_{\ell-1} = F_\ell$ pour tout $\ell = k, \ldots, \nu$, et $f|_{G_\ell} : G_\ell \to G_{\ell-1}$ est injective si $\ell > k$. Alors $f|_{G_k}$ est injective, puisque $\ker f \cap G_k = F_1 \cap G_k \subset F_{k-1} \cap G_k = \{0\}$. D'autre part, soit $y \in f(G_k) \cap F_{k-2}$. Alors il existe $x \in G_k$ tel que $y = f(x)$. Puisque $y \in F_{k-2}$ on a $x \in F_{k-1}$ donc $x \in G_k \cap F_{k-1} = \{0\}$ et $x = 0$ ce qui donne $y = 0$. Ainsi $u(G_k)$ est en somme directe avec F_{k-2}. On peut donc choisir un supplémentaire G_{k-1} de F_{k-2} dans F_{k-1} tel que $f(G_k) \subset G_{k-1}$. Ce sous-espace vérifie bien les propriétés voulues et la récurrence est établie.

Notons qu'on a $E = \ker f^\nu = F_\nu = G_\nu \oplus F_{\nu-1} = G_\nu \oplus G_{\nu-1} \oplus F_{\nu-2}$. Une récurrence immédiate donne

$$E = G_\nu \oplus \cdots \oplus G_1,$$

avec $u|_{G_k} : G_k \to G_{k-1}$ injective pour $1 < k \leqslant \nu$. Notons $s_k = \dim G_k$ pour tout $k = 1, \ldots, \nu$. On se donne une base $(e_{\nu,1}, \ldots, e_{\nu,s_\nu})$ de G_ν. Alors la famille $f(e_{\nu,1}), \ldots, f(e_{\nu,s_\nu})$ est une famille libre de $G_{\nu-1}$. On la complète en une base $(e_{\nu-1,1}, \ldots, e_{\nu-1,s_{\nu-1}})$ de $G_{\nu-1}$. Ainsi de suite, on obtient une base $(e_{k,1}, \ldots, e_{k,s_k})$ de G_k, avec $s_k \geqslant s_{k+1}$, telle que $e_{k,j} = f(e_{k+1,j})$ pour tout $j = 1, \ldots, s_{k+1}$. Notons que pour tout $1 \leqslant k \leqslant \nu$ et $1 \leqslant j \leqslant s_k$, la famille $\beta_{k,j} = (f^{k-1}(e_{k,j}), \ldots, f(e_{k,j}), e_{k,j})$ est libre, puisque $f^\ell(e_{k,j}) \in G_{k-\ell}$ pour tout $\ell < k$, et les $G_{k-\ell}$ sont en somme directe. La famille

$$\beta = \beta_{\nu,1} \oplus \cdots \oplus \beta_{\nu,s_\nu} \oplus \beta_{\nu-1,s_\nu+1} \oplus \cdots \oplus \beta_{\nu-1,s_{\nu-1}} \oplus \cdots \oplus \beta_{1,s_2+1} \oplus \cdots \oplus \beta_{1,s_1}$$

obtenue par concaténation des familles libres $\beta_{k,j}$ pour $1 \leqslant k \leqslant \nu$ et $s_{k+1} < j \leqslant s_k$ (avec la convention $s_{\nu+1} = 0$), forme exactement la famille des $e_{k,j}$ avec $1 \leqslant k \leqslant \nu$ et $1 \leqslant j \leqslant s_k$. Elle forme donc une base de E. Dans cette base, on obtient

$$[f]_\beta = \begin{pmatrix} \boxed{\mathrm{J}_\nu} & & & & & & (0) \\ & \ddots & & & & & \\ & & \boxed{\mathrm{J}_\nu} & & & & \\ & & & \ddots & & & \\ & & & & \boxed{\mathrm{J}_1} & & \\ & & & & & \ddots & \\ (0) & & & & & & \boxed{\mathrm{J}_1} \end{pmatrix}.$$

Notons que J_1 est la matrice nulle de taille 1 et que le nombre de blocs J_k de taille k est exactement $s_k - s_{k+1}$. □

Remarque 2.3.12. La preuve précédente montre que le nombre de blocs de taille k dans la forme normale d'un endomorphisme nilpotent f est donné par

$$2 \dim \ker f^k - \dim \ker f^{k-1} - \dim \ker f^{k+1}.$$

En effet, on a obtenu que ce nombre est $s_k - s_{k+1}$ où $s_k = \dim G_k$. Or G_k est un supplémentaire de $\ker f^{k-1}$ dans $\ker f^k$, donc $\dim \ker f^k - \dim \ker f^{k-1} = s_k$. Par conséquent le nombre de blocs de taille k est

$$s_k - s_{k+1} = (\dim \ker f^k - \dim \ker f^{k-1}) - (\dim \ker f^{k+1} - \dim \ker f^k).$$

2.3.4 Réduction de Jordan

Théorème 2.3.13 (Forme réduite de Jordan). *Soit $f \in \mathscr{L}(E)$ tel que χ_f est scindé. Alors il existe une base β de E des entiers $q_1, \dots, q_s \in \mathbb{N}^*$ et des scalaires $\alpha_1, \dots, \alpha_s \in \mathrm{K}$ tels que*

$$[f]_\beta = \begin{pmatrix} \boxed{\alpha_1 \mathrm{I}_{q_1} + \mathrm{J}_{q_1}} & & (0) \\ & \ddots & \\ (0) & & \boxed{\alpha_s \mathrm{I}_{q_s} + \mathrm{J}_{q_s}} \end{pmatrix}.$$

Les blocs

$$\mathrm{N}_{\lambda,q} = \lambda \mathrm{I}_q + \mathrm{J}_q = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in \mathrm{M}_q(\mathrm{K})$$

sont appelés *blocs de Jordan.*

Démonstration. Le théorème de Cayley-Hamilton et le lemme des noyaux donnent $E = \bigoplus_{\lambda \in \mathrm{sp}(f)} C_\lambda$. Comme les C_λ sont préservés par f, il suffit de montrer le théorème pour la restriction $f_\lambda = f|_{C_\lambda} \in \mathscr{L}(C_\lambda)$. Puisque $C_\lambda = \ker(f - \lambda \,\mathrm{id})^{m_\lambda}$ où m_λ est la multiplicité algébrique de f, on peut écrire $f_\lambda = \lambda \,\mathrm{id} + h_\lambda$ avec $h_\lambda = f - \lambda \,\mathrm{id} \in \mathscr{N}(C_\lambda)$. (On aurait pu aussi utiliser directement les Propositions 2.3.4 et 2.3.5.) Par le Théorème 2.3.10, on a une base β_λ telle que h_λ est une matrice diagonale par blocs, avec des blocs de la forme J_q avec $q \in \mathbb{N}^*$. Dans cette base, la matrice de $f = \lambda \,\mathrm{id} + h_\lambda$ est alors diagonale par blocs, et chacun de ses blocs est de la forme $\lambda \mathrm{I}_q + \mathrm{J}_q$ avec $q \in \mathbb{N}^*$. Ceci conclut la démonstration. □

Proposition 2.3.14 (Puissances des blocs de Jordan). *Soient $\lambda \in \mathrm{K}$ et $q \in \mathbb{N}^*$. Alors pour tout $m \in \mathbb{N}$, on a*

$$(\lambda \mathrm{I}_q + \mathrm{J}_q)^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{q-1}\lambda^{m-(q-1)} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \binom{m}{2}\lambda^{m-2} \\ \vdots & & \ddots & \ddots & \binom{m}{1}\lambda^{m-1} \\ 0 & \cdots & \cdots & 0 & \lambda^m \end{pmatrix}.$$

Démonstration. On a

$$(\lambda \mathrm{I}_q + \mathrm{J}_q)^m = \sum_{k=0}^{m} \binom{m}{k} \lambda^{m-k} \mathrm{J}_q^k.$$

Or, pour tout $k \in \mathbb{N}$, la matrice J_q^k est la matrice dont les coefficients sont nuls, sauf ceux de la k-ième sur-diagonale qui valent 1. La formule annoncée s'ensuit. □

2.4 Exponentielle de matrices

2.4.1 Définition de l'exponentielle

On dit qu'une suite $(A_k)_{k\in\mathbb{N}}$ de matrices de $\mathrm{M}_n(\mathrm{K})$ converge vers $A = (a_{i,j})$ si pour tous $1 \leqslant i, j \leqslant n$, le coefficient en place (i, j) de A_k converge vers $a_{i,j}$ quand $k \to \infty$. Sur $\mathrm{M}_n(\mathrm{K})$, on définit aussi la norme triple

$$|||A||| = \sup\Big\{ \|Ax\| \ : \ x \in \mathrm{K}^n, \ \|x\| = 1 \Big\},$$

où $\|x\| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$ est la norme euclidienne ou hermitienne canonique sur K^n, selon que $\mathrm{K} = \mathbb{R}$ ou $\mathbb{C}$. Il n'est pas dur de vérifier que $|||\cdot|||$ est une norme sur $\mathrm{M}_n(\mathrm{K})$, qui vérifie

$$|||AB||| \leqslant |||A||| \cdot |||B|||, \quad A, B \in \mathrm{M}_n(\mathrm{K}).$$

On dit que $|||\cdot|||$ est une *norme d'algèbre*. Puisque $\mathrm{M}_n(\mathrm{K})$ est de dimension finie, toutes les normes sont équivalentes, et il existe des constantes $c, C > 0$ telles que

$$c\|A\|_\infty \leqslant |||A||| \leqslant C\|A\|_\infty, \quad A \in \mathrm{M}_n(\mathrm{K}), \tag{2.10}$$

où $\|A\|_\infty = \sup_{1\leqslant i,j\leqslant n} |a_{i,j}|$. En particulier on a $A_k \to A$ dans $\mathrm{M}_n(\mathrm{K})$ si et seulement si $|||A_k - A||| \to 0$.

Théorème–Définition 2.4.1 (Exponentielle de matrice). *Soit* $A \in \mathrm{M}_n(\mathrm{K})$. *Alors la suite* (A_k) *définie par* $A_k = \displaystyle\sum_{\ell=0}^{k} \frac{A^\ell}{\ell!}$ *est convergente et converge vers une matrice que l'on note*

$$e^A = \exp A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathrm{M}_n(\mathrm{K}).$$

On dit que $\exp A$ *est l'exponentielle de la matrice* A.

Démonstration. Si $p < q$ on a

$$|||A_p - A_q||| \leqslant \sum_{\ell=p+1}^{q} \frac{|||A|||^\ell}{\ell!}.$$

Par (2.10), on obtient que pour tout (i, j), la suite $(a_{i,j}^{(k)})_{k\in\mathbb{N}}$ (où $a_{i,j}^{(k)}$ est le coefficient en place (i, j) de A_k) est de Cauchy. Donc la suite $(a_{i,j}^{(k)})_{k\in\mathbb{N}}$ elle converge, ce qui signifie que (A_k) converge. □

Remarquons que pour $A = 0$, on a

$$\exp 0 = \lim_k \sum_{\ell=0}^{k} \frac{0^\ell}{\ell!} = \lim_k \mathrm{I}_n = \mathrm{I}_n. \tag{2.11}$$

2.4.2 Propriétés de l'exponentielle

Proposition 2.4.2 (Exponentielle d'une somme de matrices qui commutent). *Soient* $A, B \in \mathrm{M}_n(\mathrm{K})$ *deux matrices qui commutent. Alors*

$$\exp(A+B) = \exp(A)\exp(B).$$

Démonstration. Puisque A et B commutent on a

$$(A+B)_k = \sum_{\ell=0}^{k} \frac{(A+B)^\ell}{\ell!} = \sum_{\ell=0}^{k}\sum_{m=0}^{\ell} \frac{A^m B^{\ell-m}}{m!(\ell-m)!} = \sum_{\substack{0\leqslant m,\ell\leqslant k\\ m+\ell\leqslant k}} \frac{A^\ell}{\ell!}\frac{B^m}{m!}.$$

Puisque $A_k B_k = \sum_{0\leqslant m,\ell\leqslant k} \frac{A^\ell}{\ell!}\frac{B^m}{m!}$, on obtient

$$|||A_kB_k - (A+B)_k||| = \left|\left|\left| \sum_{\substack{0\leqslant m,\ell\leqslant k\\ m+\ell> k}} \frac{A^\ell}{\ell!}\frac{B^m}{m!} \right|\right|\right| \leqslant \sum_{\substack{0\leqslant m,\ell\leqslant k\\ m+\ell> k}} \frac{|||A|||^\ell}{\ell!}\frac{|||B|||^m}{m!}.$$

Le terme de droite de cette inégalité est égal à

$$\left(\sum_{\ell=0}^{k}\frac{|||A|||^\ell}{\ell!}\right)\left(\sum_{m=0}^{k}\frac{|||B|||^m}{m!}\right) - \sum_{\ell=0}^{k}\frac{(|||A|||+|||B|||)^\ell}{\ell!},$$

qui tend vers $\exp(|||A|||)\exp(|||B|||) - \exp(|||A|||+|||B|||) = 0$ quand $k \to \infty$. Par suite $A_kB_k - (A+B)_k \to 0$, ce qui montre l'égalité voulue. □

Corollaire 2.4.3. *Pour toute matrice* $A \in \mathrm{M}_n(\mathrm{K})$, *la matrice* $\exp A$ *est inversible d'inverse* $\exp(-A)$.

Démonstration. En effet $\exp(A)\exp(-A) = \exp(A-A) = \exp(0) = \mathrm{I}_n$ par (2.11). □

Proposition 2.4.4 (Exponentielle de matrices semblables). *Soit* $A \in \mathrm{M}_n(\mathrm{K})$ *et* $P \in \mathrm{GL}_n(\mathrm{K})$. *Alors*

$$\exp(P^{-1}AP) = P^{-1}\exp(A)P.$$

Démonstration. En effet il suffit de remarquer que

$$\sum_{\ell=0}^{k} \frac{(P^{-1}AP)^\ell}{\ell!} = P^{-1}\left(\sum_{\ell=0}^{k} \frac{A^\ell}{\ell!}\right) P$$

et de passer à la limite $k \to \infty$. En effet, $A_k \to \exp A$ donc

$$\begin{aligned} \left|\!\left|\!\left| P^{-1}A_kP - P^{-1}\exp(A)P \right|\!\right|\!\right| &= \left|\!\left|\!\left| P^{-1}(A_k - \exp A)P \right|\!\right|\!\right| \\ &\leqslant \left|\!\left|\!\left| P^{-1} \right|\!\right|\!\right| |\!|\!| P |\!|\!| |\!|\!| A_k - \exp(A) |\!|\!| \to 0 \end{aligned}$$

quand $k \to \infty$. □

Un corollaire immédiat est le suivant.

Corollaire 2.4.5 (Exponentielle d'une matrice diagonalisable). *Soit $A \in \mathrm{M}_n(\mathrm{K})$ une matrice diagonalisable, qu'on écrit*

$$A = P^{-1}\mathrm{diag}(\alpha_1, \dots, \alpha_n)P$$

avec $P \in \mathrm{GL}_n(\mathrm{K})$. Alors

$$\exp(A) = P^{-1}\mathrm{diag}(e^{\alpha_1}, \dots, e^{\alpha_n})P.$$

Démonstration. Si $D = \mathrm{diag}(\alpha_1, \dots, \alpha_n)$ on a

$$\sum_{\ell=0}^{k} \frac{D^\ell}{\ell!} = \mathrm{diag}\left(\sum_{\ell=0}^{k} \frac{\alpha_1^\ell}{\ell!}, \dots, \sum_{\ell=0}^{k} \frac{\alpha_n^\ell}{\ell!},\right),$$

donc $\exp D = \mathrm{diag}(e^{\alpha_1}, \dots, e^{\alpha_n})$. La Proposition 2.4.4 permet de conclure. □

Proposition 2.4.6 (Exponentielle d'une matrice nilpotente). *Soit $N \in \mathrm{M}_n(\mathrm{K})$ une matrice nilpotente et ν l'indice de nilpotence de N. Alors*

$$\exp(N) = \sum_{\ell=0}^{\nu-1} \frac{N^\ell}{\ell!}.$$

Démonstration. Cela découle immédiatement du fait que $N^\ell = 0$ si $\ell \geqslant \nu$. □

Les trois propriétés précédentes impliquent le résultat suivant.

Proposition 2.4.7. *Soit $A \in \mathrm{M}_n(\mathrm{K})$ telle que χ_A est scindé. On écrit $A = \Delta + N$ la décomposition de Dunford de A, avec Δ diagonalisable et N nilpotente. Soient $\alpha_1, \dots, \alpha_n \in \mathrm{K}$ les valeurs propres de A comptées avec multiplicités. Alors il existe $P \in \mathrm{GL}_n(\mathrm{K})$ telle que*

$$\exp(A) = P^{-1}\mathrm{diag}(e^{\alpha_1}, \dots, e^{\alpha_n})P \sum_{\ell=0}^{n-1} \frac{N^\ell}{\ell!}.$$

Démonstration. On écrit $\Delta = P^{-1}\mathrm{diag}(\alpha_1, \ldots, \alpha_n)P$. Alors les trois propositions précédentes impliquent tour à tour

$$\begin{aligned}\exp A = \exp(\Delta + N) = \exp\Delta\exp N &= P^{-1}\mathrm{diag}(e^{\alpha_1}, \ldots, e^{\alpha_n})P\exp N \\ &= P^{-1}\mathrm{diag}(e^{\alpha_1}, \ldots, e^{\alpha_n})P\sum_{\ell=0}^{n-1}\frac{N^\ell}{\ell!},\end{aligned}$$

où on a utilisé que l'indice de nilpotence de N est inférieur ou égal à n. □

2.4.3 Dérivation dans l'espace des matrices

Définition 2.4.8. On dit qu'une application $F : \mathbb{R} \to \mathrm{M}_n(\mathrm{K})$, $t \mapsto f(t)$, est de classe $\mathscr{C}^1$ si pour tout (i, j) le coefficient $F_{i,j}(t)$ en place (i, j) de $F(t)$ dépend de manière $\mathscr{C}^1$ de t. Si F est de classe $\mathscr{C}^1$ on note $F'(t) = \frac{\mathrm{d}}{\mathrm{d}t}F(t)$ la matrice dont le coefficient en place (i, j) est $F'_{i,j}(t)$.

Proposition 2.4.9. *Soit $A \in \mathrm{M}_n(\mathrm{K})$ et $F(t) = \exp tA$ pour tout $t \in \mathbb{R}$. Alors $F : \mathbb{R} \to \mathrm{M}_n(\mathrm{K})$ est de classe $\mathscr{C}^1$ et*

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA) = A\exp(tA) = \exp(tA)A, \quad t \in \mathbb{R}.$$

Démonstration. Pour tous $t, h \in \mathbb{R}$ on a

$$\exp((t+h)A) - \exp(tA) = \exp(tA)\exp(hA) - \exp(tA) = \exp(tA)(\exp(hA) - \mathrm{I}_n).$$

On écrit

$$\exp(hA) - \mathrm{I}_n = hA + h^2\sum_{k=2}^{\infty}\frac{h^{k-2}A^k}{k!}$$

d'où l'on tire, pour $|h| < 1$,

$$\left|\!\left|\!\left|\frac{\exp(hA) - \mathrm{I}_n}{h} - A\right|\!\right|\!\right| \leqslant |h|\sum_{k=2}^{\infty}\frac{|||A|||^k}{k!} \leqslant |h|\exp|||A|||.$$

Il suit que

$$\left|\!\left|\!\left|\frac{F(t+h) - F(t)}{h} - \exp(tA)A\right|\!\right|\!\right| \leqslant |||\exp(tA)|||\left|\!\left|\!\left|\left(\frac{F(h) - \mathrm{I}_n}{h} - A\right)\right|\!\right|\!\right| \to 0$$

quand $h \to 0$. Par suite F est dérivable de dérivée $F'(t) = \exp(tA)A$, qui est clairement continue en t. Puisque $\exp(tA)$ et A commutent, le résultat est démontré. □

2.5 Applications

Dans ce paragraphe nous présentons deux applications des notions abordées dans ce chapitre : le calcul des suites récurrentes linéaires et la résolution des équations différentielles ordinaires.

2.5.1 Suites récurrentes linéaires

On s'intéresse aux suites définies par une relation de récurrence linéaire d'ordre n à coefficients constants. Soient $a_0, \dots, a_{n-1} \in \mathrm{K}$. On cherche à trouver les suites $(u_k)_{k\in\mathbb{N}}$ satisfaisant la relation

$$a_0 u_k + \cdots + a_{n-1} u_{k+n-1} + u_{k+n} = 0, \quad k \in \mathbb{N}. \tag{R}$$

Le *polynôme caractéristique* associé à l'équation (R) est le polynôme unitaire $P = a_0 + \cdots + a_{n-1}X^{n-1} + X^n$.

Théorème 2.5.1. *On suppose que le polynôme caractéristique de* (R) *est scindé, et on note* $P = \prod_{j=1}^{r}(X-\lambda_j)^{m_j}$ *où* $m_j \in \mathbb{N}^*$ *et les* λ_j *sont deux à deux distincts. Alors* $(u_k)_{k\in\mathbb{N}}$ *satisfait la relation* (R) *si, et seulement si, il existe des polynômes* $Q_1, \dots, Q_r$ *avec* $\deg Q_j \leqslant m_j - 1$ *tels que*

$$u_k = \sum_{j=1}^{r} \lambda_j^k Q_j(k), \quad k \in \mathbb{N}. \tag{2.12}$$

Démonstration. Soit E le K-espace vectoriel des suites $(u_k)_{k\in\mathbb{N}}$ vérifiant (R). Alors on a un isomorphisme $\Psi : E \to \mathrm{K}^n$ donné par $u \mapsto (u_0, \dots, u_{n-1})$, donc $\dim E = n$. Soit $u \in E$. En notant $\mathbf{u}_k = (u_k, \dots, u_{k+n-1})$, on voit que

$$\mathbf{u}_{k+1} = A \cdot \mathbf{u}_k, \quad k \in \mathbb{N},$$

où $A = C(P)$ est la matrice compagnon associée au polynôme caractéristique P de (R). Par récurrence immédiate, on obtient

$$\mathbf{u}_k = A^k \cdot \mathbf{u}_0, \quad k \in \mathbb{N}.$$

Soit $f \in \mathscr{L}(\mathrm{K}^n)$ l'endomorphisme $x \mapsto A \cdot x$ canoniquement associé à A. Alors

$$\chi_f = P = \prod_{j=1}^{r}(X-\lambda_j)^{m_j}.$$

Soit $\beta = \beta_1 \oplus \cdots \oplus \beta_r$ une base adaptée à la décomposition $\mathrm{K}^n = \bigoplus_{j=1}^{r} C_{\lambda_j}$. Alors $[f]_\beta$ est diagonale par blocs, et contient r blocs $A_1, \dots, A_r$ de tailles respectives

$m_1, \dots, m_r$. On peut écrire $f_{\lambda_j} = \lambda_j \operatorname{id}_{C_{\lambda_j}} + h_{\lambda_j}$ avec $h_{\lambda_j} \in \mathscr{N}(C_{\lambda_j})$, si bien que $A_j = \lambda_j \mathrm{I}_{m_j} + N_j$ avec N_j nilpotente d'indice au plus m_j. Par conséquent, si $k \in \mathbb{N}$ on a

$$A_j^k = (\lambda_j \mathrm{I}_{m_j} + N_j)^k = \sum_{\ell=0}^{k} \binom{k}{\ell} \lambda_j^{k-\ell} N_j^\ell = \lambda_j^k \sum_{\ell=0}^{m_j-1} \binom{k}{\ell} \lambda_j^{-\ell} N_j^\ell,$$

où pour tout $\ell \in \mathbb{N}$, $\binom{k}{\ell} = (\ell!)^{-1} k(k-1)\cdots(k-\ell+1)$ dépend de manière polynomiale de k. Par conséquent, il existe des matrices $M_{j,1}, \dots, M_{j,m_j-1} \in \mathrm{M}_n(\mathrm{K})$ telles que

$$A_j^k = \lambda_j^k \sum_{\ell=0}^{m_j-1} M_{j,\ell} k^\ell, \quad k \geqslant 0. \tag{2.13}$$

Écrivons à présent, pour $k \in \mathbb{N}$,

$$\mathbf{u}_k = A^k \cdot \mathbf{u}_0 = \begin{pmatrix} \boxed{A_1^k} & & (0) \\ & \ddots & \\ (0) & & \boxed{A_r^k} \end{pmatrix} \cdot \mathbf{u}_0.$$

Alors (2.13) implique immédiatement que u_k est de la forme (2.12).

Réciproquement, notons que les suites de la forme (2.12) forment un espace vectoriel de dimension $m_1 + \cdots + m_r = n$. Mais on vient de voir qu'il contient E ; cet espace est donc exactement l'espace des solutions de (R), ce qui conclut la démonstration. □

2.5.2 Équations différentielles linéaires à coefficients constants

Dans ce paragraphe on explique comment résoudre les équations différentielles linéaires à coefficients constants. Étant donnés $a_0, \dots, a_{n-1} \in \mathrm{K}$ des constantes, on cherche à résoudre l'équation différentielle d'inconnue $g \in \mathscr{C}^n(\mathbb{R}, \mathrm{K})$

$$\sum_{k=0}^{n-1} a_k g^{(k)}(t) + g^{(n)}(t) = 0, \quad t \in \mathbb{R}, \tag{ED}$$

avec conditions initiales $g^{(k)}(0) = h_k$, pour $k = 0, \dots, n-1$.

Théorème 2.5.2. *On suppose que le polynôme caractéristique de* (R) *est scindé, et on note $P = \prod_{j=1}^r (X - \lambda_j)^{m_j}$ où $m_j \in \mathbb{N}^*$ et les λ_j sont deux à deux distincts.*

Alors les solutions de (ED) *sont exactement les fonctions* $g \in \mathscr{C}^\infty(\mathbb{R}, \mathrm{K})$ *de la forme*

$$\sum_{j=1}^{r} e^{t\lambda_j} Q_j(t)$$

où $Q_j \in \mathrm{K}[X]$ *est un polynôme de degré inférieur ou égal à* $m_j - 1$ *pour tout* $j = 1, \ldots, r$.

Démonstration. Soit $E = \mathscr{C}^\infty(\mathbb{R}, \mathrm{K})$ et $D : E \to E$ l'opérateur de dérivation. Alors $g \in E$ est solution de (ED) si, et seulement si, $g \in \ker P(D)$. Le lemme des noyaux nous donne

$$\ker P(D) = \bigoplus_{j=1}^{r} C_j \quad \text{avec} \quad C_j = \ker(D - \lambda_j)^{m_j}. \tag{2.14}$$

Identifions les espaces C_j. Soit $\Psi_j : E \to E$ défini par $\Psi_j(g)(t) = g(t)e^{-t\lambda_j}$. Alors Ψ_j est un isomorphisme et $D \circ \Psi_j = \Psi_j \circ (D - \lambda_j)$, ce qui donne

$$C_j = \Psi_j^{-1}\big(\ker(D^{m_j})\big). \tag{2.15}$$

Notons que $\ker D^{m_j} \simeq \mathrm{K}_{m_j-1}[X]$. Par (2.14) et (2.15), on obtient que $g \in \ker P(D)$ si et seulement si g est de la forme

$$g(t) = \sum_{j=1}^{r} e^{t\lambda_j} Q_j(t), \quad t \in \mathbb{R},$$

avec $Q_j \in \mathrm{K}_{m_j-1}[X]$. On a donc montré que toutes les solutions de classe $\mathscr{C}^\infty$ de (ED) sont de la forme annoncée, et réciproquement. Pour conclure, il suffit donc de montrer que les solutions de classe $\mathscr{C}^n(\mathbb{R}, \mathrm{K})$ de (ED) sont en fait dans $\mathscr{C}^\infty(\mathbb{R}, \mathrm{K})$. Soit donc $g \in \mathscr{C}^n(\mathbb{R}, \mathrm{K})$ solution de (ED). Alors $g^{(k)}$ est $\mathscr{C}^1$ pour tout $k < n$, donc $g^{(n)} = -a_0 g - \cdots - a_{n-1} g^{(n-1)}$ l'est aussi. Donc g est de classe $\mathscr{C}^{n+1}$. En itérant cet argument on obtient $g \in \mathscr{C}^\infty(\mathbb{R}, \mathrm{K})$, donc toutes les solutions de (ED) sont des éléments de $\ker P(D)$. Ceci conclut la démonstration. □

La solution de (ED) est en fait explicite et donnée par une exponentielle. Nous détaillons cette autre approche, similaire à celle proposée plus haut pour les suites récurrentes.

PROPOSITION 2.5.3. *Soit* g *une solution de* (ED). *On note*

$$\mathbf{g}(t) = (g(t), \ldots, g^{(n-1)}(t)) \in \mathrm{K}^n, \quad t \in \mathbb{R}.$$

Alors $\mathbf{g}(t) = \exp(tA) \cdot \mathbf{g}(0)$ *où* A *est la matrice compagnon associée au polynôme caractéristique de* (ED).

Démonstration. En effet, l'équation différentielle (ED) s'écrit

$$\mathbf{g}'(t) = A \cdot \mathbf{g}(t),$$

où $A = C(P)$ est la matrice compagnon associée au polynôme caractéristique P de (ED). Posons

$$\mathbf{h}(t) = \exp(tA) \cdot \mathbf{g}(0), \quad t \in \mathbb{R}.$$

Alors $\mathbf{h}'(t) = A \cdot \mathbf{h}(t)$ par la Proposition 2.4.9. Ainsi on obtient $\mathbf{f}'(t) = 0$ où $\mathbf{f}(t) = \mathbf{g}(t) - \mathbf{h}(t)$. Puisque $\mathbf{f}(0) = \mathbf{h}(0) = \mathbf{g}(0)$ on obtient $\mathbf{f}(t) = 0$ donc $\mathbf{g}(t) = \mathbf{h}(t)$ pour tout $t \in \mathbb{R}$. Ceci conclut la démonstration. $\square$

Exercice 2.5.4. Retrouver le Théorème 2.5.2 en utilisant le théorème précédent et la décomposition de Dunford, dans l'esprit de la démonstration que nous avons donnée pour le Théorème 2.5.1.

Chapitre 3

Dualité

Sommaire

Dans tout le chapitre, E désigne un K-espace vectoriel qui n'est pas forcément de dimension finie.

3.1 Formes linéaires, espace dual

3.1.1 Définitions et premières propriétés

Définition 3.1.1. Une *forme linéaire* sur un espace vectoriel E est une application $\ell : E \to \mathrm{K}$ qui est linéaire. L'*espace dual* $E^\star$ de E est l'espace des formes linéaires sur E, c'est-à-dire que $E^\star = \mathscr{L}(E, \mathrm{K})$.

Exemple 3.1.2. Voici quelques exemples de formes linéaires.

(i) L'application $\ell : E \to \mathrm{K}$ donnée par $\ell(x) = 0$ pour tout x est une forme linéaire (nulle).
(ii) Si $\lambda \in \mathrm{K}$, l'application d'évaluation $\mathrm{K}[X] \to \mathrm{K}$ donnée par $P \mapsto P(\lambda)$ est une forme linéaire.
(iii) L'application $\mathscr{C}([a,b],\mathrm{K}) \to \mathrm{K}$ donnée par $f \mapsto \int_a^b f$ est une forme linéaire.

Comme K est un K-espace vectoriel de dimension 1, on a

$$\dim E = \dim E^\star$$

si $\dim E < \infty$. On notera $\langle \cdot, \cdot \rangle : E^\star \times E \to \mathrm{K}$ le crochet de dualité, défini par

$$\langle \ell, x \rangle = \ell(x), \quad \ell \in E^\star, \quad x \in E.$$

Notons que pour toute forme linéaire $\ell \in (\mathrm{K}^n)^\star$, il existe un unique n-uplet $(a_1, \ldots, a_n) \in \mathrm{K}$ tel que

$$\ell(x) = a_1 x_1 + \cdots + a_n x_n, \quad x = (x_1, \ldots, x_n) \in \mathrm{K}^n.$$

Plus généralement, on a la proposition suivante.

PROPOSITION 3.1.3. *Supposons E de dimension finie et soit $\mathbf{e} = (e_i)$ une base de E. Alors, pour toute forme linéaire $f \in E^\star$, il existe un unique n-uplet* $(a_1, \ldots, a_n) \in \mathrm{K}^n$ *tel que*

$$\ell(x) = a_1 x_1 + \cdots + a_n x_n, \quad x = x_1 e_1 + \cdots + x_n e_n \in E.$$

Ainsi, de la même manière qu'un choix de base permet d'identifier E à K^n, un choix de base permet aussi d'identifier $E^\star$ à K^n.

Démonstration. Soit $\ell \in E^\star$. Posons $a_j = \ell(e_j)$ pour tout $j = 1, \ldots, n$. Alors par linéarité de ℓ, on a

$$\ell(x_1 e_1 + \cdots + x_n e_n) = a_1 x_1 + \cdots + a_n x_n.$$

Ces coefficients sont uniques : si l'équation ci-dessus est vérifiée pour tout $x \in E$, alors on a nécessairement $a_j = \ell(e_j)$. □

REMARQUE 3.1.4. Le n-uplet $(a_1, \ldots, a_n)$ correspond en fait aux coefficients de la matrice de ℓ dans la base $\mathbf{e}$ et vers la base canonique $\mathbf{1}$ de K :

$${}_{\mathbf{1}}[\ell]_{\mathbf{e}} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}.$$

Le résultat ci-dessus revient en fait à représenter ℓ dans une certaine base $\mathbf{e}_\star$ de $E^\star$ associée à $\mathbf{e}$, comme nous allons le voir au paragraphe suivant.

3.1.2 Base duale

THÉORÈME–DÉFINITION 3.1.5 (Base duale). *Soit E de dimension finie et $\mathbf{e} = (e_j)$ une base de E. Il existe une unique base $\mathbf{e}^\star = (e_j^\star)$ de $E^\star$ telle que*

$$e_k^\star(e_\ell) = \delta_{k,\ell}, \quad 1 \leqslant k, \ell \leqslant n.$$

La base $\mathbf{e}^\star$ est la base duale *de $\mathbf{e}$, tandis que $\mathbf{e}$ est la* base anté-duale *de $\mathbf{e}^\star$.*

Dans le théorème ci-dessus, $\delta_{k,\ell}$ est le symbole de Kronecker, défini par

$$\delta_{k,\ell} = \begin{cases} 1 & \text{si } k = \ell, \\ 0 & \text{sinon.} \end{cases}$$

Démonstration. Si une telle base existe, alors nécessairement on a

$$e_j^\star(x_1 e_1 + \cdots + x_n e_n) = x_j$$

pour tout $x = x_1 e_1 + \cdots + x_n e_n \in E$. Réciproquement, si on définit $e_j^\star$ par la formule ci-dessus, on vérifie aisément que $(e_j^\star)$ est libre, donc est une base de $E^\star$ puis que $\dim E^\star = n$. En effet, si on a $\lambda_1, \ldots, \lambda_n \in \mathrm{K}$ tels que $\lambda_1 e_1^\star + \cdots + \lambda_n e_n^\star = 0$, alors en évaluant en e_j on obtient $\lambda_j = 0$. Ainsi $(e_j^\star)$ est libre, ce qui conclut la démonstration. □

PROPOSITION 3.1.6. *Soit $\mathbf{e} = (e_i)$ une base de E et $\mathbf{e}^\star = (e_j^\star)$ sa base duale. Alors on a*

$$x = \sum_{i=1}^n e_i^\star(x) e_i = \sum_{i=1}^n \langle e_i^\star, x \rangle e_i, \quad x \in E.$$

De même, on a l'égalité

$$\ell = \sum_{j=1}^n \ell(e_j) e_j^\star = \sum_{j=1}^n \langle \ell, e_j \rangle e_j^\star, \quad \ell \in E^\star.$$

Démonstration. Si $x = x_1 e_1 + \cdots x_n e_n \in E$ on a $x_j = e_j^\star(x)$ par définition de $\mathbf{e}^\star$, d'où la première affirmation. En outre, si $\tilde{\ell}$ est la forme linéaire définie par la somme à droite de la dernière égalité ci-dessus, on a, pour tout i,

$$\tilde{\ell}(e_i) = \sum_{j=1}^n \ell(e_j) e_j^\star(e_i) = \ell(e_i)$$

puisque $e_j^\star(e_i) = \delta_{j,i}$. Par suite ℓ et $\tilde{\ell}$ coïncident sur la base (e_i) donc elles coïncident partout. □

Soit $\mathbf{l} = (\ell_1, \ldots, \ell_n)$ une base de $E^\star$. On montre de la même manière que pour le Théorème-Définition 3.1.5 l'existence d'une unique base $\mathbf{e} = (e_1, \ldots, e_n)$ de E, appelée *base anté-duale* de $\mathbf{l}$, qui vérifie $\mathbf{l} = \mathbf{e}^\star$.

3.1.3 Changement de base

On donne la formule de changement pour la matrice de changement de base entre deux bases duales.

Proposition 3.1.7. *Soit E un* K*-espace de dimension finie et* $\mathbf{e}$ *et* $\mathbf{f}$ *deux bases de E. On note* $\mathbf{e}^\star$ *et* $\mathbf{f}^\star$ *les bases duales de* $\mathbf{e}$ *et* $\mathbf{f}$*. Alors on a*

$$P_{\mathbf{e}^\star,\mathbf{f}^\star} = {}^t(P_{\mathbf{e},\mathbf{f}})^{-1}$$

où $P_{\beta,\gamma}$ est la matrice de changement de base de β à γ.

Démonstration. Soit $n = \dim E$. On note $\mathbf{e} = (e_j)$, $\mathbf{f} = (f_j)$, $\mathbf{e}^\star = (e_j^\star)$ et $\mathbf{f}^\star = (f_j^\star)$. On note aussi $A = (a_{i,j}) = P_{\mathbf{f},\mathbf{e}}$ et $B = (b_{i,j}) = P_{\mathbf{e}^\star,\mathbf{f}^\star}$. Alors la Proposition 3.1.6 donne

$$f_j^\star = \sum_{i=1}^{n} f_j^\star(e_i)e_i^\star.$$

Par conséquent $b_{i,j} = f_j^\star(e_i)$. La Proposition 3.1.6 donne aussi

$$e_i = \sum_{j=1}^{n} f_j^\star(e_i)f_j.$$

Par conséquent $a_{j,i} = f_j^\star(e_i)$ et on obtient $B = {}^tA$, ce qu'on voulait démontrer. □

3.1.4 Hyperplans

Définition 3.1.8. Soit E un K-espace vectoriel. Un *hyperplan* de E est par définition le noyau d'une forme linéaire non nulle sur E.

Exemple 3.1.9. Voici quelques exemples d'hyperplans.

(i) Si $\lambda \in \mathrm{K}$, l'ensemble $\{P \in \mathrm{K}[X] \ : \ P(\lambda) = 0\}$ des polynômes qui s'annulent en λ est un hyperplan de $\mathrm{K}[X]$.

(ii) L'ensemble des fonctions continues $f : [a, b] \to \mathrm{K}$ telles que $\int_a^b f = 0$ est un hyperplan de $\mathscr{C}([a, b], \mathrm{K})$.

Proposition 3.1.10. *Soit E un* K*-espace vectoriel (pas nécessairement de dimension finie). Alors les conditions suivantes sont équivalentes :*

(i) *H est un hyperplan de E ;*

(ii) *il existe un sous-espace $F \subset E$ de dimension 1 tel que $E = H \oplus F$.*

Si, de plus, E est de dimension finie n, ces conditions sont équivalentes au fait que $\dim H = n - 1$.

Démonstration. Soit H un hyperplan de E et $\ell \in E^\star \setminus \{0\}$ telle que $H = \ker \ell$. Soit $x \in E$ tel que $\ell(x) \neq 0$. Alors on affirme que $E = H \oplus \mathrm{K}x$. En effet, on a bien sûr $H \cap \mathrm{K}x = \{0\}$ et si $y \in E$, on peut écrire

$$y = z + \lambda x \quad \text{où} \quad \lambda = \ell(y)/\ell(x) \quad \text{et} \quad z = y - \lambda x.$$

Alors $\ell(z) = \ell(y) - \lambda\ell(x) = 0$ donc $z \in H$. On a bien montré que $E = H \oplus \mathrm{K}x$. Réciproquement, supposons $E = H \oplus F$ avec $\dim F = 1$. Soit $\pi : E \to F$ la projection sur F parallèlement à H, et $\psi : F \simeq \mathrm{K}$ un isomorphisme (qui existe toujours car $\dim F = 1$). Notons $\ell = \psi \circ \pi$. Alors ℓ est une forme linéaire, non nulle, puisque $\ell(x) = \psi(x)$ pour tout $x \in F$. Enfin, montrons que $\ker \ell = H$. On a $\ker \ell = \ker(\psi \circ \pi)$. Comme ψ est injective on a $\ker \ell = \ker \pi = H$.

Enfin, si E est de dimension finie, le dernier point est une conséquence immédiate de la proposition 3.1.3. □

Corollaire 3.1.11. *Deux formes linéaires ont même noyau si et seulement si elles sont proportionnelles.*

Démonstration. Supposons que $\ell, \eta \in E^\star$ ont même noyau H. Si $H = E$, alors $\ell = \eta = 0$. Sinon soit $x \notin H$. Alors $E = H \oplus \mathrm{K}x$ par la preuve de la proposition 3.1.10 et $\ell(x), \eta(x) \neq 0$. Soit $y = z + \lambda x$ avec $z \in H$ et $\lambda \in \mathrm{K}$. Alors on a $\ell(y) = \lambda\ell(x) = a\lambda\eta(x) = a\eta(y)$ où $a = \ell(x)/\eta(x)$, donc η et ℓ sont proportionnelles. La réciproque est claire. □

Proposition 3.1.12. *Soit E un* K*-espace vectoriel de dimension finie. Soit $F \subset E$ un sous-espace de E. Alors* $\operatorname{codim} F = r$ *si et seulement si, il existe une famille libre $(\ell_1, \ldots, \ell_r)$ de $E^\star$ telle que*

$$F = \bigcap_{j=1}^{r} \ker \ell_j.$$

Démonstration. Soit F un espace vectoriel de codimension r, i.e. $\dim F = n - r$. On se donne $(e_{r+1}, \ldots, e_n)$ une base de F, que l'on complète en une base $(e_1, \ldots, e_n)$ de E. Alors $x \in F$ si, et seulement si, $e_j^\star(x) = 0$ pour tout $j = 1, \ldots, r$ si, et seulement si $x \in \ker e_j^\star$ pour tout $j = 1, \ldots, r$. En posant $\ell_j = e_j^\star$ pour $j = 1, \ldots, r$, on obtient $F = \bigcap_{j=1}^r \ker \ell_j$. Réciproquement, on se donne $(\ell_1, \ldots, \ell_r)$ une famille libre de $E^\star$ et on pose $F = \bigcap_{j=1}^r \ker \ell_j$. On complète la famille $(\ell_1, \ldots, \ell_r)$ en une base $\mathbf{l} = (\ell_1, \ldots, \ell_n)$ de $E^\star$. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ la base antéduale de $\mathbf{l}$. Alors

$$\bigcap_{j=1}^{r} \ker \ell_j = \operatorname{vect}(e_{r+1}, \ldots, e_n)$$

donc $\dim F = r$. Ceci conclut la démonstration. □

3.1.5 Espace bidual

L'espace bi-dual $E^{\star\star}$ de E est l'espace dual de $E^\star$. Pour tout $x \in E$, on note $\mathrm{ev}_x \in E^{\star\star}$ l'évaluation au point x, définie par

$$\mathrm{ev}_x : E^\star \to \mathrm{K},\ \ell \mapsto \ell(x).$$

La proposition suivante montre qu'on peut identifier x à ev_x.

Proposition 3.1.13. *Supposons que E est de dimension finie. Alors l'endomorphisme* $\mathrm{ev} : E \to E^{\star\star}$ *donné par*

$$\mathrm{ev} : x \mapsto \mathrm{ev}_x$$

est un isomorphisme.

Démonstration. Puisque les espaces ont même dimension, il suffit de montrer que ev est injective. Soit $e_1 \in E \setminus \{0\}$. On complète e_1 en une base $\mathbf{e} = (e_1, e_2, \dots, e_n)$ de E, et on note $\mathbf{e}^\star = (e_1^\star, \dots, e_n^\star)$ la base duale de $\mathbf{e}$. Alors $\mathrm{ev}_{e_1}(e_1^\star) = 1$. En particulier ev_{e_1} est non nulle et on a montré que $e_1 \mapsto \mathrm{ev}_{e_1}$ est injective. Ceci conclut. $\square$

3.2 Orthogonalité (au sens de la dualité)

3.2.1 Définitions et propriétés basiques

Définition 3.2.1. Soit E un K-espace vectoriel et $F \subset E$ un sous-espace. Le *dual orthogonal* de F est le sous-espace $F^\circ \subset E^\star$ donné par

$$F^\circ = \{\ell \in E^\star \ :\ \ell(x) = 0,\ x \in F\} = \{\ell \in E^\star \ :\ F \subset \ker \ell\}.$$

Remarque 3.2.2. Étant donnée $\ell \in E^\star$, le dual orthogonal de son noyau est la droite qu'elle engendre, i.e. $(\ker \ell)^\circ = \mathrm{K}\ell$. En effet, si $\eta \in E^\star$ vérifie $\ker \ell \subset \ker \eta$, alors $\eta = a\ell$ pour un certain $a \in \mathrm{K}$. Réciproquement $a\ell \in (\ker \ell)^\circ$ pour tout $a \in \mathrm{K}$.

Proposition 3.2.3 (Propriétés du dual orthogonal). *Soit E un K-espace vectoriel et $F, G \subset E$ des sous-espaces. Alors les propriétés suivantes sont vérifiées.*

(i) *Si E est de dimension finie alors* $\dim F^\circ + \dim F = E$.

(ii) *Si $F \subset G$ alors $G^\circ \subset F^\circ$.*

(iii) $(F + G)^\circ = F^\circ \cap G^\circ$.

(iv) $F^\circ + G^\circ \subset (F \cap G)^\circ$, *et on a l'inclusion réciproque si* $\dim E < \infty$.

(v) *Sous l'isomorphisme de la proposition 3.1.13, on a* $F \simeq (F^\circ)^\circ$.

Démonstration. On admet dans un premier temps le premier point. Pour montrer le point (ii) on remarque simplement que si $\ell \in G^\circ$, alors $\ell(x) = 0$ pour tout $x \in G$, donc en particulier $\ell(x) = 0$ pour tout $x \in F \subset G$, donc $\ell \in F^\circ$.

Pour le point (iii), on remarque que si $\ell \in F^\circ \cap G^\circ$, alors $F \subset \ker \ell$ et $G \subset \ker \ell$ donc $F + G \subset \ker \ell$ puisque $\ker \ell$ est un espace vectoriel, donc $\ell \in (F + G)^\circ$. L'inclusion réciproque est claire.

Montrons (iv). Soit $\ell = f + g \in F^\circ + G^\circ$. Si $x \in F \cap G$ on a $f(x) = 0 = g(x)$, donc $\ell(x) = 0$. Ceci implique que $\ell \in (F \cap G)^\circ$. Si E est de dimension finie, le point (i) donne $\dim(F \cap G)^\circ = n - \dim(F \cap G)$. En outre, les points (i) et (iii) impliquent

$$\begin{aligned}
\dim(F^\circ + G^\circ) &= \dim F^\circ + \dim G^\circ - \dim(F^\circ \cap G^\circ) \\
&= n - \dim F + n - \dim G - \dim(F + G)^\circ \\
&= n - (\dim F + \dim G - \dim(F + G)) \\
&= n - \dim(F \cap G),
\end{aligned}$$

donc $\dim(F^\circ + G^\circ) = \dim(F \cap G)^\circ$, donc ces deux espaces sont égaux puisqu'on a l'inclusion $F^\circ + G^\circ \subset (F \cap G)^\circ$.

Montrons à présent (v). On a

$$(F^\circ)^\circ = \{\eta \in E^{\star\star} \ : \ \eta(\ell) = 0, \ \ell \in F^\circ\} \simeq \{x \in E \ : \ \mathrm{ev}_x(\ell) = 0, \ \ell \in F^\circ\}.$$

Notons $\widetilde{F} = \{x \in E \ : \ \mathrm{ev}_x(\ell) = 0, \ \ell \in F^\circ\}$. Alors $F \subset \widetilde{F}$. En outre $\dim \widetilde{F} = \dim(F^\circ)^\circ = \dim F$ donc $F = \widetilde{F}$. On en déduit le point (v).

Il reste à montrer (i). Soit $(e_1, \ldots, e_r)$ une base de F, que l'on complète en une base $(e_1, \ldots, e_n)$ de E. Alors on affirme que

$$F^\circ = \mathrm{vect}(e^\star_{r+1}, \ldots, e^\star_n).$$

En effet, notons $\widetilde{F} \subset E^\star$ le terme de droite de l'égalité ci-dessus. On a bien sûr $\widetilde{F} \subset F^\circ$, puisque si $j > r$ on a $e^\star_j(e_i) = 0$ pour tout $i = 1, \ldots, r$, donc $e^\star_j \in F^\circ$. Réciproquement, soit $\ell \in F^\circ$. Écrivons

$$\ell = \sum_{j=1}^n a_j e^\star_j$$

avec $a_j = \ell(e_j)$. Puisque $\ell \in F^\circ$ on a $0 = \ell(e_j) = a_j$ pour tout $j = 1, \ldots, r$, donc $\ell \in \mathrm{vect}(e^\star_{r+1}, \ldots, e^\star_n) = \widetilde{F}$. On a obtenu $\dim F^\circ = \dim \widetilde{F} = n - r = n - \dim F$, ce qui conclut la démonstration. □

3.2.2 Application

Nous donnons ici une application élégante de la notion d'orthogonalité.

Proposition 3.2.4. *Soit X un ensemble et $f_1, \ldots, f_n : X \to \mathrm{K}$ des fonctions qui forment une famille libre de K^X (les applications de X vers K). Alors il existe des éléments $x_1, \ldots, x_n \in E$ tels que la matrice*

$$\begin{pmatrix} f_1(x_1) & \cdots & f_1(x_n) \\ \vdots & & \vdots \\ f_n(x_1) & \cdots & f_n(x_n) \end{pmatrix}$$

est inversible.

Démonstration. Soit $E = \mathrm{K}^X$ l'espace des applications de X vers K. Soit $F = \mathrm{vect}(f_1, \ldots, f_n)$. Pour tout $x \in X$, on note $\delta_x : F \to \mathrm{K}$ l'évaluation en x qui est la forme linéaire donnée par $F \ni f \mapsto f(x) \in \mathrm{K}$. On considère alors

$$G = \mathrm{vect}\big\{\delta_x \ : \ x \in X\big\} \subset F^\star$$

l'espace engendré toutes les formes δ_x, quand x parcourt X. Soit $f \in F$. Notons que si $\ell(f) = 0$ pour tout $\ell \in G$, alors $f(x) = 0$ pour tout $x \in X$, donc $f = 0$. Ceci montre que $\{f \in F \ : \ \mathrm{ev}_f(\ell) = 0, \ \ell \in G\} = \{0\}$. Mais cet espace est naturellement identifié à G° par l'isomorphisme de la proposition 3.1.13. En particulier $G^\circ = \{0\}$ donc $G = F^\star$. Dès lors, il existe $x_1, \ldots, x_n \in X$ tels que $\delta = (\delta_{x_1}, \ldots, \delta_{x_n})$ forme une base de $F^\star$. Pour tout $x \in X$ on peut écrire

$$\delta_x = \sum_{i=1}^n \delta_x(f_i) f_i^\star = \sum_{i=1}^n f_i(x) f_i^\star$$

où $\mathbf{f}^\star = (f_1^\star, \ldots, f_n^\star)$ est la base duale de $\mathbf{f} = (f_1, \ldots, f_n)$. En particulier, la matrice $(f_i(x_j))$ est la matrice des vecteurs de δ dans la base $\mathbf{f}^\star$, donc est inversible. □

3.3 Transposition

3.3.1 Définitions et premières propriétés

Définition 3.3.1. Soient E, F deux K-espaces vectoriels et $f \in \mathscr{L}(E, F)$. On définit l'endomorphisme transposé $f^\top : F^\star \to E^\star$ par

$$f^\top(\ell)(x) = \ell(f(x)), \quad \ell \in F^\star, \quad x \in E.$$

Proposition 3.3.2. *Soient E, F, G des K-espaces vectoriels, $\lambda \in \mathrm{K}$, $f, g \in \mathscr{L}(E, F)$, $h \in \mathscr{L}(F, G)$. Alors :*

(i) $(f+g)^\top = f^\top + g^\top$;

(ii) $(\lambda f)^\top = \lambda f^\top$;

(iii) $(h \circ f)^\top = f^\top \circ h^\top$;

(iv) *sous l'identification* $E \simeq E^\star$ *de la proposition* 3.1.13, *on a* $(f^\top)^\top = f$;

(v) $(\mathrm{id}_E)^\top = \mathrm{id}_{E^\star}$.

Démonstration. Les trois premiers points sont immédiats. Pour (iv), on vérifie que pour $\ell \in G^\star$ et $x \in E^\star$ on a

$$\left\langle (h\circ f)^\top \ell,\ x\right\rangle = \langle \ell, (h\circ f)(x)\rangle = \langle \ell, h(f(x))\rangle = \langle h^\top(\ell), f(x)\rangle = \left\langle f^\top(h^\top(\ell)), x\right\rangle.$$

Ceci montre bien l'égalité annoncée. □

Exemple 3.3.3. Soient $a < b$ des réels et $\ell : \mathbb{R}[X] \to \mathbb{R}$ définie par

$$\ell(P) = \int_a^b P(t)\mathrm{d}t, \quad P \in \mathbb{R}[X].$$

Soit $D \in \mathscr{L}(\mathbb{R}[X])$ défini par $D(P) = P'$. Alors

$$D^\top \ell = \delta_b - \delta_a \quad \text{où} \quad \delta_c(P) = P(c), \quad c \in \mathbb{R}, \quad P \in \mathbb{R}[X].$$

En effet pour tout $P \in \mathbb{R}[X]$ on a

$$\langle D^\top \ell, P\rangle = \langle \ell, D(P)\rangle = \langle \ell, P'\rangle = \int_a^b P' = P(b) - P(a) = \langle \delta_b - \delta_a, P\rangle.$$

3.3.2 Lien avec la transposition matricielle

La proposition suivante fait le lien entre la transposée d'une matrice et la transposée d'une application linéaire qu'elle représente.

Proposition 3.3.4. *Soit E un* K*-espace de dimension finie et β une base de E. Alors pour tout endomorphisme $f \in \mathscr{L}(E)$, on a*

$$\left[f^\top\right]_{\beta^\star} = {}^t[f]_\beta.$$

Démonstration. On note $\beta = (e_1, \ldots, e_n)$ et $\beta^\star = (e_1^\star, \ldots, e_n^\star)$. On note $A = (a_{i,j})$ la matrice $[f]_\beta$, et $B = (b_{i,j})$ la matrice $\left[f^\top\right]_{\beta^\star}$. Alors on a

$$f(e_j) = \sum_{i=1}^n a_{i,j} e_i \quad \text{et} \quad f^\top(e_j^\star) = \sum_{i=1}^n b_{i,j} e_i^\star, \quad j = 1, \ldots, n.$$

Par la Proposition 3.1.6 on a

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

ce qui montre bien que $B = {}^tA$. □

En conséquence du résultat précédent, on obtient la

Proposition 3.3.5. *Soit E un* K*-espace de dimension finie et $f \in \mathscr{L}(E)$. Alors*

(i) $\operatorname{rang} f = \operatorname{rang} f^\top$;

(ii) $\chi_f = \chi_{f^\top}$;

(iii) $\operatorname{tr} f = \operatorname{tr} f^\top$;

(iv) $\det f = \det f^\top$;

(v) $\dim\ker(f - \lambda \mathrm{id}_E) = \dim\ker(f^\top - \lambda \mathrm{id}_{E^\star})$ *pour tout* $\lambda \in \mathrm{K}$.

Chapitre 4

Algèbre bilinéaire

Sommaire

Dans tout le chapitre, E désigne un K-espace vectoriel qui n'est pas forcément de dimension finie.

4.1 Formes bilinéaires

4.1.1 Définitions

Définition 4.1.1. Soit E un K-espace vectoriel. Une *forme bilinéaire* sur E est une application $\varphi : E \times E \to \mathrm{K}$ qui est linéaire en chacune de ses variables, c'est-à-dire que pour tout $x \in E$, les applications

$$y \mapsto \varphi(x, y) \quad \text{et} \quad y \mapsto \varphi(y, x)$$

sont des formes linéaires. Une forme bilinéaire φ sur E est dite *symétrique* si

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

On notera $\mathscr{B}(E)$ l'espace des formes bilinéaires sur E.

Définition 4.1.2. Soit E un K-espace vectoriel. Une forme bilinéaire symétrique $\varphi : E \times E \to \mathrm{K}$ est dite *positive*, ce qu'on écrit $\varphi \geqslant 0$, si

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

Une forme bilinéaire symétrique φ est dite *définie positive* si $\varphi \geqslant 0$ et si pour tout $x \in E$,

$$\varphi(x, x) = 0 \quad \Longrightarrow \quad x = 0.$$

Une forme bilinéaire symétrique définie positive est un *produit scalaire* sur E. Un K-espace vectoriel E muni d'un produit scalaire $\varphi = (\cdot,\ \cdot)$ est appelé *espace préhilbertien*. Si de plus E est de dimension finie, alors E est un *espace euclidien*.

On notera $\mathscr{S}(E)$ l'espace des formes bilinéaires symétriques sur E et $\mathscr{S}_+(E)$ (resp. $\mathscr{S}_{++}(E)$) l'espace des formes bilinéaires symétriques positives (resp. définies positives) sur E.

Remarque 4.1.3. Si $\mathrm{K} = \mathbb{C}$, alors $\mathscr{S}_+(E) = \{0\}$ et $\mathscr{S}_{++}(E) = \emptyset$.

4.1.2 Matrice d'une forme bilinéaire

Supposons dans ce paragraphe que E de dimension finie n. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ est une base de E. À toute forme bilinéaire $\varphi : E \times E \to \mathrm{K}$, on associe la matrice $[\varphi]_{\mathbf{e}} \in \mathrm{M}_n(\mathrm{K})$ de φ dans la base $\mathbf{e}$ la matrice dont le coefficient en place (i, j) est

$$\varphi(e_i, e_j).$$

Pour $x, y \in E$, la bilinéarité de φ implique

$$\varphi(x, y) = {}^t[x]_{\mathbf{e}} \ \cdot [\varphi]_{\mathbf{e}} \cdot [y]_{\mathbf{e}}. \tag{4.1}$$

PROPOSITION 4.1.4. *Si* $\mathbf{f}$ *est une autre base de* E*, on a la formule de changement de base*

$$[\varphi]_{\mathbf{f}} = {}^tP \cdot [\varphi]_{\mathbf{e}} \cdot P \quad \textit{où} \quad P = P_{\mathbf{e},\mathbf{f}}.$$

Démonstration. La matrice de changement de base $P = (p_{i,j}) = P_{\mathbf{e},\mathbf{f}}$ de $\mathbf{e}$ vers $\mathbf{f}$ est définie par les relations

$$f_j = \sum_{i=1}^{n} p_{i,j} e_i, \quad j = 1, \ldots, n.$$

Le coefficient en place (i, j) de $[\varphi]_{\mathbf{f}}$ est donné par

$$\varphi(f_i, f_j) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} p_{k,i} p_{\ell,j} \varphi(e_k, e_\ell),$$

qui est exactement le coefficient en place (i, j) de ${}^tP[\varphi]_{\mathbf{e}}P$. □

Observons aussi que φ est symétrique si sa matrice dans n'importe quelle base est symétrique. Le sens direct est évident, et si $[\varphi]_{\mathbf{e}}$ est symétrique pour une certaine base $\mathbf{e}$, alors l'identité (4.1) montre que φ est symétrique.

La Proposition 4.1.4 implique que $\operatorname{rang}[\varphi]_{\mathbf{e}} = \operatorname{rang}[\varphi]_{\mathbf{f}}$ pour toutes bases $\mathbf{e}, \mathbf{f}$ de E. Ceci suggère la définition suivante.

DÉFINITION 4.1.5. Le *rang* d'une forme bilinéaire $\varphi : E \times E \to \mathrm{K}$ est définie par

$$\operatorname{rang} \varphi = \operatorname{rang}[\varphi]_{\mathbf{e}}$$

pour n'importe quelle base $\mathbf{e}$.

4.1.3 Formes quadratiques

DÉFINITION 4.1.6. Une *forme quadratique* sur E est une application $q : E \to \mathrm{K}$ telle que

$$q(x) = \varphi(x, x), \quad x \in E,$$

pour une forme bilinéaire symétrique φ.

On notera $\mathscr{Q}(E)$ l'espace des formes quadratiques sur E.

REMARQUE 4.1.7. On peut supprimer le terme “symétrique” dans la définition ci-dessus. En effet, si $q(x) = \varphi(x, x)$ avec φ une forme bilinéaire, on a $q(x) = \widetilde{\varphi}(x, x)$ où

$$\widetilde{\varphi}(x, y) = \frac{1}{2} \left(\varphi(x, y) + \varphi(y, x) \right), \quad x, y \in E,$$

et la forme $\widetilde{\varphi}$ ainsi définie est symétrique.

Proposition 4.1.8. *Soit φ une forme bilinéaire symétrique sur E et q la forme quadratique associée. Alors pour tous $x, y \in E$ et $\lambda \in \mathrm{K}$ on a*

(i) $q(\lambda x) = \lambda^2 q(x)$;

(ii) $2\varphi(x,y) = q(x+y) - q(x) - q(y)$;

(iii) $q(x+y) + q(x-y) = 2(q(x)+q(y))$.

Corollaire 4.1.9 (Formules de polarisation). *Soit φ une forme bilinéaire symétrique sur E et q la forme quadratique associée. Pour tous $x, y \in E$ on a*

$$\varphi(x,y) = \frac{q(x+y) - q(x) - q(y)}{2} = \frac{q(x+y) - q(x-y)}{4}.$$

En particulier, si q est une forme quadratique, il existe une unique forme bilinéaire symétrique φ telle que $\varphi(x,x) = q(x)$ pour tout $x \in E$, que l'on appelle forme polaire *de q.*

Si q est une forme quadratique, on peut donc définir $\operatorname{rang} q = \operatorname{rang} \varphi$ où φ est la forme polaire de q.

Démonstration de la Proposition 4.1.8. On a $q(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda^2 \varphi(x,x) = \lambda^2 q(x)$ par bilinéarité, d'où le point (i). Pour le second point, on remarque que $q(x+y) = \varphi(x+y, x+y) = \varphi(x,x) + \varphi(y,y) + 2\varphi(x,y)$ par symétrie de φ, d'où l'on déduit (ii). Enfin pour le troisième point on remarque que le second donne $2\varphi(x,-y) = q(x-y) - q(x) - q(y)$. En remarquant que $\varphi(x,-y) = -\varphi(x,y)$ et en combinant ce qui précède avec l'égalité (ii), on obtient le résultat voulu. ☐

Exercice 4.1.10. Montrer que $q : E \to \mathrm{K}$ est une forme quadratique si, et seulement si, pour tous $\lambda \in \mathrm{K}$ et $x, y \in E$ on a

$$q(\lambda x) = \lambda^2 q(x) \quad \text{et} \quad q(x+y) + q(x-y) = 2(q(x)+q(y)).$$

Si E est de dimension finie, $\mathbf{e} = (e_1, \ldots, e_n)$ est une base de E, et q une forme quadratique, on note $[q]_{\mathbf{e}} = [\varphi]_{\mathbf{e}}$ où φ est la forme polaire de q. Alors la formule (4.1) donne

$$q(x) = {}^t[x]_{\mathbf{e}} \cdot [q]_{\mathbf{e}} \cdot [x]_{\mathbf{e}} = \sum_{1 \leqslant i,j \leqslant n} a_{i,j} x_i x_j, \quad x = x_1 e_1 + \cdots + x_n e_n \in E,$$

où on a noté $[q]_{\mathbf{e}} = (a_{i,j})$.

4.1.4 Représentation des formes bilinéaires

Le théorème suivant nous dit qu'une forme bilinéaire n'est rien autre qu'une application linéaire $E \to E^\star$.

Théorème 4.1.11. *Soit E un* K*-espace vectoriel. Alors l'application*

$$\Phi : \mathscr{B}(E) \to \mathscr{L}(E, E^\star), \quad \varphi \mapsto \iota_\varphi$$

est un isomorphisme.

Démonstration. L'application Φ est injective : si $\iota_\varphi = 0$, alors $\varphi(v, \cdot) = 0$ pour tout $v \in E$, donc $\varphi(v, w) = 0$ pour tous $v, w \in E$, c'est-à-dire $\varphi = 0$. Montrons que Φ est surjective. Soit $f \in \mathscr{L}(E, E^\star)$. On pose

$$\varphi(x, y) = \langle f(x), y \rangle, \quad x, y \in E.$$

Alors φ est bilinéaire et pour tout $x \in E$ on a $\Phi(\varphi)(x) = \iota_\varphi(x) = \varphi(x, \cdot) = f(x)$ par définition de φ, donc $\Phi(\varphi) = f$. □

4.1.5 Formes bilinéaires non dégénérées

Définition 4.1.12. Soit E un K-espace vectoriel et φ une forme bilinéaire sur E. Alors on définit

$$G_\varphi = \{x \in E \ : \ \varphi(x, y) = 0, \ y \in E\} \quad \text{et} \quad D_\varphi = \{x \in E \ : \ \varphi(y, x) = 0, \ y \in E\}.$$

les noyaux à gauche et à droite de φ. Si de plus φ est symétrique, alors $G_\varphi = D_\varphi$ et on pose

$$\ker \varphi = \ker q = G_\varphi = D_\varphi$$

où q est la forme quadratique associée à φ.

Remarque 4.1.13. Si $\operatorname{iso} q = \{x \in E \ : \ q(x) = 0\}$ est le cône des vecteurs isotropes de q, alors on a $\ker q \subset \operatorname{iso} q$; l'inclusion inverse est fausse en général. En effet, si $E = \mathbb{R}^2$ et $q(x) = x_1^2 - x_2^2$ pour $x = (x_1, x_2) \in E$, on a

$$\ker q = \{0\} \quad \text{et} \quad \operatorname{iso} q = \{(x_1, x_2) \ : \ x_1^2 = x_2^2\}.$$

Proposition 4.1.14. *Si E est de dimension finie et φ est une forme bilinéaire sur E alors*

$$\dim G_\varphi = \dim D_\varphi = \dim E - \operatorname{rang} \varphi.$$

Démonstration. Soit $\mathbf{e}$ une base de E. Soit $x \in E$. Alors $x \in G_\varphi$ si et seulement si pour tout $y \in E$, ce qui équivaut à dire

$$^t[x]_\mathbf{f}[\varphi]_\mathbf{e}[y]_\mathbf{e} = 0, \quad y \in E.$$

Par conséquent, $x \in G_\varphi$ si et seulement si

$$^t[x]_\mathbf{e}[\varphi]_\mathbf{e} = 0,$$

c'est-à-dire si et seulement si $[x]_\mathbf{e}$ appartient au noyau de la matrice $^t[\varphi]_\mathbf{e}$. On en déduit que

$$\dim G_\varphi = \dim \ker {}^t[\varphi]_\mathbf{e} = n - \operatorname{rang} {}^t[\varphi]_\mathbf{e} = n - \operatorname{rang} \varphi.$$

De même, on montre que $\dim D_\varphi = n - \operatorname{rang} \varphi$. □

Définition 4.1.15. Soit E un K-espace vectoriel de dimension finie et φ une forme bilinéaire sur E. On dit que φ est *non dégénérée* si $G_\varphi = D_\varphi = \{0\}$.

Théorème 4.1.16 (Théorème de représentation de Riesz généralisé). *Soit E un* K-*espace vectoriel de dimension finie n et φ une forme bilinéaire non dégénérée sur E. Alors l'application*

$$\iota_\varphi : E \to E^\star, \quad v \mapsto \varphi(v, \cdot),$$

est un isomorphisme. Autrement dit, pour tout $\ell \in E^\star$, il existe un unique vecteur $\ell^{\sharp_\varphi} \in E$ tel que

$$\ell(x) = \varphi(\ell^{\sharp_\varphi}, x), \quad x \in E.$$

Démonstration. L'application $\iota_\varphi : E \to E^\star$ est linéaire. De plus on a

$$\ker \iota_\varphi = D_\varphi = G_\varphi = \{0\}$$

par hypothèse. Par suite ι_φ est injective ; comme $\dim E = n = \dim E^\star$, on en déduit que ι_φ est un isomorphisme. En notant $\ell^{\sharp_\varphi} = (\iota_\varphi)^{-1}(\ell)$ pour $\ell \in E^\star$ on a

$$\varphi(\ell^{\sharp_\varphi}, x) = \ell(x), \quad x \in E,$$

par définition de ι_φ. Par bijectivité de φ, le vecteur $\ell^{\sharp_\varphi}$ est unique. □

4.1.6 Adjoint d'un endomorphisme

Théorème–Définition 4.1.17 (Endomorphisme φ-adjoint). *Soit E un* K*-espace de dimension finie et $\varphi \in \mathscr{B}(E)$ non dégénérée. Alors pour tout $f \in \mathscr{L}(E)$, il existe un unique $f^{\star,\varphi} \in \mathscr{L}(E)$ tel que le diagramme*

$$\begin{array}{ccc} E & \xrightarrow{\iota_\varphi} & E^\star \\ {\scriptstyle f^{\star,\varphi}}\big\downarrow & & \big\downarrow{\scriptstyle f^\top} \\ E & \xrightarrow[\iota_\varphi]{} & E^\star \end{array}$$

commute, i.e. $\iota_\varphi \circ f^{\star,\varphi} = f^\top \circ \iota_\varphi$. Ceci s'écrit encore

$$\varphi(x, f(y)) = \varphi\big(f^{\star,\varphi}(x), y\big), \quad x, y \in E.$$

L'endomorphisme $f^{\star,\varphi}$ est appelé endomorphisme adjoint *de f pour φ.*

Démonstration. Si $f^{\star,\varphi}$ existe, alors nécessairement $f^{\star,\varphi} = \iota_\varphi^{-1} \circ f^\top \circ \iota_\varphi$. Réciproquement, l'endomorphisme $\iota_\varphi^{-1} \circ f^\top \circ \iota_\varphi$ vérifie les conditions demandées. On vérifie qu'on a bien, pour $x, y \in E$,

$$\varphi(x, f(y)) = \langle \iota_\varphi(x), f(y)\rangle = \langle (f^\top \circ \iota_\varphi)(x), y\rangle = \varphi\big(\big[\iota_\varphi^{-1} \circ f^\top \circ \iota_\varphi\big](x), y\big),$$

et le dernier terme est exactement $\varphi\big(f^{\star,\varphi}(x), y\big)$. □

Proposition 4.1.18. *Soit E un* K*-espace de dimension finie et* $\mathbf{e}$ *une base de E. Soit $\varphi \in \mathscr{B}(E)$ non dégénérée. Alors*

$$[f]_{\mathbf{e}} = [\varphi]_{\mathbf{e}}^{-1} \cdot {}^t[f^{\star,\varphi}]_{\mathbf{e}} \cdot \ [\varphi]_{\mathbf{e}}.$$

Démonstration. Par définition, on a $f^{\star,\varphi} = \iota_\varphi^{-1} \circ f^\top \circ \iota_\varphi$. Soit $\mathbf{e}^\star = (e_1^\star, \ldots, e_n^\star)$ la base duale de $\mathbf{e}$. Alors

$$[f^{\star,\varphi}]_{\mathbf{e}} = [\iota_\varphi]_{\mathbf{e}^\star,\mathbf{e}}^{-1} \left[f^\top\right]_{\mathbf{e}^\star} [\iota_\varphi]_{\mathbf{e}^\star,\mathbf{e}}$$

où $[\iota_\varphi]_{\mathbf{e}^\star,\mathbf{e}}$ est la matrice de $\iota_\varphi : E \to E^\star$ dans la base $\mathbf{e}$ vers la base $\mathbf{e}^\star$. Remarquons que

$$\iota_\varphi(e_j) = \sum_{i=1}^n \langle \iota_\varphi(e_j), e_i\rangle e_i^\star = \sum_{i=1}^n \varphi(e_j, e_i) e_i^\star.$$

Par suite $[\iota_\varphi]_{\mathbf{e}^\star,\mathbf{e}} = {}^t[\varphi]_{\mathbf{e}}$ et on obtient

$$[f^{\star,\varphi}]_{\mathbf{e}} = {}^t[\varphi]_{\mathbf{e}}^{-1} \cdot {}^t[f^\top]_{\mathbf{e}} \cdot \ {}^t[\varphi]_{\mathbf{e}}.$$

Puisque $\left[f^\top\right]_{\mathbf{e}^\star} = {}^t[f]_{\mathbf{e}}$, on obtient l'égalité désirée. □

Proposition 4.1.19 (Propriétés de l'adjoint). *Soit E un* K*-espace vectoriel de dimension finie et $\varphi \in \mathscr{B}(E)$ une forme bilinéaire non dégénérée. Alors pour tous $f, g \in \mathscr{L}(E)$ et $\lambda \in \mathrm{K}$, on a*

(i) $(f+g)^{\star,\varphi} = f^{\star,\varphi} + g^{\star,\varphi}$;

(ii) $(\lambda f)^{\star,\varphi} = \lambda f^{\star,\varphi}$;

(iii) $(f \circ g)^{\star,\varphi} = g^{\star,\varphi} \circ f^{\star,\varphi}$;

(iv) $(f^{\star,\varphi})^{\star,\varphi} = f$;

(v) $\mathrm{id}_E^{\star,\varphi} = \mathrm{id}_E$.

Démonstration. Soient $\lambda \in \mathrm{K}$ et $f, g \in \mathscr{L}(E)$. Soient $x, y \in E$. On a

$$\begin{aligned}\varphi((\lambda f + g)^{\star,\varphi}(x), y) &= \varphi(x, (\lambda f + g)(y)) \\ &= \varphi(x, \lambda f(y)) + \varphi(x, g(y)) \\ &= \varphi(\lambda f^{\star,\varphi}(x), y) + \varphi(g^{\star,\varphi}(x), y) \\ &= \varphi((\lambda f^{\star,\varphi} + g^{\star,\varphi})(x), y).\end{aligned}$$

Cette égalité étant vraie pour tout $y \in E$, on en déduit que

$$(\lambda f + g)^{\star,\varphi}(x) = \lambda f^{\star,\varphi}(x) + g^{\star,\varphi}(x)$$

pour tout $x \in E$, d'où les propriétés (i) et (ii). Pour (iii), on remarque que pour tous $x, y \in E$ on a

$$\varphi((f \circ g)^{\star,\varphi}(x), y) = \varphi(x, (f \circ g)(y)) = \varphi(f^{\star,\varphi}(x), g(y)) = \varphi((g^{\star,\varphi} \circ f^{\star,\varphi})(x), y).$$

Puisque φ est non dégénérée et que l'égalité ci-dessus est vraie pour tous x, y, on obtient $(f \circ g)^{\star,\varphi}(x) = g^{\star,\varphi} \circ f^{\star,\varphi}(x)$ pour tout x, d'où l'égalité souhaitée. La propriété (iv) s'obtient en appliquant deux fois la définition de l'adjoint, et (v) est immédiate. □

4.2 Loi d'inertie de Sylvester

4.2.1 Orthogonalité

Définition 4.2.1. Soit E un K-espace vectoriel et $\varphi : E \times E \to \mathrm{K}$ une forme bilinéaire. Une famille $(u_1, \ldots, u_k)$ est dite *orthogonale pour φ* si

$$\varphi(u_i, u_j) = 0, \quad 1 \leqslant i \neq j \leqslant k.$$

Proposition 4.2.2. *Soit E un* K*-espace vectoriel de dimension finie n et φ une forme bilinéaire symétrique sur E. Alors il existe une base $(e_1, \ldots, e_n)$ de E qui est orthogonale pour φ, et telle que $\varphi(e_k, e_k) = 0$ si et seulement si $k > r$, où on a noté $r = \operatorname{rang} \varphi$.*

Démonstration. On procède par récurrence sur la dimension. Si $n = 1$, c'est automatique. Supposons la propriété vraie en dimension $n-1 \geqslant 2$, et prenons φ une forme bilinéaire symétrique sur un espace E de dimension n. On considère $q : E \to \mathrm{K}$ la forme quadratique associée. Si celle-ci est nulle, alors φ l'est aussi par le Corollaire 4.1.9. En particulier $r = 0$ et n'importe quelle base est orthogonale pour φ. Sinon, il existe $e_1 \in E$ tel que $q(e_1) \neq 0$. On considère $\ell : E \to \mathrm{K}$ la forme linéaire définie par

$$\ell(x) = \varphi(e_1, x), \quad x \in E.$$

(Autrement dit $\ell = \iota_\varphi(e_1)$.) Alors ℓ est non nulle, puisque $\ell(e_1) \neq 0$; par conséquent $G = \ker \ell$ est de dimension $n-1$. Comme $q(e_1) \neq 0$ on a $G \cap \mathrm{K}e_1 = \{0\}$ d'où l'on tire

$$E = G \oplus \mathrm{K}e_1.$$

La restriction $\varphi|_{G\times G}$ est une forme bilinéaire symétrique sur G. En appliquant l'hypothèse de récurrence, il existe $(e_2, \dots, e_n)$ une base de G telle que $\varphi(e_i, e_j) = 0$ si $2 \leqslant i \neq j \leqslant n$, et $\varphi(e_j, e_j) \neq 0$ ssi $j \leqslant r$ pour un certain $r \geqslant 1$. En outre $\varphi(e_1, e_j) = 0$ si $j > 1$ puisque dans ce cas on a $e_j \in G = \ker \ell$. Ainsi la base $\mathbf{e} = (e_1, \dots, e_n)$ est orthogonale pour φ. Dans cette base, la matrice de φ est diagonale et ses éléments diagonaux sont les $\varphi(e_j, e_j)$ avec $1 \leqslant j \leqslant n$, et on a $\varphi(e_j, e_j) \neq 0$ ssi $j \leqslant r$. On a alors $\operatorname{rang} \varphi = \operatorname{rang}[\varphi]_{\mathbf{e}} = r$, ce qui conclut la démonstration. □

Définition 4.2.3. Soit E un K-espace vectoriel, $F \subset E$ et $\varphi : E \times E \to \mathrm{K}$ une forme bilinéaire symétrique. L'espace *φ-orthogonal* de F, noté $F^{\perp_\varphi} \subset E$, est définie par

$$F^{\perp,\varphi} = \{x \in E \ : \ \varphi(x, y) = 0, \ y \in F\}.$$

Proposition 4.2.4. *Si E est un* K*-espace vectoriel, $\varphi \in \mathscr{B}(E)$ est symétrique et $F \subset E$ alors*

$$\dim F^{\perp,\varphi} + \dim F = n + \dim(F \cap \ker \varphi).$$

Démonstration. □

4.2.2 Loi d'inertie de Sylvester

Définition 4.2.5. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, q une forme quadratique sur E et φ sa forme polaire. La *signature* de φ (ou de q) est le couple $(p, q) \in \mathbb{N}^2$ où p et q sont respectivement donnés par la dimension maximale d'un sous-espace sur lequel φ est définie positive (respectivement définie négative). Autrement dit, on a

$$p = \max\bigl\{\dim F \ : \ F \subset E, \ q(x) > 0, \ x \in F \setminus \{0\}\bigr\}$$

tandis que q est donné par

$$q = \max\big\{\dim F \ : \ F \subset E,\ q(x) < 0,\ x \in F \setminus \{0\}\big\}.$$

Le théorème suivant montre que deux formes quadratiques réelles sont équivalentes si et seulement si elles ont même signature.

Théorème 4.2.6 (Loi d'inertie de Sylvester). *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n, φ une forme bilinéaire symétrique sur E, $r = \operatorname{rang}\varphi$ et (p, q) la signature de φ. Alors $p + q = r$ et il existe une base $\mathbf{e}$ de E la matrice $[\varphi]_{\mathbf{e}}$ est diagonale par blocs et donnée par*

$$[\varphi]_{\mathbf{e}} = \begin{pmatrix} \boxed{\mathrm{I}_p} & & 0 \\ & \boxed{-\mathrm{I}_q} & \\ 0 & & \boxed{0} \end{pmatrix},$$

Autrement dit, si $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n$ sont des éléments de E avec $x_j, y_j \in \mathbb{R}$, on a

$$\varphi(x, y) = \sum_{j=1}^{p} x_jy_j - \sum_{j=p+1}^{q} x_jy_j.$$

Démonstration. Soit $\widetilde{\mathbf{e}}$ une base de E qui est orthogonale pour φ, et telle que $\varphi(\widetilde{e}_j, \widetilde{e}_j)$ est non nul ssi $1 \leqslant j \leqslant r$. Soit $\tilde{p}$ le cardinal de l'ensemble des j vérifiant $\varphi(\widetilde{e}_j, \widetilde{e}_j) > 0$; quitte à réordonner les e_j, on peut supposer $\varphi(\widetilde{e}_j, \widetilde{e}_j) > 0$ ssi $1 \leqslant j \leqslant p$ de sorte que $\varphi(\widetilde{e}_j, \widetilde{e}_j) < 0$ pour tout $\tilde{p} + 1 \leqslant j \leqslant r$. On pose

$$e_j = |\varphi(e_j, e_j)|^{-1/2}\ \widetilde{e}_j, \quad j = 1, \ldots, r.$$

Alors $|\varphi(e_j, e_j)| = 1$ pour tout $1 \leqslant j \leqslant r$, donc

$$\varphi(e_j, e_j) = 1 \quad \text{si} \quad 1 \leqslant j \leqslant \tilde{p} \qquad \text{et} \qquad \varphi(e_j, e_j) = -1 \quad \text{si} \quad \tilde{p} + 1 \leqslant j \leqslant r.$$

Il reste à montrer que $\tilde{p} = p$ et que $r - p = q$ où (p, q) est la signature de φ. Si $G = \operatorname{vect}(e_1, \ldots, e_{\tilde{p}})$, on a $q(x) > 0$ pour pour tout $x \in G$ non nul. Ainsi on a bien $p \geqslant \dim G = \tilde{p}$. Réciproquement, montrons que $\tilde{p} \geqslant p$. Soit F un sous-espace sur lequel φ est définie positive et posons $G = \operatorname{vect}(e_{p+1}, \ldots, e_n)$. Puisque φ est définie positive sur F, on a $G \cap F = \{0\}$. Par conséquent

$$\dim F \leqslant n - \dim G = p,$$

d'où l'on tire $\tilde{p} \leqslant p$. En appliquant ce qui précède à la forme quadratique φ on obtient $r - \tilde{p} = q$, d'où $p + q = r$. Ceci achève la démonstration. □

4.2.3 Formes quadratiques complexes

Théorème 4.2.7 (Forme normale pour les formes quadratiques complexes). *On suppose ici* $\mathrm{K} = \mathbb{C}$. *Soit* E *un* $\mathbb{C}$*-espace vectoriel de dimension finie* n, φ *une forme bilinéaire symétrique sur* E *et* $r = \operatorname{rang}\varphi$. *Alors il existe une base* $\mathbf{e}$ *de* E *telle que*

$$[\varphi]_{\mathbf{e}} = \begin{pmatrix} \boxed{\mathrm{I}_r} & 0 \\ 0 & \boxed{0} \end{pmatrix}.$$

Autrement dit, si $x = x_1e_1 + \cdots + x_ne_n$ *et* $y = y_1e_1 + \cdots + y_ne_n$ *sont des éléments de* E *avec* $x_j, y_j \in \mathbb{C}$, *on a*

$$\varphi(x, y) = \sum_{j=1}^{r} x_jy_j.$$

Démonstration. On se donne $\widetilde{\mathbf{e}} = (\widetilde{e}_1, \ldots, \widetilde{e}_n)$ une base orthogonale pour φ telle que $\varphi(\widetilde{e}_j, \widetilde{e}_j) \neq 0$ est non nul ssi $1 \leqslant q \leqslant r$. Pour tout j, soit $\alpha_j \in \mathbb{C}$ une racine complexe de $\varphi(\widetilde{e}_j, \widetilde{e}_j)$, c'est-à-dire qui vérifie $\alpha_j^2 = \varphi(\widetilde{e}_j, \widetilde{e}_j)$. On pose $e_j = \alpha_j^{-1}\widetilde{e}_j$ pour tout $1 \leqslant j \leqslant r$ et $e_j = \widetilde{e}_j$ si $j > r$. Alors par bilinéarité on a

$$\varphi(e_j, e_j) = \alpha_j^{-2}\varphi(\widetilde{e}_j, \widetilde{e}_j) = 1$$

pour tout $j = 1, \ldots, r$ et $\varphi(e_j, e_j) = 0$ si $j > r$. En outre la base $\mathbf{e} = (e_1, \ldots, e_n)$ est orthogonale pour φ, ce qui achève la démonstration. □

4.3 Espaces euclidiens

4.3.1 Définitions et premières propriétés

Définition 4.3.1. Un *espace euclidien* $\big(E, (\,\cdot\,|\,\cdot\,)\big)$ est la donnée d'un $\mathbb{R}$-espace vectoriel E de dimension finie, muni d'un produit scalaire $(\,\cdot\,|\,\cdot\,)$ sur E — c'est-à-dire une forme bilinéaire symétrique définie positive. Pour tout $x \in E$, la *norme euclidienne* de x est donnée par

$$\|x\| = \sqrt{(x|x)}$$

Il est facile de vérifie que $\|\cdot\|$ est bien une norme sur E, c'est-à-dire qu'on a $x = 0$ ssi $\|x\| = 0$ et pour tous $\lambda \in \mathbb{R}$ et $x, y \in E$,

$$\|\lambda x\| = |\lambda| \cdot \|x\| \quad \text{et} \quad \|x + y\| \leqslant \|x\| + \|y\|.$$

Une *base orthonormale* $(e_1, \ldots, e_n)$ de E est une base orthogonale pour $(\,\cdot\,|\,\cdot\,)$ telle que $\|e_j\| = 1$ pour tout $j = 1, \ldots, n$.

Remarque 4.3.2. Par le Théorème 4.2.6, on sait que dans tout espace euclidien, il existe toujours une base ortho-normale. Dans une telle base, la matrice de $\varphi = (\cdot|\cdot)$ est donnée par I_n.

Théorème 4.3.3 (Pythagore). *Soit* $(E, (\cdot|\cdot))$ *un espace euclidien et* $(u_1, \ldots, u_k)$ *une famille orthogonale. Alors*

$$\left\|\sum_{\ell=0}^{k} u_\ell\right\|^2 = \sum_{\ell=0}^{k} \|u_\ell\|^2.$$

Démonstration. On a par bilinéarité et par orthogonalité

$$\left\|\sum_{\ell=0}^{k} u_\ell\right\|^2 = \left(\sum_{\ell=0}^{k} u_\ell \middle| \sum_{\ell=0}^{k} u_\ell\right) = \sum_{0\leqslant j,\ell\leqslant k} (u_j|u_\ell) = \sum_{\ell=0}^{k} \|u_\ell\|^2,$$

ce qui est l'égalité voulue. □

4.3.2 Endomorphisme adjoint

Soit $\big(E, (\cdot|\cdot)\big)$ un espace euclidien. Si $f \in \mathscr{L}(E)$, on notera $f^\star = f^{\star,\varphi}$ où $\varphi = (\cdot|\cdot)$, cf. le Théorème-Définition 4.1.17. Autrement dit, $f^\star \in \mathscr{L}(E)$ est défini par

$$(x|f(y)) = (f^\star(x)|y), \quad x, y \in E.$$

On notera aussi $F^\perp = F^{\perp,\varphi}$ pour tout sous-espace $F \subset E$. Une application immédiate de la Proposition 4.1.19 est la suivante.

Proposition 4.3.4 (Propriétés de l'adjoint euclidien). *Soit* $\big(E, (\cdot|\cdot)\big)$ *un espace euclidien. Alors pour tous* $f, g \in \mathscr{L}(E)$ *et* $\lambda \in \mathbb{R}$, *on a*

(i) $(f + g)^\star = f^\star + g^\star$;

(ii) $(\lambda f)^\star = \lambda f^\star$;

(iii) $(f \circ g)^\star = g^\star \circ f^\star$;

(iv) $(f^\star)^\star = f$;

(v) $\mathrm{id}_E^\star = \mathrm{id}_E$.

Définition 4.3.5. On dit que $f \in \mathscr{L}(E)$ est auto-adjoint ou symétrique (resp. anti auto-adjoint ou anti-symétrique) si $f = f^\star$ (resp. $f = -f^\star$). On notera $\mathscr{S}(E)$ (resp. $\mathscr{A}(E)$) l'espace des endomorphismes symétriques.

Une application directe de la Proposition 4.1.18 est la suivante.

PROPOSITION 4.3.6. *Soit* $\mathbf{e}$ *une base orthonormale de* E*. Alors pour tout* $f \in \mathscr{S}(E)$*, on a*

$$[f^\star]_{\mathbf{e}} = {}^t[f]_{\mathbf{e}}.$$

Ainsi l'espace $\mathscr{S}(E)$ *(resp.* $\mathscr{A}(E)$*) est un sous-espace de dimension* $n(n+1)/2$ *(resp.* $n(n-1)/2$*) de* $\mathscr{L}(E)$*.*

PROPOSITION 4.3.7 (Préservation des orthogonaux). *Soit* $f \in \mathscr{S}(E)$ *et* $F \subset E$ *qui est stable par* f*. Alors* $F^\perp$ *est stable par* f*.*

Démonstration. Soit $x \in F^\perp$. Pour tout $y \in F$, on a $f(y) \in F$ par hypothèse et comme $f \in \mathscr{S}(E)$, on obtient

$$(f(x)|y) = (x|f(y)) = 0.$$

Ainsi $f(x) \in F^\perp$ et $f(F^\perp) \subset F^\perp$. □

4.3.3 Théorème spectral

THÉORÈME 4.3.8 (Théorème spectral pour les endomorphismes). *Soit* $\big(E, (\cdot\,|\,\cdot)\big)$ *un espace euclidien et* $f \in \mathscr{S}(E)$*. Alors il existe une base orthonormale* $\mathbf{e}$ *de* E *telle que* $[f]_{\mathbf{e}}$ *est diagonale.*

Avant de donner la preuve du théorème spectral, énonçons trois résultats intermédiaires.

LEMME 4.3.9 (Existence d'une droite ou d'un plan stable pour les endomorphismes réels). *Soit* E *un* $\mathbb{R}$*-espace vectoriel de dimension* $n \geqslant 1$ *et* $f \in \mathscr{L}(E)$*. Alors il existe* $F \subset E$ *un sous-espace de dimension* 1 *ou* 2 *qui est stable par* f*.*

Démonstration. On peut supposer $E = \mathbb{R}^n$. Si χ_f admet une racine, alors f admet un vecteur propre x et $F = \mathbb{R}x$ est stable par f. Sinon, χ_f n'a que des racines dans $\mathbb{C}\setminus\mathbb{R}$. Comme χ_f est un polynôme réel, on a $\overline{\chi_f(\lambda)} = \chi_f(\bar\lambda)$ pour tout $\lambda \in \mathbb{C}$, donc λ est une racine de χ_f ssi $\bar\lambda$ en est une. Par suite on peut écrire

$$\chi_f = \prod_{j=1}^{r} (X-\lambda_j)^{\alpha_j}(X-\bar\lambda_j)^{\beta_j}$$

où les $\lambda_j \in \mathbb{C}$ sont deux à deux distincts. En utilisant encore que χ_f est réel, on voit que $\alpha_j = \beta_j$. Le théorème de Cayley–Hamliton et lemme des noyaux donnent alors

$$E = \bigoplus_{j=1}^{r} \ker(Q_j(f)^{\alpha_j}) \quad \text{où} \quad Q_j = (X-\lambda_j)(X-\bar\lambda_j) \in \mathbb{R}[X].$$

En particulier il existe j tel que $\ker(Q_j(f)^{\alpha_j})$ n'est pas réduit à $\{0\}$ donc $\ker Q_j(f)$ non plus. Mais alors si $x \in \ker Q_j(f) \setminus \{0\}$, on a

$$0 = Q_j(f)(x) = f^2(x) - 2\mathrm{Re}(\lambda_j)f(x) + |\lambda_j|^2 x,$$

donc $f^2(x)$ est combinaison linéaire de $f(x)$ et de x. Ainsi $F = \mathrm{vect}(x, f(x))$ est stable par f. □

Lemme 4.3.10 (Orthogonalité des vecteurs propres d'un endomorphisme symétrique). *Deux vecteurs propres d'un endomorphisme symétrique associés à des valeurs propres (réelles) distinctes sont orthogonaux.*

Démonstration. Soit E un espace euclidien et $f \in \mathscr{S}(E)$. Soient $\lambda, \mu \in \mathbb{R}$ distincts et $x, y \in E$ tels que $f(x) = \lambda x$ et $f(y) = \mu y$. On a

$$\lambda(x|y) = (f(x)|y) = (x|f(y)) = \mu(x|y),$$

donc $(\lambda - \mu)(x|y)$ ce qui donne $(x|y) = 0$. □

Lemme 4.3.11 (Théorème spectral en dimension 2). *Si* $\dim F = 2$ *et* $g \in \mathscr{S}(F)$, *alors il existe une base orthonormale de vecteurs propres pour* g.

Démonstration. Soit $\mathbf{f}$ une base orthonormale de F, et $A = [g]_{\mathbf{f}} \in \mathrm{M}_2(\mathbb{R})$. Alors $A = (a_{i,j})$ est symétrique, donc son polynôme caractéristique s'écrit

$$\chi_g = X^2 - \mathrm{tr}(A)X + \det(A) = X^2 - (a+c) + ac - b^2$$

où on a noté $a = a_{1,1}, c = a_{2,2}$ et $b = a_{1,2} = a_{2,1}$. Si $b = 0$, alors A est diagonale et il n'y a rien à démontrer. Sinon, le discriminant $(a+c)^2 - 4ac + 4b^2 = (a-c)^2 + 4b^2$ de χ_g est strictement positif, donc χ_g a deux racines distinctes. Par conséquent il existe une base (e_1, e_2) de F formée de vecteurs propres de g. Le lemme 4.3.10 implique alors que la famille (e_1, e_2) est orthogonale. Quitte à remplacer e_j par $e_j/\|e_j\|$, on peut les supposer les e_j de norme 1, de sorte que (e_1, e_2) est orthonormale. □

Démonstration du théorème spectral. On raisonne par récurrence sur la dimension. Pour $n = 1$, c'est trivial, et le cas $n = 2$ découle du Lemme 4.3.11. On suppose donc le résultat vrai en toute dimension $1 \leqslant k \leqslant n-1$ pour $n - 1 \geqslant 2$. Soit E un espace euclidien de dimension n. Le lemme 4.3.9 nous donne l'existe d'un espace F de dimension 1 ou 2 tel que $f(F) \subset F$. Soit $g = f|_F \in \mathscr{L}(F)$. Alors par hypothèse de récurrence, il existe une base orthonormale $\mathbf{f}$ de F formée de vecteurs propre de g (donc de f). Notons que $F^\perp$ est stable par f par la Proposition 4.3.7. Comme $\dim F^\perp \leqslant n-1$, l'hypothèse de récurrence nous donne une base ortho-normale $\mathbf{g}$ de $F^\perp$ qui est formée de vecteurs propres de f. Mais alors $\mathbf{f} \oplus \mathbf{g}$ est une base orthonormale de E formée de vecteurs propres de f. La récurrence est établie. □

On conclut ce paragraphe avec une démonstration alternative du théorème spectral basée sur le calcul différentiel, qui n'utilise pas les trois lemmes précédents. En particulier, elle n'utilise ni le lemme des noyaux ni le théorème de Cayley–Hamilton.

Démonstration alternative du théorème spectral. On procède encore par récurrence sur la dimension. Pour $n = 1$, c'est trivial, et on suppose le résultat vraie en dimension $n - 1 \geqslant 1$. Soit $f \in \mathscr{L}(\mathbb{R}^n)$. Soit $S = \{x \in \mathbb{R}^n \ : \ \|x\| = 1\}$ la sphère unité de $\mathbb{R}^n$. Alors S est une partie compacte de $\mathbb{R}^n$, donc l'application

$$\Phi : S \to \mathbb{R}, \quad x \mapsto (f(x)|x)$$

admet un maximum qu'elle atteint en un point $x_\star \in S$. On affirme que $x_\star$ est un vecteur propre de S. En effet, soit $F = \mathbb{R}x_\star$ et $x \in F^\perp$ de norme 1. On pose

$$\gamma(t) = \cos(t)x_\star + \sin(t)x, \quad t \in \mathbb{R}.$$

Comme $x \in F^\perp$ on a $\|\gamma(t)\|^2 = \cos^2(t)\|x_\star\|^2 + \sin^2(t)\|x\|^2 = 1$ pour tout t, donc $\gamma(t) \in S$. En outre, on a

$$\begin{aligned}\Phi(\gamma(t)) &= \Big(\cos(t)f(x_\star) + \sin(t)f(x)\Big|\cos(t)x_\star + \sin(t)x\Big) \\ &= \cos(t)^2\Phi(x_\star) + \sin(t)^2\Phi(x) + 2\cos(t)\sin(t)(f(x_\star)|x).\end{aligned}$$

Comme γ admet un maximum au point $t = 0$, on a $\gamma'(0) = 0$ ce qui donne $0 = 2(f(x_\star)|x)$. Par conséquent $f(x_\star) \perp x$. Ceci étant vrai pour tout $x \in F^\perp$ on obtient $f(x_\star) \in (F^\perp)^\perp = F = \mathbb{R}x_\star$ donc $x_\star$ est vecteur propre de f. En outre f préserve aussi $F^\perp$ et par hypothèse de récurrence il existe une base orthonormale $\mathbf{g}$ de $F^\perp$ formée de vecteurs propres de f. Comme $x_\star \in S$, la base $\mathbf{e} = x_\star \oplus \mathbf{g}$ est orthonormale et formée de vecteurs propres de f. La récurrence est établie. $\square$

4.4 Endomorphismes orthogonaux

Dans cette section, on se donne $\big(E, (\,\cdot\,|\,\cdot\,)\big)$ un espace euclidien. En particulier, $\mathrm{K} = \mathbb{R}$.

4.4.1 Définitions, premières propriétés

Définition 4.4.1. Un endomorphisme $f \in \mathscr{L}(E)$ est dit *orthogonal* si pour tous $x, y \in E$, on a

$$(f(x)|f(y)) = (x|y).$$

On note $\mathscr{O}(E)$ l'ensemble des endomorphismes orthogonaux de E.

En particulier pour tout $f \in \mathscr{O}(E)$, on a

$$\|f(x)\|^2 = \|x\|^2, \quad x \in E.$$

Définition 4.4.2. Une matrice $A \in \mathrm{M}_n(\mathbb{R})$ est dite *orthogonale* si ${}^tAA = \mathrm{I}_n$. On note $\mathrm{O}_n(\mathbb{R})$ l'espace des matrice orthogonales.

Proposition 4.4.3. *Soit $f \in \mathscr{L}(E)$. Alors les propriétés suivantes sont équivalentes :*

(i) *f est orthogonal ;*

(ii) *f est un isomorphisme et $f^{-1} = f^\star$;*

(iii) *il existe une base de E dont l'image par f est une base orthonormale de E ;*

(iv) *l'image par f de toute base orthonormale de E est une base orthonormale de E ;*

(v) *il existe une base orthonormale $\mathbf{e}$ de E telle que la matrice $[f]_{\mathbf{e}}$ est orthogonale ;*

(vi) *pour toute base orthonormale $\mathbf{e}$ de E, la matrice $[f]_{\mathbf{e}}$ est orthogonale.*

Démonstration. L'équivalence entre (i), (ii), (iii) et (iv) est immédiate. L'équivalence entre (iii) et (v) découle de la Proposition 4.3.6. Les implications (iii) $\Rightarrow$ (iv) et (v) $\Rightarrow$ (vi) sont triviales. $\square$

Proposition 4.4.4. *Une matrice est orthogonale si et seulement si c'est la matrice de passage entre deux bases orthogonales de E.*

Démonstration. Soit $A \in \mathrm{O}_n(\mathbb{R})$. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base orthonormée de E et $f \in \mathscr{L}(E)$ telle que $[f]_{\mathbf{e}} = A$. Alors f est un endomorphisme orthogonal, donc $\mathbf{f} = (f(e_1), \ldots, f(e_n))$ est une base orthonormée de E. Il suit que $[f]_{\mathbf{e}}$ est la matrice de passage de $\mathbf{e}$ à $\mathbf{f}$. $\square$

Théorème 4.4.5 (Théorème spectral pour les matrices). *Soit $A \in \mathrm{M}_n(\mathbb{R})$ une matrice symétrique. Alors il existe $Q \in \mathrm{O}_n(\mathbb{R})$ et une matrice diagonale D telles que*

$$D = Q^{-1}AQ = {}^tQAQ.$$

Démonstration. Soit A symétrique. Soit $f \in \mathscr{L}(\mathbb{R}^n)$ l'endomorphisme canoniquement associé à A. Alors f est symétrique donc il existe une base orthonormée $\mathbf{e}$ formée de vecteurs propres de f, de sorte que $D = [f]_{\mathbf{e}}$ est diagonale. Si $\mathbf{b}$ est la base canonique de $\mathbb{R}^n$, on note Q la matrice de passage de $\mathbf{b}$ à $\mathbf{e}$. Alors

$$D = [f]_{\mathbf{e}} = Q^{-1}[f]_{\mathbf{b}}Q = Q^{-1}AQ = {}^tQAQ$$

puisque Q est orthogonale. $\square$

PROPOSITION 4.4.6. *Soit* $f \in \mathscr{O}(E)$. *Alors* $\det(f) = \pm 1$.

Démonstration. Soit $\mathbf{e}$ une base orthonormale de E. Alors $[f]_{\mathbf{e}}$ est une matrice orthogonale, donc

$$1 = \det\big([f]_{\mathbf{e}}^{-1}[f]_{\mathbf{e}}\big) = \det\big({}^t[f]_{\mathbf{e}}[f]_{\mathbf{e}}\big) = \det([f]_{\mathbf{e}})^2,$$

d'où le résultat. □

4.4.2 Classification des endomorphismes orthogonaux en dimension 3

Dans ce paragraphe nous démontrons le résultat suivant.

THÉORÈME 4.4.7 (Endomorphismes orthogonaux en dimension 3). *On suppose que E est un espace euclidien de dimension $n = 3$. Soit $f \in \mathscr{O}(E)$. Alors il existe une base orthonormale $\mathbf{e}$ et des réels $\lambda, \eta \in \{-1, 1\}$ et $\theta \in \mathbb{R}$ tels que*

$$[f]_{\mathbf{e}} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\eta\sin\theta \\ 0 & \sin\theta & \eta\cos\theta \end{pmatrix}.$$

REMARQUE 4.4.8. Le Théorème 4.4.7 dit que tout endomorphisme orthogonal de $\mathbb{R}^3$ est soit une rotation autour d'un axe, soit une symétrie orthogonale par rapport à un plan suivi d'une rotation autour de l'axe perpendiculaire à ce plan.

Nous aurons besoin du résultat analogue en dimension 2.

THÉORÈME 4.4.9 (Endomorphismes orthogonaux en dimension 2). *On suppose que E est un espace euclidien de dimension $n = 2$. Soit $f \in \mathscr{O}(E)$. Alors il existe une base orthonormale $\mathbf{e}$ et $\theta \in \mathbb{R}$ tels que*

$$[f]_{\mathbf{e}} = \begin{pmatrix} \cos\theta & -\eta\sin\theta \\ \sin\theta & \eta\cos\theta \end{pmatrix}.$$

où $\eta = \det f \in \{-1, 1\}$.

Démonstration. Soit $\mathbf{e} = (e_1, e_2)$ une base orthonormale de E. Écrivons $f(e_1) = \alpha e_1 + \beta e_2$ avec $\alpha, \beta \in \mathbb{R}$. Comme f est orthogonal, on a

$$1 = \|f(e_1)\|^2 = \alpha^2 + \beta^2.$$

Par suite il existe $\theta \in \mathbb{R}$ tel que $\alpha = \cos\theta$ et $\beta = \sin\theta$. Comme $f(e_2)$ est orthogonal à e_1 et $\|f(e_2)\|^2 = 1$ on a $f(e_2) = \eta(-\sin(\theta)e_1 + \cos(\theta)e_2)$ avec $\eta = \pm 1$. Alors $[f]_{\mathbf{e}}$ est de la forme annoncée et on calcule $\eta = \det[f]_{\mathbf{e}} = \det f$. □

Démonstration du Théorème 4.4.7. On suppose que E est un espace euclidien de dimension 3. Soit $f \in \mathscr{O}(E)$. Alors $\chi_f \in \mathbb{R}_3[X]$ donc χ_f admet au moins une racine réelle, que l'on note $\lambda \in \mathbb{R}$. Soit $x \in E$ non nul un vecteur propre associé de norme 1. On a

$$1 = \|x\|^2 = \|f(x)\|^2 = \|\lambda x\|^2 = |\lambda|$$

donc $\lambda = \pm 1$. Soit $F = (\mathbb{R}x)^{\perp}$. Alors pour tout $y \in F$, on a, puisque $x = \eta f(x)$,

$$(f(y)|x) = \eta(f(y)|f(x)) = \eta(y|x) = 0.$$

Par conséquent f préserve F. Par le Théorème 4.4.9, il existe une base $\mathbf{f}$ de F et des réels $\eta \in \{-1, 1\}$ et $\theta \in \mathbb{R}$ tels que

$$[f|_F]_{\mathbf{f}} = \begin{pmatrix} \cos\theta & -\eta\sin\theta \\ \sin\theta & \eta\cos\theta \end{pmatrix}.$$

Soit $\mathbf{e} = x \oplus \mathbf{f}$. Alors on a

$$[f]_{\mathbf{e}} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\eta\sin\theta \\ 0 & \sin\theta & \eta\cos\theta \end{pmatrix},$$

et on calcule $\lambda\eta = \det f$. $\square$