HAX301X – Algèbre III
Réduction des endomorphismes

Damien Calaque

Table des matières

Introduction

1. Objectifs

L'objectif principal de ce cours est d'étudier et comprendre en détail l'ensemble $\mathrm{End}(E)$ des endomorphismes d'un espace vectoriel E de dimension finie. On rappelle que si E est un $\mathbb{K}$-espace vectoriel de dimension finie alors $E \cong \mathbb{K}^n$, où $n = \dim(E)$.
On s'intéressera plus précisément à deux sous-objectifs de notre objectif principal :

(1) Étant donné $L \in \mathrm{End}(E)$, on fera en sorte de trouver une base $\mathcal{B}$ de E qui soit adaptée pour que la matrice $M_{\mathcal{B},\mathcal{B}}(L)$ de L dans cette base soit le plus simple possible (par exemple, une matrice diagonale). Ce problème peut se reformuler de la manière (équivalente) qui suit : étant donnée une matrice carrée M de taille $n \times n$, trouver une matrice N semblable (on dit aussi conjuguée ; c'est-à-dire qu'il existe une matrice inversible $P \in GL_n(\mathbb{K})$ telle que $N = P^{-1}MP$) qui soit le plus simple possible.

(2) Trouver des fonctions "intéressantes" $\alpha : \mathrm{End}(E) \to X$ (X un ensemble), qui vont nous permettre de distinguer des endomorphismes : si $\alpha(L_1) \neq \alpha(L_2)$ alors $L_1 \neq L_2$. Pour trouver une telle fonction lorsque E est de dimension finie n, il suffit d'exhiber une fonction $a : \mathrm{Mat}_{n\times n}(\mathbb{K}) \to X$ qui est **invariante par conjugaison** :

$$\boxed{\text{Pour tout } P \in GL_n(\mathbb{K}) \text{ et tout } M \in \mathrm{Mat}_{n\times n}(\mathbb{K}),\ a(P^{-1}MP) = a(M).}$$

En effet, on définit alors $\alpha(L) := a\big(M_{\mathcal{B},\mathcal{B}}(L)\big)$, où $\mathcal{B}$ est n'importe quelle base de E. Le résultat est indépendant du choix de $\mathcal{B}$: pour toute autre base $\mathcal{B}'$, si on pose $P := P_{\mathcal{B}}^{\mathcal{B}'}$ alors $M_{\mathcal{B}',\mathcal{B}'}(L) = P^{-1}M_{\mathcal{B},\mathcal{B}}(L)P$ et donc $a\big(M_{\mathcal{B}',\mathcal{B}'}(L)\big) = a\big(M_{\mathcal{B},\mathcal{B}}(L)\big)$.

Exemple 1.1 (trace d'une matrice)**.** On définit la trace d'une matrice carrée comme la somme de ses coefficients diagonaux :

$$\begin{array}{rcl} \mathrm{tr}\colon \mathrm{Mat}_{n\times n}(\mathbb{K}) & \longrightarrow & \mathbb{K} \\ A = (a_{ij})_{1\le i,j\le n} & \longmapsto & \sum_{i=1}^{n} a_{ii}\,. \end{array}$$

On peut montrer par le calcul que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$:

$$\begin{aligned} \mathrm{tr}(AB) &= \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \sum_{j=1}^{n}(BA)_{jj} = \mathrm{tr}(BA)\,. \end{aligned}$$

On en déduit que si P est inversible alors $\mathrm{tr}(P^{-1}MP) = \mathrm{tr}(MPP^{-1}) = \mathrm{tr}(M)$. △

On déduit de l'exemple précédent qu'on peut définir sans ambiguité la trace d'un endomorphisme d'un espace vectoriel de dimension finie. La trace nous permet également de montrer que des matrices ne sont pas semblables : $\mathrm{tr}\begin{pmatrix}1 & 3\\ 3 & 1\end{pmatrix} = 2 \neq 6 = \mathrm{tr}\begin{pmatrix}3 & 1\\ 1 & 3\end{pmatrix}$, et donc $\begin{pmatrix}1 & 3\\ 3 & 1\end{pmatrix}$ n'est pas conjuguée à $\begin{pmatrix}3 & 1\\ 1 & 3\end{pmatrix}$.

Exercice 1.2. *Vérifier que pour tout $k \in \mathbb{N}$, l'application $M \mapsto \mathrm{tr}(M^k)$ est invariante par conjugaison.*

2. Applications

Voyons maintenant deux applications du premier sous-objectif.

2.1. Puissances d'une matrice. Supposons qu'on se donne une matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ ou un endomorphisme $L \in \mathrm{End}(E)$, et qu'il existe
— $P \in GL_n(\mathbb{K})$ telle que $A := P^{-1}MP$ soit "simple",
— ou une base $\mathcal{B}$ de E telle que $A := M_{\mathcal{B},\mathcal{B}}(L)$ soit "simple".
Dans les deux cas on peut exprimer facilement les puissances de M et L en fonction de celles de A : en effet, pour tout $k \in \mathbb{N}$,
— $M^k = (PAP^{-1})^k = PA\underbrace{P^{-1}P}_{=\mathrm{I}_n}AP\cdots PAP^{-1} = PA^kP^{-1}$, et
— $L^{\circ k}$ est l'unique endomorphisme de E tel que $A^k = M_{\mathcal{B},\mathcal{B}}(L^{\circ k})$.
Dans les deux situations, comme A est simple (quoi que cela veuille dire), le calcul en sera facilité.

Exemple 2.1. On considère la matrice $M = \begin{pmatrix}0 & 2 & 1\\ -1 & 3 & 1\\ -2 & 2 & 3\end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$. On remarque que $M\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $M\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}2\\2\\0\end{pmatrix}$ et $M\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}3\\3\\3\end{pmatrix}$. On en déduit que l'endomorphisme L de $\mathbb{R}^3$ qui est donné par M dans la base canonique $\mathcal{C}$ est donné par $A := diag(1,2,3)$ dans la base

$$\mathcal{B} := \left(\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right).$$

Par conséquent si on note $P = P_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 1 & 0 & 1\end{pmatrix}$, alors on peut montrer que $P^{-1} = P_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix}1 & -1 & 0\\ 1 & 0 & -1\\ -1 & 1 & 1\end{pmatrix}$ (le faire en exercice!). On en déduit que $M = PAP^{-1}$, et donc que

$$\begin{aligned} M^k = PA^kP^{-1} &= \begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 2^k & 0\\ 0 & 0 & 3^k\end{pmatrix}\begin{pmatrix}1 & -1 & 0\\ 1 & 0 & -1\\ -1 & 1 & 1\end{pmatrix}\\ &= \begin{pmatrix}1 & 2^k & 3^k\\ 0 & 2^k & 3^k\\ 1 & 0 & 3^k\end{pmatrix}\begin{pmatrix}1 & -1 & 0\\ 1 & 0 & -1\\ -1 & 1 & 1\end{pmatrix}\end{aligned}$$

$$= \begin{pmatrix} 1+2^k-3^k & 3^k-1 & 3^k-2^k \\ 2^k-3^k & 3^k & 3^k-2^k \\ 1-3^k & 3^k-1 & 3^k \end{pmatrix}$$

△

2.2. Suites récurrentes. Voyons comment les choses fonctionnent sur un exemple. Considérons deux suites $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ qui satisfont la relation de récurrence linéaire qui suit :

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix} \quad \text{avec} \quad \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Autrement dit : $u_{n+1} = v_n$ et $u_{n+2} = v_{n+1} = u_n + v_n = u_n + u_{n+1}$. Certain·es d'entre vous reconnaitront en u_n la **suite de Fibonacci**. Notons $\varphi := \frac{1+\sqrt{5}}{2}$ et $\psi := -\varphi^{-1} = \frac{1-\sqrt{5}}{2}$. On peut vérifier que

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \varphi \end{pmatrix} = \begin{pmatrix} \varphi \\ 1+\varphi \end{pmatrix} = \begin{pmatrix} \varphi \\ \varphi^2 \end{pmatrix} = \varphi \begin{pmatrix} 1 \\ \varphi \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \psi \end{pmatrix} = \begin{pmatrix} 1 \\ 1+\psi \end{pmatrix} = \begin{pmatrix} \psi \\ \psi^2 \end{pmatrix} = \psi \begin{pmatrix} 1 \\ \psi \end{pmatrix}$$

Par conséquent, l'endomorphisme L de $\mathbb{R}^2$ qui est donné par $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ dans la base canonique $\mathcal{C}$ est donné par $diag(\varphi, \psi)$ dans la base

$$\mathcal{B} := \left(\begin{pmatrix} 1 \\ \varphi \end{pmatrix}, \begin{pmatrix} 1 \\ \psi \end{pmatrix} \right) .$$

La matrice de changement de base est $P := P_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ \varphi & \psi \end{pmatrix}$, d'inverse $P^{-1} = P_{\mathcal{B}}^{\mathcal{C}} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix}$.

On en déduit que $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = P \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix} P^{-1}$, et donc

$$\begin{aligned} \begin{pmatrix} u_n \\ v_n \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = P \begin{pmatrix} \varphi^n & 0 \\ 0 & \psi^n \end{pmatrix} P^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= P \begin{pmatrix} \varphi^n & 0 \\ 0 & \psi^n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{pmatrix} = P \begin{pmatrix} \frac{\varphi^n}{\sqrt{5}} \\ -\frac{\psi^n}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\varphi^n - \psi^n}{\sqrt{5}} \\ \frac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}} \end{pmatrix} . \end{aligned}$$

On obtient ainsi une formule pour le terme général de la suite de Fibonacci :

$$u_n = \frac{\varphi^n - \psi^n}{\sqrt{5}} .$$

Chapitre 1

Déterminant

1. Groupe symétrique

On rappelle[1] qu'un **groupe** est un ensemble G muni d'une loi de composition interne $\star : G \times G \to G$ qui est associative, admet un élément neutre, et telle que tout élément de G admet un inverse.

Exercice 1.1 (rappel de HAX202X)**.** *Écrire ce que chacune de ces trois propriétés signifie. Montrer que lorsque qu'il existe, l'élément neutre est unique. Même chose avec l'inverse d'un élément.*

1.1. Le vocabulaire du groupe symétrique.

Soit $X = \{1, \dots, n\}$, $n \geq 1$. On note $\mathfrak{S}_n$ (ou aussi S_n) l'ensemble $\mathrm{Bij}(X)$ des bijections de X dans X, muni de la loi donnée par la composition des bijections $\circ$. On appelle $\mathfrak{S}_n$ le **groupe symétrique** d'indice n. Les éléments de $\mathfrak{S}_n$ sont appelées **permutations** (de X). Une permutation $\sigma \in \mathfrak{S}_n$ est parfois notée de manière explicite de la manière suivante :

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

Le **support** d'une permutation σ est l'ensemble

$$supp(\sigma) := \{x \in X \,|\, \sigma(x) \neq x\} .$$

Exemple 1.2. Pour $n = 4$, $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$ définit un élément de $\mathfrak{S}_4$, dont le support est $\{1, 2, 3\}$. Cet exemple est un cas particulier de permutation appelée **cycle** (ici de longueur 3). △

Exercice 1.3 (rappel de HAX203X)**.** *Montrer que $\mathfrak{S}_n$ a $n!$ éléments.*

Définition 1.4 (cycle)**.** Soit $k \in \{1, \dots, n\}$. Un **cycle de longueur** k est une permutation $\sigma \in \mathfrak{S}_n$ telle qu'il existe k éléments deux à deux distincts $(i_1, \dots, i_k)$ de X satisfaisant[2] :

- $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \dots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1$;
- $supp(\sigma) = \{i_1, \dots, i_k\}$.

On note alors $\sigma = (i_1\, i_2\, \cdots\, i_k)$.
Les cycles de longueur 2 sont appelés **transpositions**.

Le cycle de l'Exemple 1.2 s'écrit $(1\,2\,3)$.

Remarque 1.5. Il n'y a pas de cycle de longueur 1, car le support d'une permutation ne peut pas être un singleton.

Remarque 1.6. L'écriture d'un cycle ne dépend que de l'ordre cyclique des éléments de son support : $(i_1\, i_2\, \cdots\, i_k) = (i_2\, \cdots\, i_k\, i_1)$. Pour le cycle de l'Exemple 1.2 on a $(1\,2\,3) = (2\,3\,1) = (3\,1\,2)$; mais attention, $(1\,2\,3) \neq (2\,1\,3)$.

1. HAX202X – Algèbre II
2. Autrement dit : un **arrangement** de k éléments parmi n.

Exercice 1.7. *Soit $n = 4$. Montrer que $(1\,3) \circ (2\,4) = (2\,4) \circ (1\,3) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$.*

Exercice 1.8. *Montrer que tous les éléments de $\mathfrak{S}_2$ sont des cycles. Même chose avec $\mathfrak{S}_3$. Montrer qu'en revanche $\mathfrak{S}_4$ contient des éléments qui ne sont pas des cycles.*

À partir de maintenant on notera $\sigma_1\sigma_2 := \sigma_1 \circ \sigma_2$ la composition de deux permutations, qu'on appellera **produit**.

1.2. Décomposition en produit de cycles à supports disjoints.

Nous allons maintenant démontrer le résultat qui suit :

Théorème 1.9 (décomposition en produit de cycles)**.** *Toute permutation $\sigma \in \mathfrak{S}_n$ est égale à un produit de cycles à supports (deux à deux) disjoints. Ce produit est unique à permutation près des facteurs.*

On a vu un tel exemple d'une factorisation en produit de cycles à l'Exemple 1.7.

Exercice 1.10. *Factoriser la permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 4 & 5 & 3 & 2 & 1 \end{pmatrix}$ en produit de cycles à supports disjoints.*

Remarque 1.11. La permutation identité est le produit vide. En effet, de même que la somme (indexée par l'ensemble) vide de réels vaut 0 par convention, que le produit vide de réels vaut 1 par convention, le produit vide dans un groupe est l'élément neutre du groupe.

Définition 1.12 (orbite)**.** Étant donné $\sigma \in \mathfrak{S}_n$ et $i \in X$, on définit l'**orbite de i selon** σ comme l'ensemble

$$\mathcal{O}_\sigma(i) := \{\sigma^k(i) \,|\, k \in \mathbb{Z}\} \subset X = \{1, \ldots, n\}\,.$$

Posons

$$k_i := min\{\ell \in \mathbb{N}^* \,|\, \sigma^\ell(i) = i\}\,.$$

Si $k_i > 1$ alors à une telle orbite on peut associer un cycle $c_{\sigma,i} := \big(i\,\sigma(i)\,\cdots\sigma^{k_i-1}(i)\big)$. Sinon $k_i = 1$, ce qui équivaut à $\mathcal{O}_\sigma(i) = \{i\}$, ou encore à $i \notin supp(\sigma)$.

Remarque 1.13. Il faut quand même vérifier que k_i est bien défini, c'est-à-dire que l'ensemble

$$\{\ell \in \mathbb{N}^* \,|\, \sigma^\ell(i) = i\} \subset \mathbb{N}^*$$

est non vide. C'est bien le cas : l'orbite $\mathcal{O}_\sigma(i)$ est un ensemble fini (puisque c'est une partie de X, qui a lui-même un nombre fini n d'éléments), donc il existe deux entiers distincts m et n tels que $\sigma^m(i) = \sigma^n(i)$; par conséquent si on pose $\ell = |m - n| \in \mathbb{N}^*$ alors on obtient $\sigma^\ell(i) = i$.

Lemme 1.14. *L'orbite de i selon σ a exactement k_i éléments, et*

$$\mathcal{O}_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \ldots, \sigma^{k_i-1}(i)\}\,.$$

Démonstration. Supposons que $\sigma^\ell(i) = i$. Pour tout $n \in \mathbb{Z}$, on note $n = p\ell + r$ la division euclidienne de n par ℓ (en particulier, $0 \le r < \ell$). Alors $\sigma^n(i) = \sigma^r(\sigma^\ell)^p(i) = \sigma^r(i)$. Par conséquent

$$\mathcal{O}_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \ldots, \sigma^{\ell-1}(i)\}\,.$$

Il nous reste à démontrer que si $\ell = k_i$, alors il n'y a aucune répétition dans la liste des éléments $i, \sigma(i), \sigma^2(i), \ldots, \sigma^{k_i-1}(i)$. Si c'était le cas, on aurait $\sigma^m(i) = \sigma^n(i)$, avec $0 \le m \ne n < k_i$, et par conséquent $\sigma^{|m-n|}(i) = i$, avec $0 < |m - n| < k_i$, ce qui contredirait la minimalité de k_i. □

DÉMONSTRATION DU THÉORÈME 1.9.

Exercice 1.15. *Montrer que pour tout $i, j \in X$ il n'y a que deux possibilités : soit $\mathcal{O}_\sigma(i) = \mathcal{O}_\sigma(j)$, soit $\mathcal{O}_\sigma(i) \cap \mathcal{O}_\sigma(j) = \emptyset$.*

L'ensemble des orbites selon σ forme donc une partition de X :

$$\{1, \ldots, n\} = \mathcal{O}_\sigma(j_1) \sqcup \cdots \sqcup \mathcal{O}_\sigma(j_m)\,.$$

En retirant les orbites constituées d'un seul élément, on obtient une partition du support de σ :

$$supp(\sigma) = \mathcal{O}_\sigma(i_1) \sqcup \cdots \sqcup \mathcal{O}_\sigma(i_k)\,.$$

On a alors $\sigma = c_{\sigma,i_1} \cdots c_{\sigma,i_k}$. On peut permuter les facteurs :

Exercice 1.16. *Montrer que deux permutations (et donc en particulier deux cycles) de supports disjoints commutent.*

Démontrons maintenant l'unicité de la décomposition (à l'ordre des facteurs près). Toute décomposition $\sigma = c_1 \cdots c_k$ en produit de cycles à supports disjoints induit une partition du support de σ :

$$supp(\sigma) = supp(c_1) \sqcup \cdots \sqcup supp(c_k)\,.$$

Notons d'abord que si $i \in supp(c_1) \subset supp(\sigma)$ alors $\sigma(i) = c_1(i) \in supp(c_1)$, et en itérant on trouve que $\mathcal{O}_\sigma(i) = \mathcal{O}_{c_1}(i) = supp(c_1)$. Supposons qu'on ait une autre telle décomposition $\sigma = d_1 \cdots d_\ell$. Alors il existe $s \in \{1, \ldots, \ell\}$ tel que $i \in supp(d_s)$. Quitte à permuter les facteurs de la décomposition, on peut supposer que $s = 1$ et que $i \in supp(d_1)$. Ainsi $supp(d_1) = \mathcal{O}_\sigma(i) = supp(c_1)$, et $c_1 = \sigma_{|supp(c_1)} = \sigma_{|supp(d_1)} = d_1$. En simplifiant à gauche on obtient une nouvelle permutation $\sigma' = c_2 \cdots c_k = c_1^{-1}\sigma = d_1^{-1}\sigma = d_2 \cdots d_\ell$ qui a deux décompositions, chacune ayant un facteur de moins. On itère le raisonnement pour démontrer que $c_1 = d_1$, $c_2 = d_2$, ..., jusqu'à tomber sur la permutation identité (qui ne peut s'écrire que comme le produit vide, cf. Remarque 1.11). □

Définition 1.17 (ordre)**.** L'**ordre** d'une permutation σ est le plus petit entier $\ell \in \mathbb{N}^*$ tel que $\sigma^\ell = \mathrm{id}$.

Exercice 1.18. *Montrer qu'un tel ℓ existe (on peut utiliser que $\mathfrak{S}_n$ a un nombre fini d'éléments).*

Exercice 1.19. *Soit σ une permutation d'ordre ℓ. On note $\sigma = c_1 \cdots c_m$ sa décomposition en produit de cycles à supports disjoints, et on note ℓ_i l'ordre de c_i. Montrer que $\ell = ppcm(\ell_1, \ldots, \ell_m)$.*

1.3. Signature.

Définition 1.20 (inversion)**.** Soit $\sigma \in \mathfrak{S}_n$. Une **inversion** de σ est une partie à deux éléments $\{i, j\} \subset X = \{1, \ldots, n\}$ telle que $\frac{\sigma(i) - \sigma(j)}{i - j} < 0$.

Autrement dit, $\{i, j\}$ est une inversion de σ si σ change l'ordre de i et j.

Exemple 1.21. $\{1, 2\}$, $\{1, 7\}$ et $\{4, 5\}$ sont des inversions de $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 4 & 5 & 3 & 2 & 1 \end{pmatrix}$, mais pas $\{3, 4\}$. △

Définition 1.22 (signature)**.** Soit $\sigma \in \mathfrak{S}_n$. La **signature** de σ est $\varepsilon(\sigma) := (-1)^{q_\sigma}$, où q_σ est le nombre d'inversions de σ.

Exercice 1.23. *Montrer que la signature de toute transposition vaut -1. [Indication : on pourra montrer que le nombre d'inversions de la transposition $(i\,j)$ est $2|i - j| - 1$.]*

Proposition 1.24 (formule pour la signature). *Pour tout $\sigma \in \mathfrak{S}_n$,*

$$\varepsilon(\sigma) = \prod_{\substack{\{i,j\}\subset X\\ i\neq j}} \frac{\sigma(i)-\sigma(j)}{i-j}$$

DÉMONSTRATION. On note C_X^2 l'ensemble des combinaisons (=parties) de 2 éléments de X, et on pose

$$\begin{array}{rcl} F : C_X^2 & \longrightarrow & \{-1,+1\} \\ \{i,j\} & \longmapsto & \begin{cases} -1 & \text{si } \{i,j\} \text{ est une inversion de } \sigma \\ +1 & \text{sinon} \end{cases} \end{array}$$

Ainsi

$$\varepsilon(\sigma) = \prod_{\{i,j\}\in C_X^2} F(\{i,j\}).$$

On rappelle ensuite que pour tout $x \in \mathbb{R}^*$, $\frac{x}{|x|}$ vaut 1 si $x > 0$ et -1 si $x < 0$. On remarque finalement que $\frac{\sigma(i)-\sigma(j)}{i-j} < 0$ si et seulement si $\{i,j\}$ est une inversion. Par conséquent, $F(\{i,j\}) = \frac{|i-j|\left(\sigma(i)-\sigma(j)\right)}{(i-j)|\sigma(i)-\sigma(j)|}$. On conclut grâce au calcul suivant :

$$\begin{aligned} \prod_{\{i,j\}\in C_X^2} \frac{|i-j|\left(\sigma(i)-\sigma(j)\right)}{(i-j)|\sigma(i)-\sigma(j)|} &= \prod_{\{i,j\}\in C_X^2} \frac{\sigma(i)-\sigma(j)}{i-j} \times \frac{\prod_{\{i,j\}\in C_X^2} |i-j|}{\prod_{\{i,j\}\in C_X^2} |\sigma(i)-\sigma(j)|} \\ &= \prod_{\{i,j\}\in C_X^2} \frac{\sigma(i)-\sigma(j)}{i-j} \times \frac{\prod_{\{i,j\}\in C_X^2} |i-j|}{\prod_{\{k,l\}\in C_X^2} |k-l|} \\ &= \prod_{\{i,j\}\in C_X^2} \frac{\sigma(i)-\sigma(j)}{i-j} \end{aligned}$$

□

Attention : il y a une subtilité importante dans la démonstration qui précède. On peut écrire $\prod_{\{i,j\}\in C_X^2} |i-j|$ car la valeur absolue $|i-j|$ est bien définie pour une partie $\{i,j\} \in C_X^2$. En revanche, la différence $i-j$ ne dépend pas que de $\{i,j\}$ mais aussi de l'ordre entre i et j (il en va de même pour $\sigma(i)-\sigma(j)$) ; néanmoins la fraction $\frac{\sigma(i)-\sigma(j)}{i-j}$ ne dépend plus de l'ordre entre i et j car les variations de signe se compensent. C'est pourquoi le symbole $\prod_{\{i,j\}\in C_X^2}$ doit absolument rester à l'extérieur de la fraction dans le résultat final !

Théorème 1.25 (signature d'un produit). *L'application signature $\varepsilon : \mathfrak{S}_n \to \{-1,+1\}$ est un morphisme de groupe, où la loi de groupe sur $\{-1,+1\}$ est la multiplication usuelle.*

DÉMONSTRATION. Soient $\sigma, \tau \in \mathfrak{S}_n$. Calculons :

$$\begin{aligned} \varepsilon(\sigma)\varepsilon(\tau) &= \prod_{\{i,j\}\in C_X^2} \frac{\sigma(i)-\sigma(j)}{i-j} \prod_{\{i,j\}\in C_X^2} \frac{\tau(i)-\tau(j)}{i-j} \\ &= \prod_{\{i,j\}\in C_X^2} \frac{\sigma(\tau(i))-\sigma(\tau(j))}{\tau(i)-\tau(j)} \prod_{\{i,j\}\in C_X^2} \frac{\tau(i)-\tau(j)}{i-j} \end{aligned}$$

$$= \prod_{\{i,j\}\in C_X^2} \frac{\sigma\tau(i)-\sigma\tau(j)}{i-j} = \varepsilon(\sigma\tau)$$

□

Corollaire 1.26. *La signature d'un cycle de longueur k vaut $(-1)^{k-1}$. Par conséquent, pour tout $\sigma \in \mathfrak{S}_n$, $\varepsilon(\sigma) = (-1)^{p_\sigma}$, où p_σ est le nombre de cycles de longueur paire dans la décomposition de σ en cycles de supports disjoints.*

DÉMONSTRATION. D'après l'exercice 1.23 la signature d'une transposition vaut -1. On remarque qu'un cycle $(i_1\, i_2\, \cdots\, i_k)$ s'écrit facilement comme produit de $k-1$ transpositions :

$$(i_1\, i_2\, \cdots\, i_k) = (i_1\, i_2)(i_2\, i_3)\cdots(i_{k-1}\, i_k)\,.$$

Par conséquent, en utilisant le Théorème 1.25 on trouve que

$$\varepsilon(i_1\, i_2\, \cdots\, i_k) = \varepsilon(i_1\, i_2)\varepsilon(i_2\, i_3)\cdots\varepsilon(i_{k-1}\, i_k) = (-1)^{k-1} = \begin{cases} -1 & \text{si } k \text{ est pair} \\ +1 & \text{sinon} \end{cases}.$$

□

Dans certaines références, les gens notent $(-1)^\sigma := \varepsilon(\sigma)$.

2. Déterminant d'une matrice carrée

Dans toute cette section, on fixe un corps $\mathbb{K}$. On rappelle (HAX202X) que les éléments de $\mathbb{K}$ sont appelés **scalaires**.

Définition 2.1 (déterminant)**.** Soit $A = (a_{ij})_{0\leq i,j\leq n} \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ une matrice carrée de taille $n \times n$ à coefficients dans $\mathbb{K}$. Le **déterminant** de A est le scalaire

$$\det(A) := \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} = \sum_{\sigma\in\mathfrak{S}_n}\left(\varepsilon(\sigma)\prod_{i=1}^n a_{\sigma(i)i}\right)$$

On note aussi $\det(A) = |a_{ij}|_{0\leq i,j\leq n}$.

2.1. Le déterminant en petite dimension.

Pour $n = 2$, il n'y a que deux éléments dans $\mathfrak{S}_2$: l'identité et la transposition (12). Par conséquent,

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

C'est le produit en croix, qui permet de déterminer si deux vecteurs de $\mathbb{K}^2$ sont colinéaires (le produit en croix/déterminant est nul) ou pas (le produit en croix/déterminant est non nul).

Proposition 2.2 (aire algébrique)**.** *Soient $u, v \in \mathbb{R}^2$ deux vecteurs, et $M := (u\, v)$ la matrice formée par ces deux vecteurs en colonne. Alors $\det(M)$ est l'aire algébrique du parallélogramme formé par u et v.*

DÉMONSTRATION. Supposons que $u = \begin{pmatrix} a \\ b\end{pmatrix}$ et $v = \begin{pmatrix} c \\ d\end{pmatrix}$, de sorte que $\det(M) = ad - bc$.

Il faut montrer que $ad - bc$ est l'aire algébrique du parallélogramme formé par u et v. Cette aire algébrique est le nombre réel $aire(u, v)$

- dont la valeur absolue est l'aire usuelle ;
- dont le signe est "+" si $(u; v) \in [0, \pi[$ et "−" si $(u; v) \in [\pi, 2\pi[$.

Exemple 2.3. Si (e_1, e_2) est la base canonique de $\mathbb{R}^2$ alors $aire(e_1, e_2) = 1$ tandis que $aire(e_2, e_1) = -1$. △

Remarque 2.4. Si l'angle vaut 0 ou π, le signe n'a en réalité pas beaucoup d'importance car l'aire est nulle.

Représentons nos deux vecteur sous forme de nombres complexes : $u = a + ib$ et $v = c + id$. On remarque que $ad - bc = \mathrm{Im}(\bar{u}v)$. Passons en forme polaire : $u = r_1 e^{i\theta_1}$ et $v = r_2 e^{i\theta_2}$. Alors $\mathrm{Im}(\bar{u}v) = r_1 r_2 \sin(\theta_2 - \theta_1)$, qui coïncide exactement avec l'aire algébrique du parallélogramme formé par u et v (faites un dessin !). □

Pour $n = 3$, il y a six éléments dans $\mathfrak{S}_3$: l'identité, trois transpositions (12), (23) et (13), et deux 3-cycles[3] (123) et (321). Par conséquent,

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

Remarque 2.5. Dans ce calcul on observe un cas particulier d'un phénomène général que nous verrons plus loin : on peut développer un déterminant selon une colonne.

Remarque 2.6. On reconnait le **produit vectoriel**

$$\begin{pmatrix} \begin{vmatrix} a_{21} & a_{32} \\ a_{31} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{31} & a_{12} \\ a_{11} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} \wedge \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix},$$

de sorte que le déterminant d'une matrice 3×3 est le **produit mixte** de ses trois vecteurs colonnes.

Proposition 2.7 (volume algébrique)**.** *Soient $u, v, w \in \mathbb{R}^3$ trois vecteurs, et $M := (u\, v\, w)$ la matrice formée par ces vecteurs en colonne. Alors $\det(M)$ est le volume algébrique du parallélépipède formé par u, v et w.*

DÉMONSTRATION. Non traitée dans ce cours (bien que ce soit de la géométrie élémentaire, ça n'est pas l'objet de cette UE). □

2.2. Quelques propriétés utiles pour le calcul.

Soit $A = (a_{ij})_{0 \leq i,j \leq n}$ une matrice carrée. On rappelle sa **transposée** $A^{\mathrm{T}} = (b_{ij})_{0 \leq i,j \leq n}$, dont les coefficients sont $b_{ij} = a_{ji}$ (autrement dit, les lignes de la transposée sont les colonnes de la matrice de départ).

Proposition 2.8 (déterminant de la transposée)**.** $\det(A) = \det(A^{\mathrm{T}})$.

3. C'est une abbréviation pour "cycle de longueur 3".

DÉMONSTRATION. On calcule :

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} \\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma^{-1}(1)}a_{2\sigma^{-1}(2)}\cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma^{-1}) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = \det(A^{\mathrm{T}})\end{aligned}$$

Dans l'égalité permettant de passer à la dernière ligne du calcul, nous avons utilisé que $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$. Pouvez-vous le démontrer ? □

On rappelle que la matrice A est dite **triangulaire supérieure** si $a_{ij} = 0$ lorsque $i > j$ (et triangulaire inférieure si $a_{ij} = 0$ lorsque $i < j$).

Proposition 2.9 (déterminant triangulaire)**.** *Si A est triangulaire (supérieure ou inférieure), alors* $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

DÉMONSTRATION. Il suffit de le démontrer pour les matrices triangulaires supérieures [4]. Supposons donc que A est triangulaire supérieure. Dans la définition du déterminant, chaque terme $\varepsilon(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$ est non nul si et seulement si $\sigma(i) \leq i$ quel que soit $i \in \{1,\ldots,n\}$. Or ça n'arrive que si σ est l'identité. Ainsi $\det(A) = a_{11}a_{22}\cdots a_{nn}$. □

On dit qu'une matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ est **diagonale par blocs** si elle est de la forme suivante :

$$M = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$$

avec A une matrice carrée de taille $q \times q$ et D une matrice carrée de taille $r \times r$ (on a nécesairement $n = q + r$). Ça signifie concrètement que les coefficients m_{ij} de M sont nuls dans les deux cas qui suivent :
(a) si $i > q$ et $j \leq q$
(b) si $i \leq q$ et $j > q$.
On remarque que $A = (m_{ij})_{1\leq i,j\leq q}$ et $B = (m_{q+i\,q+j})_{1\leq i,j\leq r}$.

Exercice 2.10. *Démontrer qu'une matrice M est diagonale par blocs de type (q,r), avec $q + r = n$, si et seulement si*

$$M\big(\mathbb{K}^q \times \{(0,\ldots,0)\}\big) \subset \mathbb{K}^q \times \{(0,\ldots,0)\} \quad \textit{et} \quad M\big(\{(0,\ldots,0)\} \times \mathbb{K}^r\big) \subset \{(0,\ldots,0)\} \times \mathbb{K}^r.$$

Dans ce cas, on peut démontrer que $\det(M) = \det(A)\det(D)$. C'est un cas particulier d'un résultat plus général qui concerne les matrices **triangulaires par blocs**. On dit qu'une matrice M est

- triangulaire supérieure par blocs lorque ses coefficients m_{ij} s'annulent dans le cas (a) ci-dessus.
- triangulaire inférieure par blocs lorque ses coefficients m_{ij} s'annulent dans le cas (b) ci-dessus.

4. Le cas des matrices triangulaires inférieures se déduit par passage à la transposée.

En particulier, une matrice triangulaire supérieure par blocs est de la forme

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \tag{1}$$

avec $A \in \mathrm{Mat}_{q\times q}(\mathbb{K})$, $B \in \mathrm{Mat}_{q\times r}(\mathbb{K})$ et $D \in \mathrm{Mat}_{r\times r}(\mathbb{K})$.

Exercice 2.11. *Démontrer qu'une matrice M est triangulaire supérieure par blocs de type (q,r), avec $q+r=n$, si et seulement si*

$$M\big(\mathbb{K}^q \times \{(0,\dots,0)\}\big) \subset \mathbb{K}^q \times \{(0,\dots,0)\}\,.$$

Proposition 2.12 (déterminant triangulaire par blocs)**.** *Soit M une matrice triangulaire supérieure par bloc comme dans l'équation* (1)*. Alors* $\det(M) = \det(A)\det(D)$.

DÉMONSTRATION. Commençons par observer que si $(\alpha,\beta) \in \mathfrak{S}_q \times \mathfrak{S}_r$ alors on a la permutation de $n = q+r$ éléments qui suit :

$$\alpha \sqcup \beta := \begin{pmatrix} 1 & \cdots & q & q+1 & \cdots & q+r \\ \alpha(1) & \cdots & \alpha(q) & q+\beta(1) & \cdots & q+\beta(r) \end{pmatrix}.$$

Démontrez en exercice que $\varepsilon(\alpha \sqcup \beta) = \varepsilon(\alpha)\varepsilon(\beta)$. Les permutations σ de cette forme sont exactement celles qui satisfont

$$\sigma\big(\{1,\dots,q\}\big) = \{1,\dots,q\} \qquad \text{et} \qquad \sigma\big(\{q+1,\dots,n\}\big) = \{q+1,\dots,n\}\,.$$

On remarque que si σ n'est pas de cette forme, alors il existe $i \le q$ tel que $\sigma(i) > q$, de sorte que $m_{\sigma(i)i} = 0$. Par conséquent,

$$\begin{aligned} \det(M) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\left(\prod_{i=1}^n m_{\sigma(i)i}\right) = \sum_{\substack{\alpha\in\mathfrak{S}_q\\ \beta\in\mathfrak{S}_r}} \varepsilon(\alpha\sqcup\beta)\left(\prod_{i=1}^q m_{\alpha(i)i}\right)\left(\prod_{k=1}^r m_{q+\beta(k)\,q+k}\right) \\ &= \left(\sum_{\alpha\in\mathfrak{S}_q} \varepsilon(\alpha)\left(\prod_{i=1}^q m_{\alpha(i)i}\right)\right)\cdot\left(\sum_{\beta\in\mathfrak{S}_q} \varepsilon(\beta)\left(\prod_{k=1}^r m_{q+\beta(k)\,q+k}\right)\right) \\ &= \det(A)\det(B)\,. \end{aligned}$$

□

2.3. Deux exemples instructifs : matrices de vecteurs canoniques.

Dans cette partie on calcule le déterminant de matrices dont les colonnes sont des vecteurs de la base canonique.

Exemple 2.13 (déterminant d'une matrice de permutation)**.** On rappelle (HAX202X) qu'à toute permutation $\tau \in \mathfrak{S}_n$ on peut associer la matrice de permutation $P_\tau = (p_{ij})_{0\le i,j\le n}$ dont les coefficients sont définis comme suit[5] : $p_{ij} = \delta_{i\tau(j)}$. En particulier, si $(e_1,\dots,e_n)$ est la base canonique de $\mathbb{K}^n$, alors

$$P_\tau \cdot e_j = \begin{pmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{pmatrix} = e_{\tau(j)}\,.$$

5. On utilise le symbole de Kronecker δ_{kl}, qui vaut 1 si $k = l$ et 0 sinon.

Calculons le déterminant de P_τ :

$$\begin{aligned}\det(P_\tau) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) p_{\sigma(1)1}p_{\sigma(2)2}\cdots p_{\sigma(n)n}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\delta_{\sigma(1)\tau(1)}\delta_{\sigma(2)\tau(2)}\cdots\delta_{\sigma(n)\tau(n)} = \varepsilon(\tau)\,.\end{aligned}$$

△

En particulier, le déterminant de la matrice identité vaut 1.

Exemple 2.14. Considérons une fonction $h : \{1,\ldots,n\} \to \{1,\ldots,n\}$ qui n'est **pas** une bijection. En particulier, il existe $i \neq j$ tel que $h(i) = h(j)$. Considérons la matrice $M_h = (m_{ij})_{0\leq i,j\leq n}$ dont les coefficients sont définis comme suit : $m_{ij} = \delta_{ih(j)}$. Ici encore, si $(e_1,\ldots,e_n)$ est la base canonique de $\mathbb{K}^n$, alors

$$M_h \cdot e_j = \begin{pmatrix} m_{1j}\\ m_{2j}\\ \vdots\\ m_{nj}\end{pmatrix} = e_{h(j)}\,.$$

Calculons le déterminant de M_h :

$$\begin{aligned}\det(M_h) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) m_{\sigma(1)1}m_{\sigma(2)2}\cdots m_{\sigma(n)n}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\delta_{\sigma(1)h(1)}\delta_{\sigma(2)h(2)}\cdots\delta_{\sigma(n)h(n)} = 0\,.\end{aligned}$$

△

Exercice 2.15. *Pour chacun des deux exemples qui précèdent, effectuez quelques calculs explicites avec $n = 2$, $n = 3$, et éventuellement $n = 4$.*

3. Formes alternées

3.1. Définition abstraite.

Soit E un $\mathbb{K}$-espace vectoriel et soit n un entier naturel.

Définition 3.1 (forme n-linéaire alternée)**.** Une **forme n-linéaire alternée** est une application $f : E^n \to \mathbb{K}$ telle que

— f est **linéaire en chaque argument**[6] : pour tout $i \in \{1,\ldots,n\}$, quels que soient $u_1,\ldots,$ $u_{i-1}, u_{i+1},\ldots,u_n \in E$, l'application partielle $E \to \mathbb{K}$ définie par

$$\mathbf{u} \mapsto f(u_1,\ldots,u_{i-1},\mathbf{u},u_{i+1},\ldots,u_n)$$

est linéaire ;

— f est **alternée** : pour tous $u_1,\ldots,u_n \in E$, si il existe $i \neq j$ tels que $u_i = u_j$ alors $f(u_1,\ldots,u_n) = 0$.

On note $\mathrm{Alt}_n(E)$ l'ensemble des formes n-linéaires alternées de E.

Proposition 3.2 (antisymétrie des formes alternées)**.** *Toute forme alternée $f \in \mathrm{Alt}_n(E)$ est **antisymétrique** : pour tous $u_1,\ldots,u_n \in E$,*

$$f(u_{\sigma(1)},\ldots,u_{\sigma(n)}) = \epsilon(\sigma) f(u_1,\ldots,u_n)\,.$$

6. On dit qu'elle est **multi-linéaire**, ou **n-linéaire**.

DÉMONSTRATION. Commençons par montrer que le lemme est vérifié pour les transpositions. Soit $\tau = (ij)$ une transposition (on peut supposer que $i < j$). Comme f est alternée on a

$$\begin{aligned} f(u_1,\dots,u_n) + f(u_{\tau(1)},\dots,u_{\tau(n)}) &= f(u_1,\dots,u_i,\dots,u_j,\dots,u_n) + f(u_1,\dots,u_j,\dots,u_i,\dots,u_n) \\ &= f(u_1,\dots,u_i,\dots,u_j,\dots,u_n) + f(u_1,\dots,u_j,\dots,u_i,\dots,u_n) \\ &\quad + f(u_1,\dots,u_i,\dots,u_i,\dots,u_n) + f(u_1,\dots,u_j,\dots,u_j,\dots,u_n) \\ &= f(u_1,\dots,u_i+u_j,\dots,u_i+u_j,\dots,u_n) = 0 \end{aligned}$$

Donc $f(u_{\tau(1)},\dots,u_{\tau(n)}) = -f(u_1,\dots,u_n) = \varepsilon(\tau) f(u_1,\dots,u_n)$.
Si $\sigma \in \mathfrak{S}_n$ est une permutation, alors on a une décomposition (non unique !) en produit de transpositions $\sigma = \tau_1 \cdots \tau_m$. Par conséquent,

$$\begin{aligned} f(u_{\sigma(1)},\dots,u_{\sigma(n)}) &= f(u_{\tau_1\cdots\tau_m(1)},\dots,u_{\tau_1\cdots\tau_m(n)}) \\ &= \varepsilon(\tau_m) f(u_{\tau_1\cdots\tau_{m-1}(1)},\dots,u_{\tau_1\cdots\tau_{m-1}(n)}) \\ &\cdots \\ &= \varepsilon(\tau_m)\cdots\varepsilon(\tau_1) f(u_1,\dots,u_n) = \varepsilon(\sigma) f(u_1,\dots,u_n)\,. \end{aligned}$$

□

Remarque 3.3. Si la caractéristique de $\mathbb{K}$ est différente de 2, alors on peut démontrer une réciproque de la proposition précédente : une application $f : E^n \to \mathbb{K}$ qui est linéaire en chaque argument est alternée si et seulement si elle est antisymétrique. Si vous êtes très à l'aise, vous pouvez le démontrer en exercice. Ça repose sur le fait que $x = -x$ implique $x = 0$ lorsque la caractéristique est différente de 2.

Exercice 3.4. *Montrer que l'ensemble* $\mathrm{Alt}_n(E)$ *des formes linéaires n-alternées est un sous-espace vectoriel de l'espace vectoriel* $\mathbb{K}^E$ *des applications de E dans* $\mathbb{K}$ *(autrement dit : toute combinaison linéaire de formes n-linéaires alternées est une forme n-linéaire alternée).*

Exemple 3.5 (une forme bilinéaire alternée)**.** Soit $E = \mathbb{K}^n \times \mathbb{K}^n$. Un élément de E est un couple (u, u') de vecteurs de $\mathbb{K}^n$. On définit $\omega : E^2 \to \mathbb{K}$ de la manière suivante :

$$\omega\big((u,u'),(v,v')\big) := \langle u, v'\rangle - \langle v, u'\rangle\,.$$

Vous pouvez vérifier en exercice que ω est linéaire chaque argument (u, u'). Montrons maintenant qu'elle est alternée :

$$\omega\big((u,u'),(u,u')\big) = \langle u, u'\rangle - \langle u, u'\rangle = 0\,.$$

△

Dans la suite on va uniquement s'intéresser au cas où E est de dimension finie et où $n = \dim(E)$. Fixons une base $\mathcal{B} = (v_1,\dots,v_n)$ de E.

Proposition 3.6 (l'espace des formes alternées est de dimension 1)**.** *L'application* $\mathrm{Alt}_n(E) \to \mathbb{K}$ *définie par* $f \mapsto f(v_1,\dots,v_n)$ *est un isomorphisme d'espaces vectoriels.*

DÉMONSTRATION. Vous pouvez aisément vérifier en exercice que cette application est linéaire. Pour démontrer qu'elle est bijective, on commence par observer qu'une application $f : E^n \to \mathbb{K}$ qui est linéaire en chaque argument est uniquement déterminée par les scalaires $f_{i_1,\dots,i_n} := f(v_{i_1},\dots,v_{i_n})$ avec $i_1,\dots,i_n \in \{1,\dots,n\}$.

LEMME 3.7. *Une forme n-linéaire f est de plus alternée si et seulement si les deux conditions qui suivent sont satisfaites :*

(a) si un indice se répète (c'est-à-dire que $i_s = i_t$ avec $s \neq t$), alors $f_{i_1,\dots,i_n} = 0$.
(b) si tous les indices sont distincts, alors on a une permutation

$$\sigma := \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

et $f_{i_1,\dots,i_n} = \varepsilon(\sigma) f_{1,\dots,n}$.

Par conséquent, on en déduit qu'une forme n-linéaire alternée f est uniquement définie par $f_{1,\dots,n} = f(v_1, \dots, v_n)$. □

DÉMONSTRATION DU LEMME. Soit f une forme n-linéaire.
$\boxed{\Rightarrow}$ C'est le sens facile de la démonstration. On suppose que f est alternée. Ainsi, par définition, $f_{i_1,\dots,i_n} = f(v_{i_1}, \dots, v_{i_n}) = 0$ si un indice (et donc un vecteur) se répète. La condition (a) est donc vérifiée. Ensuite, toute forme alternée est antisymétrique (c'est la Proposition 3.2), de sorte que la condition (b) est également vérifiée.
$\boxed{\Leftarrow}$ La réciproque est un peu plus délicate. On suppose que les conditions (a) et (b) sont vérifiées. On doit démontrer que f est alternée, c'est-à-dire que si $(u_1, \dots, u_n)$ est une famille de vecteurs dans laquelle un vecteur se répète alors $f(u_1, \dots, u_n) = 0$. Nous allons faire la démonstration dans le cas où $u_1 = u_2 = u$, c'est-à-dire que la famille est $(u, u, u_3, \dots, u_n)$[7]. Par linéarité de f en chaque variable, il suffit de le démontrer lorsque $u_3, \dots, u_n$ sont des vecteurs de la base : $u_s = v_{i_s}$ pour $3 \leq s \leq n$. On écrit alors $u = \sum_{j=1}^n \lambda_j v_j$ et on calcule :

$$\begin{aligned}
f(u, u, v_{i_3}, \dots, v_{i_n}) &= \sum_{i_1, i_2 = 1}^{n} \lambda_j \lambda_k f(v_{i_1}, v_{i_2}, v_{i_3}, \dots, v_{i_n}) \\
&= \sum_{1 \leq i_1 < i_2 \leq n} \lambda_{i_1} \lambda_{i_2} f_{i_1, i_2, i_3, \dots, i_n} + \sum_{1 \leq i_2 < i_1 \leq n} \lambda_{i_1} \lambda_{i_2} f_{i_1, i_2, i_3, \dots, i_n} + \sum_{i=1}^{n} \lambda_i^2 \underbrace{f_{i,i,i_3,\dots,i_n}}_{=0 \text{ d'après (a)}} \\
&= \sum_{1 \leq i_1 < i_2 \leq n} \lambda_{i_1} \lambda_{i_2} \underbrace{(f_{i_1, i_2, i_3, \dots, i_n} + f_{i_2, i_1, i_3, \dots, i_n})}_{=0 \text{ d'après (b)}} = 0
\end{aligned}$$

□

3.2. Le déterminant est une forme alternée.
Soit $E = \mathbb{K}^n$. Considérons l'application composée

$$E^n \xrightarrow{\sim} \mathrm{Mat}_{n \times n}(\mathbb{K}) \xrightarrow{\det} \mathbb{K}$$

où la première application envoie une famille de vecteurs $u = (u_1, \dots, u_n)$ sur sa matrice des vecteurs colonnes $M_u = (u_1 \cdots u_n)$.

Théorème 3.8 (le déterminant est une forme alternée)**.** *Cette application, définie par* $(u_1, \dots, u_n) \mapsto \det(u_1 \cdots u_n)$*, est une forme n-linéaire alternée.*

DÉMONSTRATION. Notons u_{ij} les coordonnées du vecteur u_j. Autrement dit, u_{ij} est le coefficient (i, j) de la matrice M_u. Montrons dans un premier temps que l'application est linéaire en la première variable : si $u_1 = v_1 + \lambda w_1$, alors $u_{i1} = v_{i1} + \lambda w_{i1}$ quel que soit i, et ainsi

$$\det(u_1 \cdots u_n) = \sum_{\sigma \in \mathfrak{S}_n} \left(\varepsilon(\sigma) \prod_{i=1}^{n} u_{\sigma(i)i} \right)$$

7. La démonstration est la même en général ; c'est simplement plus aisé à comprendre quand les deux vecteurs identiques sont au début

$$\begin{aligned}
&= \sum_{\sigma\in\mathfrak{S}_n}\left(\varepsilon(\sigma)u_{\sigma(1)1}\prod_{i=2}^n u_{\sigma(i)i}\right)\\
&= \sum_{\sigma\in\mathfrak{S}_n}\left(\varepsilon(\sigma)(v_{\sigma(1)1}+\lambda w_{\sigma(1)1})\prod_{i=2}^n u_{\sigma(i)i}\right)\\
&= \sum_{\sigma\in\mathfrak{S}_n}\left(\varepsilon(\sigma)v_{\sigma(1)1}\prod_{i=2}^n u_{\sigma(i)i}\right)+\lambda\sum_{\sigma\in\mathfrak{S}_n}\left(\varepsilon(\sigma)w_{\sigma(1)1}\prod_{i=2}^n u_{\sigma(i)i}\right)\\
&= \det(v_1\,u_2\,\cdots\,u_n)+\lambda\det(w_1\,u_2\,\cdots\,u_n)\,.
\end{aligned}$$

Le raisonnement pour démontrer la linéarité dans chacune des autres variables est le même. Pour voir que l'application est alternée, il faut (et il suffit de) vérifier les conditions (a) et (b) du Lemme 3.7 avec la base canonique $(e_1,\dots,e_n)$. La condition (b) est garantie par l'Exemple 2.13 et la condition (a) par l'Exemple 2.14. □

Remarque 3.9. Nous avons démontré que le déterminant est une forme multilinéaire alternée des vecteurs colonnes. Mais comme le déterminant reste le même après passage à la transposée (c'est la Proposition 2.8), il est également une forme multilinéaire alternée des vecteurs lignes.

Exercice 3.10. *Montrer que si f est une forme multilinéaire alternée, alors $f(u_1,\dots,u_n)=0$ pour toute famille liée $(u_1,\dots,u_n)$. En déduire que le déterminant d'une matrice s'annule lorsqu'il existe une relation de dépendance linéaire non triviale entre ses colonnes (ou entre ses lignes, d'après la Remarque 3.9).*

Corollaire 3.11 (unicité du déterminant)**.** *L'application $(u_1,\dots,u_n)\mapsto\det(u_1\,\cdots\,u_n)$ est l'unique forme n-linéaire alternée $f\in\mathrm{Alt}_n(E)$ telle que $f(e_1,\dots,e_n)=1$. Par conséquent, pour tout $f\in\mathrm{Alt}_n(E)$, $f(u_1,\dots,u_n)=f(e_1,\dots,e_n)\det(u_1\,\cdots\,u_n)$ quels que soient $u_1,\dots,u_n\in E$.*

Démonstration. D'après la Proposition 3.6, quel que soit $\lambda\in\mathbb{K}$ il existe un unique $f_\lambda\in\mathrm{Alt}_n(f)$ tel que $f_\lambda(e_1,\dots,e_n)=\lambda$, et si $\lambda,\mu\in\mathbb{K}$ alors $\mu f_\lambda=\lambda f_\mu$. Or d'après l'Exemple 2.13, $\det(e_1\,\cdots\,e_n)=1$. □

3.3. Déterminant d'un endomorphisme.

Théorème 3.12 (déterminant d'un produit)**.** *Soient M et N deux matrices carrées de taille $n\times n$. Alors $\det(MN)=\det(M)\det(N)$.*

Démonstration. Fixons M, posons $E=\mathbb{K}^n$ et considérons l'application $g:E^n\to\mathbb{K}$ définie par $(u_1,\dots,u_n)\mapsto\det(MN)$, où $N:=(u_1\,\cdots\,u_n)$.

Lemme 3.13. *L'application g est une forme n-linéaire alternée.*

Démonstration. Commençons par la multi-linéarité, et supposons que la j-ième colonne de N est une combinaison linéaire $u_j=v_j+\lambda w_j$. On note N' (resp. N'') la matrice obtenue en remplaçant dans N la j-ième colonne u_j par v_j (resp. par w_j). On rappelle que la j-ième colonne C_j du produit $C=MN$ est donnée par le produit de M avec le vecteur colonne u_j : $C_j=Mu_j=Mv_j+\lambda Mw_j$. On note C' (resp. C'') la matrice obtenue en remplaçant dans C la j-ième colonne $C_j=Mu_j$ par Mv_j (resp. par Mw_j). Par multi-linéarité du déterminant dans les colonnes, on trouve que

$$\det(MN)=\det(C)=\det(C')+\lambda\det(C'')=\det(MN')+\lambda\det(MN'')\,,$$

c'est-à-dire que $g(u_1,\dots,u_n)=g(u_1,\dots,v_i,\dots,u_n)+\lambda g(u_1,\dots,w_i,\dots,u_n)$.

Montrons maintenant que g est alternée. Si deux colonnes de N sont égales, c'est-à-dire que $u_i = u_j$ pour $i \neq j$, alors les colonnes correspondantes de $C = MN$ sont égales aussi : $C_i = C_j$. Ainsi $\det(MN) = \det(C) = 0$. □

Par conséquent $g(u_1, \ldots, u_n) = g(e_1, \ldots, e_n)\det(u_1 \cdots u_n)$, d'après le Corollaire 3.11. Ainsi $\det(MN) = \det(M)\det(N)$. □

Corollaire 3.14 (invariance par conjugaison du déterminant)**.** *Si* $P \in GL_n(\mathbb{K})$ *et* $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$, *alors* $\det(P) \in \mathbb{K}^\times$ *et* $\det(P^{-1}MP) = \det(M)$.

DÉMONSTRATION. À faire en exercice. □

Une conséquence directe de ce Corollaire est que si E est un espace vectoriel de dimension n et $\phi \in \mathrm{End}(E)$ alors le **déterminant de la matrice** $M_{\mathcal{B},\mathcal{B}}(\phi)$ **est indépendant du choix d'une base** $\mathcal{B}$ **de** E. On a donc une application bien définie

$$\det \,:\, \mathrm{End}(E) \longrightarrow \mathbb{K}\,.$$

Exercice 3.15. *Soit* $\phi : \mathbb{K}[X]_{\leq 3} \to \mathbb{K}[X]_{\leq 3}$ *l'application linéaire donnée par la dérivée. Calculer* $\det(\phi)$.

4. Aspects calculatoires

Dans cette partie nous explorons plusieurs conséquences calculatoires intéressantes du fait que le déterminant est une forme multi-linéaire alternée, à la fois dans les lignes et dans les colonnes.

4.1. Développement suivant une colonne/ligne.

Soit $A = (a_{ij})_{1\leq i,j\leq n}$ une matrice carrée. Quels que soient deux indices i et j, on note $A^{i,j}$ la matrice obtenue à partir de A en retirant la i-ième ligne et la j-ième colonne : $A^{i,j} \in \mathrm{Mat}_{n-1,n-1}(\mathbb{K})$. Autrement dit, le coefficient (k,ℓ) de la matrice $A^{i,j}$ est
- $a_{k\,\ell}$ si $k < i$ et $\ell < j$;
- $a_{k+1\,\ell}$ si $k \geq i$ et $\ell < j$;
- $a_{k\,\ell+1}$ si $k < i$ et $\ell \geq j$.
- $a_{k+1\,\ell+1}$ si $k \geq i$ et $\ell \geq j$.

Proposition 4.1 (développement selon une colonne)**.** *Pour tout* $j \in \{1, \ldots, n\}$,

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

DÉMONSTRATION. On note $C_j = \begin{pmatrix} a_{1j} \\ \cdots \\ a_{nj} \end{pmatrix}$ la j-ième colonne de A. Par linéarité en la j-ième colonne, on trouve que

$$\det(A) = \sum_{i=1}^{n} a_{ij}\det(C_1 \,\cdots C_{j-1}\, e_i\, C_{j+1}\, \cdots\, C_n)\,.$$

Par antisymétrie, en échangeant d'abord les colonnes numéros 1 et j puis les lignes numéros 1 et i de la matrice $(C_1 \,\cdots C_{j-1}\, e_i\, C_{j+1}\, \cdots\, C_n)$, on trouve

$$\det(C_1 \,\cdots C_{j-1}\, e_i\, C_{j+1}\, \cdots\, C_n) \quad = \quad (-1)^{j-1}(-1)^{i-1}\begin{vmatrix} 1 & * \\ 0 & A^{i,j} \end{vmatrix}$$

$$= \quad (-1)^{i+j}\det(A^{i,j})$$

Pour la seconde égalité, nous avons utilisé la Proposition 2.9 qui calcule le déterminant d'une matrice triangulaire par blocs. □

Remarque 4.2. On peut aussi développer le déterminant suivant une ligne (cela découle l'invariance du déterminant par passage à la transposée). À titre d'exercice, écrivez la formule du développement du déterminant selon la i-ième ligne.

Exemple 4.3.

$$\begin{aligned}
\begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 2 & 3 & 2 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{vmatrix} &= 2 \cdot \begin{vmatrix} 2 & 3 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 3 & 0 \\ 2 & 3 & 2 \\ 1 & 1 & 0 \end{vmatrix} \\
&= 2 \cdot 1 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} - 1 \cdot (-2) \cdot \begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix} \\
&= 2 \cdot (2 \cdot 1 - 1 \cdot 3) + 2 \cdot (1 \cdot 1 - 1 \cdot 3) \\
&= -6
\end{aligned}$$

△

Soient $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ deux vecteurs de $\mathbb{K}^n$

Corollaire 4.4 (formule de Cramer)**.** *Si $AX = Y$ alors*

$$\det(A)x_j = \det(C_1 \, \cdots \, C_{j-1} \, Y \, C_{j+1} \, \cdots \, C_n)$$

quel que soit $1 \leq j \leq n$.

Démonstration. En développant selon la j-ième colonne on trouve

$$\begin{aligned}
\left|C_1 \, \cdots \, C_{j-1} \, Y \, C_{j+1} \, \cdots \, C_n\right| &= \sum_{i=1}^{n} (-1)^{i+j} y_i |A^{i,j}| \\
&= \sum_{i=1}^{n} (-1)^{i+j} \underbrace{\left(\sum_{k=1}^{n} a_{ik} x_k \right)}_{=y_i} |A^{i,j}| \\
&= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} (-1)^{i+j} a_{ik} |A^{i,j}| \right) x_k
\end{aligned}$$

Or, toujours en développant selon la j-ième ligne, on trouve

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ik} |A^{i,j}| = \left|C_1 \, \cdots \, C_{j-1} \, C_k \, C_{j+1} \, \cdots \, C_n\right| = \begin{cases} |A| & \text{si } j = k \\ 0 & \text{sinon} \end{cases}$$

On trouve donc bien $\left|C_1 \, \cdots \, C_{j-1} \, Y \, C_{j+1} \, \cdots \, C_n\right| = |A| x_j$. □

Remarque 4.5 (critère d'inversibilité des matrices)**.** Si $|A| \in \mathbb{K}^\times$ alors on obtient la formule $x_j = \frac{|C_1 \cdots C_{j-1}\, Y\, C_{j+1} \cdots C_n|}{|A|}$, qui nous donne une unique solution X du système linéaire inhomogène $AX = Y$ quel que soit Y, ce qui signifie que A est inversible. On a donc que A est inversible si et seulement si son déterminant l'est [8].

Exemple 4.6. Considérons le cas où $n = 2$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{K})$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ est un vecteur d'inconnues et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{K}^2$ est fixé. On suppose que $|A| = ad - cb \in \mathbb{K}^\times$, c'est-à-dire que A est inversible. L'unique solution du système linéaire inhomogène

$$AX = Y \Leftrightarrow \begin{cases} ax_1 + bx_2 = y_1 \\ cx_1 + dx_2 = y_2 \end{cases}$$

est donnée par les formules qui suivent :

$$x_1 = (ad - cb)^{-1}(y_1 d - y_2 b) \quad \text{et} \quad x_2 = (ad - cb)^{-1}(ay_2 - cy_1)\,.$$

Vérifier en exercice que ces formules donnent bien une solution du système. △

4.2. Comatrice.

On a vu dans la partie précédente (dont on conserve les notations) que les scalaires $(-1)^{i+j}|A^{i,j}|$ jouent un rôle important dans les calculs. on peut les voir comme les coefficients d'une nouvelle matrice, qu'on appelle la **comatrice** de A.

Définition 4.7 (comatrice)**.** La **comatrice** de A est la matrice $\mathrm{com}(A) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ dont le coefficient (i, j) est $(-1)^{i+j}|A^{i,j}|$.

UN PEU DE VOCABULAIRE : la comatrice est parfois aussi appelée **matrice adjointe** [a] ou **matrice des cofacteurs** ; les cofacteurs sont les scalaires $(-1)^{i+j}|A^{i,j}|$. À un signe près, ces cofacteurs sont les **mineurs** $|A^{i,j}|$ de la matrice A.

a. Surtout par les anglophones ; en français on évitera de l'utiliser car ce terme désigne aussi la matrice transconjuguée, qui n'a rien à voir.

Proposition 4.8. $\mathrm{com}(A)^{\mathrm{T}} \cdot A = \det(A)\mathrm{I}_n = A \cdot \mathrm{com}(A)^{\mathrm{T}}$.

DÉMONSTRATION. Le coefficient (i, j) de $\mathrm{com}(A)^{\mathrm{T}} \cdot A$ est

$$\sum_{k=1}^{n} (-1)^{i+k}|A^{k,i}|a_{kj} = \left|C_1 \cdots C_{i-1}\, C_j\, C_{i+1} \cdots C_n\right| = \begin{cases} \det(A) & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

On obtient bien $\mathrm{com}(A)^{\mathrm{T}} \cdot A = \det(A)\mathrm{I}_n$. L'autre égalité se démontre de la même manière □

Ainsi, lorsque la matrice A est inversible, on obtient une formule pour son inverse, dite **formule de Laplace** :

$$A^{-1} = |A|^{-1}\mathrm{com}(A)^{\mathrm{T}}\,.$$

8. Comme $\mathbb{K}$ est un corps, être inversible signifie être différent de 0. Mais en réalité nous n'avons jusqu'à présent utilisé que les propriétés d'anneau commutatif de $\mathbb{K}$, et tous les énoncés fonctionnent dans cette généralité. Ainsi, par exemple, une matrice à coefficients dans $\mathbb{Z}$ admet un inverse **à coefficients entiers** si et seulement si son déterminant est inversible dans $\mathbb{Z}$, c'est-à-dire vaut ± 1. Exemple : la matrice $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ est inversible dans $\mathrm{Mat}_{2\times 2}(\mathbb{Q})$ mais pas dans $\mathrm{Mat}_{2\times 2}(\mathbb{Z})$, car son déterminant est 3 (qui est inversible dans $\mathbb{Q}$ mais pas dans $\mathbb{Z}$).

Exemple 4.9. Dans le cas d'une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de taille 2×2, la comatrice vous est familière :
$$\mathrm{com}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$
Ainsi, si A est inversible alors (**n'oubliez pas la transposée !**)
$$A^{-1} = (ad - bc)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
△

4.3. Le retour du pivot de Gauss.
Nous terminons ce chapitre avec le calcul du déterminant à l'aide de l'algorithme de Gauss (encore lui !). Commençons avec quelques rappels à propos des opérations élémentaires sur les lignes $L_1, \ldots, L_n$ d'une matrice M :

— On peut permuter des lignes. Ça revient à multiplier M **à gauche** par une matrice de permutation P_σ.
EFFET SUR LE DÉTERMINANT : le déterminant de M est multiplié par la signature $\varepsilon(\sigma)$ de la permutation. Notez que, dans la pratique, on se contente d'échange deux lignes L_i et L_j, c'est-à-dire qu'on applique une transposition (ij), dont la signature est -1. Donc, très concrètement, à chaque fois qu'on échange deux lignes le déterminant est multiplié par -1.

— On peut ajouter un multiple d'une ligne à une autre ligne, c'est-à-dire remplacer L_i par $L_i + \lambda L_j$ ($i \neq j$ et $\lambda \in \mathbb{K}$). Ça revient à multiplier M **à gauche** par la matrice $\mathrm{I}_n + \lambda E_{ij}$, où E_{ij} est la matrice dont le coefficient (i, j) est le seul non nul et vaut 1.
EFFET SUR LE DÉTERMINANT : le déterminant de M ne change pas. En effet, vous pouvez vérifier en exercice que $\det(\mathrm{I}_n + E_{ij}) = 1.$.

— On peut enfin multiplier une ligne par un scalaire inversible $\lambda \in \mathbb{K}^\times$. Ça revient à multiplier M **à gauche** par une matrice diagonale de la forme $diag(1, \ldots 1, \lambda, 1, \ldots, 1)$.
EFFET SUR LE DÉTERMINANT : le déterminant est multiplié par λ. Notez que, dans la pratique, on n'utilisera pas cette opération, qui sert uniquement si on souhaite avoir des pivots égaux à 1.

Si on se contente des opérations élémentaires des deux premiers types, on peut appliquer notre algorithme de Gauss en comptant le nombre q de fois où on a échangé deux lignes. Si on note N la matrice échelonnée à laquelle l'algorithme aboutit, on trouve
$$\det(M) = (-1)^q \det(N).$$
Le déterminant d'une matrice échelonnée carrée est très facile à calculer, car celle-ci est nécessairement triangulaire supérieure. Ainsi, le déterminant de N est le produit de ses éléments diagonaux. On distingue deux situations :

(1) si il y a un pivot par ligne (et donc par colonne, puisque la matrice est carrée) : le déterminant de la matrice échelonnée N est le produit de ses pivots.
(2) sinon le déterminant est nul.

Exemple 4.10. Reprenons la matrice de l'Exemple 4.3 :
$$\begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 2 & 3 & 2 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 3 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 3 & 0 \end{vmatrix} \qquad (L_1 \leftrightarrow L_4)$$

$$
\begin{aligned}
&= \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 3 & 2 \\ 2 & 1 & 3 & 0 \end{vmatrix} \qquad (L_2 \leftrightarrow L_3) \\
&= \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 \end{vmatrix} \qquad (L_4 \mapsto L_4 - 2L_1) \\
&= \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & -2 \end{vmatrix} \qquad \left(\begin{smallmatrix} L_3 \mapsto L_3 - 2L_2 \\ L_4 \mapsto L_4 - L_2 \end{smallmatrix}\right) \\
&= \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -6 \end{vmatrix} = -6
\end{aligned}
$$

Ouf, on trouve la même chose ! $\triangle$

Chapitre 2

Spectre, polynôme caractéristique et réduction

Soit E un $\mathbb{K}$-espace vectoriel.

1. Valeur/vecteur/espace propre et spectre

Soit $f \in \mathrm{End}(E)$ un endomorphisme de E.

1.1. Définitions et exemples.

Définition 1.1 (vocabulaire propre)**.**

— Un **vecteur propre** de f est un vecteur non nul $v \in E \backslash \{0_E\}$ tel qu'il existe $\lambda \in \mathbb{K}$ satisfaisant $f(v) = \lambda v$.

— Une **valeur propre** de f est un scalaire $\lambda \in \mathbb{K}$ tel qu'il existe un vecteur non nul $v \in E \backslash \{0_E\}$ satisfaisant $f(v) = \lambda v$.

Dans les situations qui précèdent, on dit que λ est la valeur propre associée au vecteur propre v, et que v est le vecteur propre associé à la valeur propre λ.

— Le **spectre** $\mathrm{Sp}(f) \subset \mathbb{K}$ de f est l'ensemble de ses valeurs propres.

— L'**espace propre** $E_\lambda(f)$ associé à $\lambda \in \mathrm{Sp}(f)$ est

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

Si $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$, on parle de valeur propre, vecteur propre, spectre et espace propre de la matrice A lorsque $E = \mathbb{K}^n$ et l'endomorphisme f est défini par $f(v) := A \cdot v$.

Exemple 1.2. Considérons la matrice $A := \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$. Pour tout scalaire $\lambda \in \mathbb{R}$ et tout vecteur $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$, on a :

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} 2a + b = \lambda a \\ 2b = \lambda b \\ 3c = \lambda c \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} (2-\lambda)a + b = 0 \\ (2-\lambda)b = 0 \\ (3-\lambda)c = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} & \lambda = 3 \text{ et } b = a = 0 \\ \text{ou} & \lambda = 2 \text{ et } c = b = 0 \\ \text{ou} & c = b = a = 0 . \end{cases}$$

Donc les valeurs propres de A sont 2 et 3, et leurs espaces propres respectifs sont

$$E_2(A) = \mathbb{R}\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \text{et} \quad E_3(A) = \mathbb{R}\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

△

Remarque 1.3. Il est important de bien rappeler le corps $\mathbb{K}$ sur lequel on travaille. Suivant qu'on est par exemple sur $\mathbb{R}$ ou $\mathbb{C}$, la matrice $M := \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$ n'a pas les mêmes valeurs propres. En effet, $\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \lambda\begin{pmatrix}a\\b\end{pmatrix}$ si et seulement si $-b = \lambda a$ et $a = \lambda b$. En exigeant que $(a,b) \neq (0,0)$, on obtient que λ est une valeur propre de M si et seulement si $\lambda^2 = -1$. Par conséquent :

- sur $\mathbb{R}$, M n'a pas de valeur propre ;
- sur $\mathbb{C}$, M a deux valeurs propres qui sont i et $-i$, ayant pour sous-espaces propres
$$E_i(M) = \mathbb{C}\begin{pmatrix}i\\1\end{pmatrix} \quad \text{et} \quad E_{-i}(M) = \mathbb{C}\begin{pmatrix}1\\i\end{pmatrix};$$
- sur $\mathbb{F}_2$, M n'a qu'une valeur propre qui est 1, ayant pour sous-espace propre $E_1(M) = \mathbb{F}_2\begin{pmatrix}1\\1\end{pmatrix}$.

Exemples 1.4 (exemples en dimension infinie)**.** Soit $E = \mathbb{K}[X]$ l'espace vectoriel des polyômes.

- L'application dérivée n'a qu'une seule valeur propre, qui est 0 (l'espace propre associé est l'espace des polynômes constants).
- Il en va de même pour l'endomorphisme de E donné par $P(X) \mapsto P(X+1)$.
- Les valeurs propres de l'endomorphisme de E défini par $P(X) \mapsto XP'(X)$ sont les entiers naturels. △

Exercice 1.5. *Dans l'exemple qui précède, que se passe-t-il si on remplace $\mathbb{K}[X]$ par l'ensemble des fonctions $C^\infty(\mathbb{R},\mathbb{R})$?*

1.2. Premières propriétés.

Proposition 1.6 (invariance du spectre par isomorphisme)**.** *Soient $f \in \mathrm{End}(E)$, F un $\mathbb{K}$-espace vectoriel, et $\varphi : E \to F$ un isomorphisme. Si on pose $g := \varphi \circ f \circ \varphi^{-1} \in \mathrm{End}(F)$ alors*

- $\mathrm{Sp}(g) = \mathrm{Sp}(f)$;
- *pour tout $\lambda \in \mathrm{Sp}(f)$, $v \in E$ est un vecteur propre associé à λ (pour f) si et seulement si $\varphi(v)$ est un vecteur propre associé à λ (pour g). Autrement dit, $E_\lambda(g) = \varphi\big(E_\lambda(f)\big)$.*

DÉMONSTRATION. Soit (λ, v) une paire constituée d'une valeur propre de λ et d'un vecteur propre v associé à λ, pour l'endomorphisme f. Autrement dit : $f(v) = \lambda v$, $v \neq 0_E$. Un simple calcul montre que $\varphi(v)$ (qui est $\neq 0_F$ car φ est un isomorphisme) est un vecteur propre de g, de valeur propre λ :

$$g\big(\varphi(v)\big) = (\varphi \circ f \circ \varphi^{-1})\big(\varphi(v)\big) = \varphi\big(f(v)\big) = \varphi(\lambda v) = \lambda\varphi(v).$$

Par conséquent $\mathrm{Sp}(f) \subset \mathrm{Sp}(g)$ et φ induit une application de l'ensemble des vecteurs propres de f associés à λ dans l'ensemble les vecteurs propres de g associés à λ. En effectuant le même raisonnement avec $f = \varphi^{-1} \circ g \circ \varphi$ on trouve que $\mathrm{Sp}(g) \subset \mathrm{Sp}(f)$ et que φ^{-1} induit une application de l'ensemble des vecteurs propres de f associés à λ dans l'ensemble les vecteurs propres de g associés à λ, inverse de celle induite par φ. $\square$

Corollaire 1.7. *Deux matrices conjuguées ont le même spectre.* $\square$

Proposition 1.8 (propriétés des espaces propres)**.** *Soient $\lambda, \mu \in \mathrm{Sp}(f)$. Alors $E_\lambda(f)$ est un sous-espace vectoriel de E, de dimension ≥ 1, qui est stable par f. De plus $E_\lambda(f) \cap E_\mu(f) = \{0_E\}$ si et seulement si $\lambda \neq \mu$.*

On rappelle qu'une partie $Y \subset E$ est dite **stable par** f lorsque $f(Y) \subset Y$.

DÉMONSTRATION. L'application $f - \lambda \mathrm{id}_E$ étant linéaire, son noyau $E_\lambda(f)$ est un sous-espace vectoriel de E. Comme λ est une valeur propre (par hypothèse) alors par définition il existe un vecteur non nul dans $E_\lambda(f)$, qui est donc de dimension au moins 1. Montrons maintenant que $E_\lambda(f)$ est stable par f : pour tout $v \in E_\lambda(f)$, $f(v) = \lambda v \in E_\lambda(f)$. Pour finir :

$\boxed{\Rightarrow}$ Si $\lambda = \mu$ alors $E_\lambda(f) \cap E_\mu(f) = E_\lambda(f) \neq \{0_E\}$.

$\boxed{\Leftarrow}$ Si $\lambda \neq \mu$ alors que que soit $v \in E_\lambda(f) \cap E_\mu(f)$, $\lambda v = f(v) = \mu(v)$, donc $(\lambda - \mu)v = 0_E$ et ainsi $v = 0_E$.

$\square$

Remarque 1.9. On peut en fait montrer par récurrence sur $k \geq 2$ que si $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ sont deux à deux distincts (autrement dit, $\lambda_i \neq \lambda_j$ si $i \neq j$), alors les sous-espaces propres $E_{\lambda_1}(f), \ldots, E_{\lambda_k}(f)$ sont en somme directe[1] :

— L'initialisation de la récurrence est donnée par la Proposition 1.8.

— Supposons que l'affirmation est vérifiée pour k, et donnons-nous $\lambda_0, \lambda_1 \ldots, \lambda_k \in \mathbb{K}$, $k+1$ scalaires deux à deux distincts. Par hypothèse de récurrence, $E_{\lambda_1}(f), \ldots, E_{\lambda_k}(f)$ sont en somme directe. Il nous suffit donc de démontrer que $E_{\lambda_0}(f)$ est en somme directe avec $E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_k}(f)$. Soit $v \in E_{\lambda_0}(f) \cap \big(E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_k}(f)\big)$. Alors $f(v) = \lambda_0 v$, et $v = v_1 + \cdots + v_k$ avec $f(v_i) = \lambda_i v_i$ pour tout $1 \leq i \leq k$. Par conséquent,

$$\sum_{i=1}^{k} \lambda_0 v_i = \lambda_0 v = f(v) = \sum_i \lambda_i v_i \,.$$

Par unicité de la décomposition dans la somme directe $E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_k}(f)$, on a $\lambda_0 v_i = \lambda_i v_i$. Pour tout $1 \leq i \leq k$, $\lambda_0 \neq \lambda_i$, donc $v_i = 0_E$. Ainsi $v = 0_E$.

1.3. Quelques exemples.

Exemple 1.10 (projection)**.** On considère une projection $p : E \to E$, c'est-à-dire une application linéaire telle que $p \circ p = p$. On rappelle (HAX202X) que $E = F \oplus G$, avec $F := \mathrm{im}(p)$ et $G := \ker(p)$. Mettons de côté deux cas extrêmes :

(1) si $p = \mathrm{id}_E$ (auquel cas $G = \{0_E\}$) alors p a une seule valeur propre qui est 1, et $E_1(p) = E$.

(2) si p est l'application nulle (auquel cas $F = \{0_E\}$) alors p a une seule valeur propre qui est 0, et $E_0(p) = E$.

On suppose maintenant que $F \neq \{0_E\}$ et $G \neq \{0_E\}$. Pour tout $v = p(w) \in F \neq \{0_E\}$, $p(v) = p(p(w)) = p(w) = v$, donc 1 est une valeur propre de p et $E_1(p) = F$. Comme $\ker(p) = G \neq \{0_E\}$, 0 est valeur propre de p et $E_0(p) = F$. Puisque $E = F \oplus G$, d'après la Proposition 1.8 p n'a pas d'autre valeur propre.

1. Attention, il ne suffit pas de montrer qu'ils sont deux à deux en somme directe.

Supposons que E est de dimension finie $n \in \mathbb{N}$, et choisissons une base $(v_1, \ldots, v_k)$ de F et une base $(v_{k+1}, \ldots, v_n)$ de G. Alors dans la base $\mathcal{B} = (v_1, \ldots, v_n)$ de E, la matrice de p s'écrit $M_{\mathcal{B},\mathcal{B}}(p) = \begin{pmatrix} \mathrm{I}_k & 0 \\ 0 & 0 \end{pmatrix}$. △

Exemple 1.11 (symétrie)**.** On considère une symétrie $s : E \to E$, c'est-à-dire une application linéaire telle que $s \circ s = \mathrm{id}_E$. On rappelle (HAX202X) que lorsque la caractéristique de $\mathbb{K}$ est différente de 2, ce que nous supposerons dans cet exemple, $E = F \oplus G$, avec $F := \ker(s - \mathrm{id}_E)$ et $G := \ker(s + \mathrm{id}_E)$. Mettons de côté deux cas extrêmes :

(1) si $s = \mathrm{id}_E$ (auquel cas $G = \{0_E\}$) alors s a une seule valeur propre qui est 1, et $E_1(p) = E$.
(2) si $s = -\mathrm{id}_E$ (auquel cas $F = \{0_E\}$) alors s a une seule valeur propre qui est -1, et $E_{-1}(p) = E$.

On suppose maintenant que $F \neq \{0_E\}$ et $G \neq \{0_E\}$. Pour tout $v =\in F \neq \{0_E\}$, $s(v) = v$, donc 1 est une valeur propre de p et $E_1(s) = F$. De même, $v =\in F \neq \{0_E\}$, $s(v) = -v$, donc -1 est une valeur propre de p et $E_{-1}(s) = G$. Puisque $E = F \oplus G$, d'après la Proposition 1.8 s n'a pas d'autre valeur propre.
Supposons que E est de dimension finie $n \in \mathbb{N}$, et choisissons une base $(v_1, \ldots, v_k)$ de F et une base $(v_{k+1}, \ldots, v_n)$ de G. Alors dans la base $\mathcal{B} = (v_1, \ldots, v_n)$ de E, la matrice de s s'écrit $M_{\mathcal{B},\mathcal{B}}(s) = \begin{pmatrix} \mathrm{I}_k & 0 \\ 0 & -\mathrm{I}_{n-k} \end{pmatrix}$. △

Dans l'exemple qui précède, l'hypothèse de caractéristique différente de 2 est importante car en caractéristique 2, $-\mathrm{id}_E = \mathrm{id}_E$ et donc $F = G$.

Exemple 1.12 (une symétrie en caractéristique 2)**.** On suppose que $\mathbb{K}$ est un corps de caractéristique 2, et on considère l'endomorphisme s de $\mathbb{K}^2$ donné par la matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Commencez par vérifier en exercice que s est une symétrie. Cet endomorphisme n'a qu'une seule valeur propre qui est 1, et dont l'espace propre associé est $E_1(s) = \mathbb{K}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. En effet, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}$ si et seulement si $a + b = \lambda a$ et $b = \lambda b$, ce qui arrive si et seulement si $\lambda = 1$ et $b = 0$ ou $a = b = 0$. △

Exercice 1.13 (endomorphisme d'ordre k)**.** *Soit k un entier naturel. Supposons que $f \in \mathrm{End}(E)$ est tel que $f^{\circ k} = \mathrm{id}_E$. Montrer que $\mathrm{Sp}(f) \subset \{\lambda \in \mathbb{K} \mid \lambda^k = 1\}$.*

2. Diagonalisation

Dans toute la suite du chapitre, on supposera que E est un $\mathbb{K}$-espace vectoriel de dimension finie.

2.1. Endomorphisme diagonalisable.

Définition 2.1 (diagonalisabilité)**.** On dit qu'un endomorphisme $f \in \mathrm{End}(E)$ est **diagonalisable** s'il existe une base $\mathcal{B}$ de E telle que $M_{\mathcal{B},\mathcal{B}}(f)$ est une matrice diagonale. On dit qu'une matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ est diagonalisable si l'endomorphisme de $\mathbb{K}^n$ qui lui est correspond l'est ; ça revient à dire que M est conjuguée à une matrice diagonale : $\exists P \in GL_n(\mathbb{K})$, PMP^{-1} diagonale.

On a vu dans l'Exemple 1.10 que toute projection est diagonalisable, et dans l'Exemple 1.11 qu'en caractéristique différente de 2 toute symétrie est diagonalisable.

Exemple 2.2 (une matrice non diagonalisable)**.** Considérons la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Elle ne peut pas être diagonalisable. En effet, si il existait une matrice inversible $P \in GL_2(\mathbb{K})$ telle que $PAP^{-1} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ alors $a^2 = d^2 = 0$ (Pouvez-vous jusifier pourquoi ?), de sorte que $a = d = 0$. Mais tout matrice conjuguée à la matrice nulle est nulle. Par l'absurde, A n'est pas diagonalisable. △

Proposition 2.3. *Si f possède $n = \dim(E)$ valeurs propres distinctes alors f est diagonalisable. Plus précisément, il existe une base $v = (v_1, \ldots, v_n)$ constituées de vecteurs propres telle que $M_{v,v}(f) = diag(\lambda_1, \ldots, \lambda_n)$, où λ_i est la valeur propre associée au vecteur propre v_i.*

DÉMONSTRATION. Nous avons vu (Proposition 1.8 et Remarque 1.9) que les espaces propres associés à des valeurs propres distinctes sont en somme directe. Donc, si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres (disctinctes, par hypthèse) de f alors

$$E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_n}(f) \subset E\,.$$

Si on note $n_i := \dim\big(E_{\lambda_i}(f)\big) \geq 1$, $i \in \{1, \ldots, n\}$, alors $\sum_{i=1}^n n_i \leq n$ et ainsi on a nécessairement $n_i = 1$ quelque soit i, et l'inclusion devient une égalité :

$$E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_n}(f) = E\,.$$

Pour chaque $i \in \{1, \ldots, n\}$ on choisit un vecteur propre v_i associé à la valeur propre λ_i, et on obtient

$$\mathbb{K}v_1 \oplus \cdots \oplus \mathbb{K}v_n = E\,.$$

Autrement dit $(v_1, \ldots, v_n)$ est une base de E et, pour tout $i \in \{1, \ldots, n\}$, $f(v_i) = \lambda_i$. Donc $f = diag(\lambda_1, \ldots, \lambda_n)$. □

Comme dans la partie précédédente, il est extrêmement important de toujours préciser sur quel corps on travaille.

Exemple 2.4 (permutation circulaire)**.** La matrice de permutation circulaire

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

est diagonalisable dans $\mathrm{Mat}_{3\times 3}(\mathbb{C})$. En effet, si on note $j = e^{2i\pi/3}$ alors

$$M\begin{pmatrix} 1 \\ j \\ j^2 \end{pmatrix} = \begin{pmatrix} j^2 \\ 1 \\ j \end{pmatrix} = j^2\begin{pmatrix} 1 \\ j \\ j^2 \end{pmatrix}\,,\quad M\begin{pmatrix} 1 \\ j^2 \\ j \end{pmatrix} = \begin{pmatrix} j \\ 1 \\ j^2 \end{pmatrix} = j\begin{pmatrix} 1 \\ j^2 \\ j \end{pmatrix} \quad \text{et} \quad M\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

La matrice M ayant donc 3 valeurs propre distinctes, la Proposition 2.3 nous apprend qu'elle est diagonalisable. Trouvez la matrice de changement de base qui permet de diagonaliser M. △

2.2. Polynôme caractéristique.

Notons $R = \mathbb{K}[X]$ l'anneau des polynômes à coefficients dans $\mathbb{K}$. On considère la matrice $X \cdot \mathrm{I}_n := diag(X, \ldots, X) \in \mathrm{Mat}_{n\times n}(R)$. On remarque aussi que toute matrice à coefficients dans $\mathbb{K}$ peut être vue comme une matrice à coefficients dans R (ses coefficients sont des polynômes constants).

Définition 2.5 (polynôme caractéristique)**.** Le **polynôme caractéristique** d'une matrice $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ est le polynôme

$$\chi_A(X) := \det(X \cdot \mathrm{I}_n - A) \in \mathbb{K}[X]\,.$$

Exemples 2.6 (exemples et propriétés immédiates)**.**

(1) Si $A = diag(a_1, \ldots a_n)$ alors $\chi_A(X) = \prod_{i=1}^n (X - a_i)$.

(2) Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ alors $\chi_A(X) = (X-a)(X-d) - cb$.

(3) D'après la Proposition 2.8, $\chi_A(X) = \chi_{A^\intercal}(X)$.

(4) D'après la Proposition 2.9, si $A = (a_{ij})_{1\le i,j\le n}$ est triangulaire (supérieure ou inférieure) alors $\chi_A(X) = \prod_{i=1}^n (X - a_{ii})$.

(5) D'après la Proposition 2.12, si M est triangulaire par blocs comme dans l'équation (1) alors $\chi_M(X) = \chi_A(X)\chi_D(X)$.

(6) Si M est la matrice de permutation circulaire de l'Exemple 2.4 alors

$$\chi_M(X) = \begin{vmatrix} X & 0 & -1 \\ -1 & X & 0 \\ 0 & -1 & X \end{vmatrix} = X \cdot (X^2 + 0) + 1 \cdot (0 - 1) = X^3 - 1$$

△

Remarque 2.7. Vérifier en exercice que si $P \in GL_n(\mathbb{K})$ alors $P(X \cdot \mathrm{I}_n)P^{-1} = X \cdot \mathrm{I}_n$. Par conséquent le polynôme caractéristique est invariant par conjugaison : $\chi_{PAP^{-1}}(X) = \chi_A(X)$. On peut donc parler du **polynôme caractéristique d'un endomorphisme de** E (on rappelle que E est supposé de dimension finie depuis le début de la Section 2 du présent Chapitre).

Théorème 2.8 (spectre = racines de χ)**.** *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Un scalaire $\lambda \in \mathbb{K}$ est racine de $\chi_A(X)$ (c'est-à-dire que $\chi_A(\lambda) = 0$) si et seulement si c'est une valeur propre de A.*

DÉMONSTRATION. On raisonne par équivalences :

$$\begin{aligned} \lambda \in \mathrm{Sp}(A) &\Leftrightarrow \exists v \in \mathbb{K}^n \backslash \{0_{\mathbb{K}^n}\}\,,\, Av = \lambda v \\ &\Leftrightarrow \ker(A - \lambda \mathrm{I}_n) \neq \{0_{\mathbb{K}^n}\} \\ &\Leftrightarrow A - \lambda \mathrm{I}_n \text{ non inversible} \\ &\Leftrightarrow \chi_A(\lambda) \overset{def}{=} \det(A - \lambda \mathrm{I}_n) = 0 \end{aligned}$$

□

Exemple 2.9. On rappelle (Exemple 2.6(6)) que le polynôme caractéristique de la matrice de permutation circulaire de l'Exemple 2.4 est $X^3 - 1$. Ses racines sont les 3 racines cubiques de l'unité dans $\mathbb{C}$, ce qui coïncide bien avec les trois valeurs propres trouvées dans l'Exemple 2.4. △

Exemple 2.10. Si $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ alors $\chi_A(X) = (X-1)^2 - 1 = X^2 - 2X = X(X-2)$. Cette matrice possède donc deux valeurs propres distinces (0 et 2), et est en conséquence diagonalisable, conjuguée à la matrice $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. On remarque que $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est un vecteur propre pour la valeur propre 2 et que $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est un vecteur propre pour la valeur propre 0. Oups, nous n'avons pas été assez vigilants ; pouvez-vous dire ce qui se passe si le corps $\mathbb{K}$ est de caractéristique 2 ? [Indication : essayez de montrer que A est conjuguée à la matrice de l'Exemple 2.2.] $\triangle$

Corollaire 2.11. *Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie a au moins une valeur propre (complexe).*

DÉMONSTRATION. Le corollaire peut se reformuler comme suit : toute matrice à coefficients complexes admet au moins une valeur propre complexe. Ce résultat découle du Théorème 2.8 et du **théorème fondamental de l'algèbre**[2], qui affirme que tout polynôme non constant à coefficients complexes admet au moins une racine complexe[3]. $\square$

Exemple 2.12 (une matrice sans valeur propre)**.** Revenons sur la Remarque 1.3 ; pour voir que la matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'a pas de valeur propre réelle, on observe simplement que son polynôme caractéristique est $X^2 + 1$, qui n'a pas de racine réelle. $\triangle$

Corollaire 2.13. *Pour tout endomorphisme f d'un $\mathbb{R}$-espace vectoriel E de dimension finie, il existe au moins un sous-espace vectoriel de E de dimension 1 ou 2 qui est stable par f (i.e. tel que $f(V) \subset V$).*

DÉMONSTRATION. Le corollaire peut se reformuler ainsi : pour toute matrice A à coefficients réels de taille $n \times n$, il existe un sous-espace vectoriel $V \subset \mathbb{R}^n$ de dimension 1 ou 2 tel que $AV \subset V$. Comme $\mathbb{R} \subset \mathbb{C}$ on peut voir A comme une matrice à coefficients complexes, et d'après le Corollaire 2.11 A admet une valeur propre complexe $\lambda = a + ib \in \mathbb{C}$, avec $a, b \in \mathbb{R}$. Choisissons $v = u + iw \in \mathbb{C}^n$ un vecteur propre associé à cette valeur propre ; $u, w \in \mathbb{R}^n$ sont non tous deux nuls. On a :

$$Au + iAw = Av = \lambda v = au - bw + i(aw + bu) \quad \Longrightarrow \quad \begin{cases} Au = au - bw \\ Aw = aw + bu \,. \end{cases}$$

Par conséquent $V := \mathrm{Vect}(u, w)$ est stable par A, et comme u et w sont non tous deux nuls, $1 \leq \dim(V) \leq 2$. $\square$

2.3. Un critère de diagonalisabilité.

On rappelle que pour tout polynôme $P \in \mathbb{K}[X]$, $\lambda \in \mathbb{K}$ est une racine de P si et seulement si $X - \lambda$ divise P.

Définition 2.14 (multiplicité)**.** Soit $\lambda \in \mathbb{K}$ un scalaire, $P \in \mathbb{K}[X]$ un polynôme, et $f \in \mathrm{End}(E)$ un endomorphisme (on rappelle que E est supposé de dimension finie).
- La **multiplicité de λ dans** P, notée $m_\lambda(P)$ est le plus grand nombre entier naturel m tel que $(X - \lambda)^m$ divise P.
- La **multiplicité algébrique de** λ **dans** f est la multiplicité de λ dans χ_f.
- La **multiplicité géométrique de** λ **dans** f est $\dim\big(E_\lambda(f)\big)$.

Si l'endomorphisme est clairement identifié, on note $m_a(\lambda)$ la multiplicité algébrique de λ (dans f) et $m_g(\lambda)$ sa multiplicité algébrique (toujours dans f).

Proposition 2.15. *Pour tout endomorphisme $f \in \mathrm{End}(E)$ et tout $\lambda \in \mathbb{K}$, $m_g(\lambda) \leq m_a(\lambda)$.*

2. Non démontré dans ce cours. Il est aussi appelé "théorème de d'Alembert" ou encore "théorème de d'Alembert-Gauss".

3. En faisant une récurrence sur le degré, on en déduit facilement que tout polynôme à coefficients complexes est scindé (sur $\mathbb{C}$), c'est-à-dire constant ou produit de polynômes de degré 1.

DÉMONSTRATION. On peut supposer que λ est une valeur propre de f ; en effet, si ça n'est pas le cas d'après le Théorème 2.8 $m_g(\lambda) = m_a(\lambda) = 0$. On sait (c'est la Proposition 1.8) que $E_\lambda(f)$ est stable par f.

LEMME 2.16. *Soit $G \subset E$ un sous-espace vectoriel stabe par f. On note $g \in \mathrm{End}(G)$ la restriction de f à G. Alors χ_g divise χ_f.*

DÉMONSTRATION DU LEMME. Soit $u = (u_1, \ldots, u_k)$ une base de G que l'on complète en une base $v = (u_1, \ldots, u_k, v_{k+1}, \ldots, v_n)$ de E. Alors la matrice de f dans cette base est triangulaire par blocs :

$$M_{v,v}(f) = \begin{pmatrix} M_{u,u}(g) & B \\ 0 & D \end{pmatrix}.$$

Par conséquent, d'après l'Exemple 2.6(5), $\chi_f(X) = \chi_g(X)\chi_D(X)$. □

Ainsi, d'après ce Lemme, le polynôme caractéristique χ_g de la restriction g de f au sous-espace stable $E_\lambda(f)$ divise χ_f. Or, par définition, $g = \lambda \mathrm{id}_{E_\lambda(f)}$, donc $\chi_g(X) = (X-\lambda)^{m_g(\lambda)}$. Il en découle que $m_g(\lambda) \leq m_a(\lambda)$. □

Exemple 2.17 (un endomorphisme sans sous-espace stable non trivial)**.** Considérons la matrice

$$N := \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{Q})$$

dont le polynôme caractéristique est $\chi_N(X) = X^3 - 2$ (à vérifier en exercice). Ce polynôme n'ayant aucune racine dans $\mathbb{Q}$, il n'a aucun facteur de degré < 3 dans $\mathbb{Q}[X]$. Il résulte donc du Lemme 2.16 que que N n'a aucun sous-espace stable de dimension 1 ou 2 dans $\mathbb{Q}^3$. △

Théorème 2.18 (critère de diagonalisabilité)**.** *Un endomorphisme $f \in \mathrm{End}(E)$ est diagonalisable si et seulement si χ_f est scindé et pour tout $\lambda \in \mathrm{Sp}(f)$, $m_a(\lambda) = m_g(\lambda)$.*

Exemples 2.19.

(1) Reprenons pour commencer l'Exemple 2.2. Le polynôme caractéristique de la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est X^2, qui est scindé, ayant pour unique racine 0, de multiplicité $m_a(0) = 2$. Or le noyau de l'application linéaire canoniquement associée à A est $\mathbb{K}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, qui est de dimension 1. Donc $m_g(0) = 1 < 2 = m_a(0)$. La matrice n'est pas diagonalisable, comme nous l'avions déjà démontré dans l'Exemple 2.2.

(2) Reprenons maintenant l'Exemple 2.12 de la matrice $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, dont le polynôme caractéristique est X^2+1. Sur $\mathbb{C}$ ce polynôme est scindé (comme tous les polynômes à coefficients complexes) avec deux valeurs propres distinctes de multiplicité 1 : $X^2+1 = (X-i)(X+i)$. Vue comme matrice à coefficients complexes, B est diagonalisable. À titre d'exercice, écrivez la matrice de passage permettant de diagonaliser B. Néanmoins, le polynôme X^2+1 n'est pas scindé sur $\mathbb{R}$, de sorte que B n'est pas diagonalisable comme matrice à coefficients réels.

(3) Reprenons la même matrice B du point précédent, que l'on voit comme une matrice à coefficients dans un corps $\mathbb{K}$ de caractéristique 2, auquel cas $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Dans cette situation le polynôme caractéristique se factorise, $X^2+1 = (X-1)^2$, et a donc

une racine double ; 1 est ainsi l'unique valeur propre de B et a pour multiplicité algébrique 2. Or la multiplicité géométrique de cette valeur propre est $1 = 2 - \mathrm{rg}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.
Conclusion : B n'est pas diagonalisable sur un corps de caractéristique 2.

(4) Vérifiez en exercice que la matrice de permutation circulaire M de l'exercice 2.4 n'est pas diagonalisable dans $\mathrm{Mat}_{3\times 3}(\mathbb{R})$. △

Démonstration du Théorème. $\boxed{\Rightarrow}$ Supposons que f est diagonalisable. Alors il existe une base $v = (v_1, \dots, v_n) \in E^n$ et des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tels que $M_{v,v}(f) = diag(\lambda_1, \dots, \lambda_n)$. Par conséquent $\chi_f(X) = \prod_{i=1}^n (X - \lambda_i)$ est scindé. Soit $\lambda \in \mathbb{K}$ une valeur propre de f ; d'après le Théorème 2.8, $\lambda \in \{\lambda_i \,|\, 1 \leq i \leq n\}$. La multiplicité algébrique de λ est

$$m_a(\lambda) = \#\{1 \leq i \leq n \,|\, \lambda = \lambda_i\}\,.$$

Calculons la multiplicité géométrique de λ. On sait que $M_{v,v}(f - \lambda \mathrm{id}_E) = diag\big((\lambda_i - \lambda)_{1 \leq i \leq n}\big)$, ce qui signifie que $M_{v,v}(f - \lambda \mathrm{id}_E)$ est une matrice diagonale avec exactement $m_a(\lambda)$ coefficients diagonaux nuls. Donc $m_g(\lambda) \overset{def}{=} \dim\big(\ker(f - \lambda \mathrm{id}_E)\big) = m_a(\lambda)$.
$\boxed{\Leftarrow}$ Supposons que χ_f est scindé : $\chi_f(X) = \prod_{i=1}^k (X - \lambda_i)^{n_i}$, avec $\lambda_i \neq \lambda_j$ si $i \neq j$. En particulier, $n_i = m_a(\lambda_i)$ et $\sum_{i=1}^k n_i = \dim(E)$. On rappelle que

$$G := \bigoplus_{i=1}^k E_\lambda(f) \subset E$$

est stable par f, et que la restriction $g \in \mathrm{End}(G)$ de f à G est diagonalisable (en effet, $g = \sum_{i=1}^k \lambda_i \mathrm{id}_{E_{\lambda_i}(f)}$). Si on suppose de surcroît que $m_g(\lambda_i) = m_a(\lambda_i) = n_i$ quel que soit i, alors $\dim(G) = \dim(E)$ et donc $G = E$, de sorte que f est diagonalisable. □

Exemple 2.20. Reprenons l'Exemple 1.2 de la matrice $A := \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$. Son polynôme caractéristique est scindé : $\chi_A(X) = (X-2)^2(X-3)$. Néanmoins la valeur propre 2 a pour multiplicité algébrique 2 et pour multiplicité géométrique 1 (en effet, on a vu dans l'Exemple 1.2 que $E_2(A)$ est la droite vectorielle engendrée par e_1). Donc A n'est pas diagonalisable (sur $\mathbb{R}$). Vérifier en exercice qu'elle n'est pas non plus diagonalisable sur $\mathbb{C}$. △

3. Trigonalisation

3.1. Endomorphisme trigonalisable.

Définition 3.1 (trigonalisabilité)**.** On dit qu'un endomorphisme $f \in \mathrm{End}(E)$ est **trigonalisable** s'il existe une base $\mathcal{B}$ de E telle que $M_{\mathcal{B},\mathcal{B}}(f)$ est une matrice triangulaire supérieure. On dit qu'une matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ est trigonalisable si l'endomorphisme de $\mathbb{K}^n$ qui lui est correspond l'est ; ça revient à dire que M est conjuguée à une matrice triangulaire supérieure.

Exemple 3.2. Toute matrice triangulaire inférieure est trigonalisable. En effet, si

$$P := \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}$$

alors, pour toute matrice $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ qui est triangulaire inférieure A, PAP^{-1} est triangulaire supérieure (à vérifier en exercice[4]). △

Exemple 3.3. La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ de l'Exemple 2.12 n'est pas trigonalisable sur $\mathbb{R}$. En effet, nous avons vu qu'elle n'a pas de valeur propre. Or toute matrice triangulaire supérieure (et donc toute matrice trigonalisable) admet au moins une valeur propre (pouvez-vous dire pourquoi ?). △

3.2. Propriétés des endomorphismes trigonalisables.
Soit $f \in \mathrm{End}(E)$ un endomorphisme trigonalisable. Alors il existe une base $\mathcal{B}$ de E telle que $M := M_{\mathcal{B},\mathcal{B}}(f)$ est une matrice triangulaire supérieure. Notons $\lambda_1, \ldots, \lambda_n$ les éléments diagonaux de M. Résumons ce que nous savons :
— $\det(f) = \lambda_1 \cdots \lambda_n$.
— $\mathrm{tr}(f) = \lambda_1 + \cdots + \lambda_n$.
— $\chi_f(X) = (X - \lambda_1)\cdots(X - \lambda_n)$.

Proposition 3.4 (invariants et valeurs propres)**.** *Pour tout $k \in \mathbb{N}$, $\mathrm{tr}(f^{\circ k}) = \lambda_1^k + \cdots + \lambda_n^k$.*

DÉMONSTRATION. Il suffit de montrer que M^k est une matrice trianglaire supérieure dont les coefficients diagonaux sont $\lambda_1^k, \ldots, \lambda_n^k$ (à faire en exercice). □

Dans le reste de ce paragraphe, on supposera que $\mathbb{K}$ est un sous-corps de $\mathbb{C}$, et on se donnera un scalaire $t \in \mathbb{K}$ tel que $|t| < \frac{1}{\alpha}$, où

$$\alpha = \max_{1\le i\le n}(|\lambda_i|)\,.$$

Cela garantit que pour tout i, $1 - t\lambda_i$ est inversible dans $\mathbb{K}$, d'inverse

$$(1 - t\lambda_i)^{-1} = \sum_{k=0}^{+\infty} \lambda_i^k t^k\,.$$

On a une autre formule pour l'inverse :

$$(1 - t\lambda_i)^{-1} = \exp\big(-\log(1 - t\lambda_i)\big) = \exp\left(\sum_{k=1}^{+\infty} \frac{\lambda_i^k}{k} t^k\right)\,.$$

Comme $|t\lambda_i| < 1$, les deux séries précédentes convergent.

Proposition 3.5 (déterminant d'une matrice triangulaire proche de l'identité)**.** *Sous les hypothèses qui précèdent, $\mathrm{I}_n - t\cdot M$ est inversible et*

$$\det(\mathrm{I}_n - t\cdot M) = \exp\left(-\sum_{k=1}^{+\infty} \frac{\mathrm{tr}(M^k)}{k} t^k\right)\,.$$

DÉMONSTRATION. D'après ce qui précède,

$$\det(\mathrm{I}_n - t\cdot M) = (1 - t\lambda_1)\cdots(1 - t\lambda_n)$$

est inversible. Donc $\mathrm{I}_n - t\cdot M$ est inversible. On calcule :

$$\det(\mathrm{I}_n - t\cdot M)^{-1} \;=\; \prod_{i=1}^{n}(1 - t\lambda_i)^{-1} = \prod_{i=1}^{n}\left(\exp\left(\sum_{k=1}^{+\infty}\frac{\lambda_i^k}{k}t^k\right)\right)$$

4. Vous pouvez utiliser l'Exercice no 5 du livret d'exercices.

$$= \exp\left(\sum_{i=1}^{n}\sum_{k=1}^{+\infty}\frac{\lambda_i^k}{k}t^k\right) = \exp\left(\sum_{k=1}^{+\infty}\frac{\sum_{i=1}^{n}\lambda_i^k}{k}t^k\right)$$
$$= \exp\left(\sum_{k=1}^{+\infty}\frac{\mathrm{tr}(M^k)}{k}t^k\right) .$$

En passant à l'inverse, on en déduit l'identité voulue. □

3.3. Un critère de trigonalisabilité.

Théorème 3.6 (critère de trigonalisabilité)**.** *Un endomorphisme $f \in \mathrm{End}(E)$ est trigonalisable si et seulement si χ_f est scindé.*

DÉMONSTRATION. On a déjà vu que si f est trigonalisable alors son polynôme caractéristique est scindé. Pour démontrer l'implication réciproque, on raisonne par récurrence sur la dimension de E.
Initialisation : c'est évident. Si $\dim(E) = 1$ alors tout endomorphisme f de E est un multiple de l'identité : $f = \lambda \mathrm{id}_E$ est trigonalisable et son polynôme caractéristique $\chi_f(X) = X - \lambda$ est scindé.
Hérédité : supposons qu'on a démontré qu'un endomorphisme d'un espace vectoriel de dimension n est trigonalisable si son polynôme caractéristique est scindé. Soit E un espace vectoriel de dimension $n+1$ et soit f un endomorphisme de E dont le polynôme caractéristique est scindé : $\chi_f(X) = (X-\lambda_1)\cdots(X-\lambda_{n+1})$, les $\lambda_i \in \mathbb{K}$ n'étant pas forcément distincts. On choisit v_1 un vecteur propre associé à la valeur propre λ_1 et on complète ce vecteur en une base $\mathcal{B} = (v_1, v_2, \ldots, v_{n+1})$ de E. On a

$$M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & L \\ 0 & N \end{pmatrix} ,$$

avec $L \in \mathrm{Mat}_{1,n}(\mathbb{K})$ une matrice ligne et $N \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Par conséquent (cf. Exemple 2.6(5)) $\chi_f(X) = (X-\lambda_1)\chi_f(N)$. D'après l'hypothèse de récurrence, comme $\chi_f(N) = (X-\lambda_2)\cdots(X-\lambda_{n+1})$ est scindé, la matrice N est trigonalisable : il existe $Q \in GL_n(\mathbb{K})$ tel que $Q^{-1}NQ$ soit triangulaire supérieure. On a donc

$$\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}^{-1} M_{\mathcal{B},\mathcal{B}}(f) \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & L \\ 0 & N \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & LQ \\ 0 & Q^{-1}NQ \end{pmatrix} ,$$

qui est triangulaire supérieure. □

Corollaire 3.7. *Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie (ou toute matrice à coefficients complexes) est trigonalisable.*

DÉMONSTRATION. Sur $\mathbb{C}$, tout polynôme est scindé. □

Le résultat qui suit nous apprend que la Proposition 3.5 reste vraie même lorsque M n'est pas trigonalisable.

Corollaire 3.8 (déterminant d'une matrice proche de l'identité)**.** *Supposons que $\mathbb{K}$ est un sous-corps de $\mathbb{C}$ et que $t \in \mathbb{K}$ satisfait $|t| < 1/|\lambda|$ pour toute valeur propre complexe λ de $M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Alors $\mathrm{I}_n - t\cdot M \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ est inversible et*

$$\det(\mathrm{I}_n - t\cdot M) = \exp\left(-\sum_{k=1}^{+\infty}\frac{\mathrm{tr}(M^k)}{k}t^k\right) .$$

DÉMONSTRATION. D'après le Corollaire 3.7, la matrice M est trigonalisable sur $\mathbb{C}$. Donc d'après la Proposition 3.5 l'égalité souhaitée est vérifiée. En particulier, $\det(\mathrm{I}_n - t \cdot M) \neq 0$, et donc M est inversible dans $\mathrm{Mat}_{n\times n}(\mathbb{K})$. □

Le fait de passer par un corps plus grand pour effectuer des calculs et raisonnements, puis revenir au corps de départ, peut s'avérer très utile. Ci-dessous on donne des exemples avec $\mathbb{R} \subset \mathbb{C}$.

Exemple 3.9. On rappelle (Exemple 3.3) que la matrice $B := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'est pas trigonalisable sur $\mathbb{R}$. En effet, son polynôme caractéristique $\chi_B(X) = X^2 + 1$ n'est pas scindé sur $\mathbb{R}$. Mais sur $\mathbb{C}$, $\chi_B(X) = (X - i)(X + i)$, et la matrice B a donc deux valeurs propres distinctes. On en déduit que pour tout $k \in \mathbb{N}$,

$$\mathrm{tr}(B^k) = i^k + (-i)^k = \begin{cases} 0 & \text{si } k \text{ est impair}\,; \\ (-1)^\ell \cdot 2 & \text{si } k = 2\ell\,. \end{cases}$$

On trouve également que $\det(\mathrm{I}_2 - tB) = (1 + it)(1 - it) = 1 + t^2$. Ainsi la formule de la Proposition 3.5 donne

$$(1 + t^2)^{-1} = \exp\left(\sum_{\ell=1}^{+\infty} \frac{(-1)^\ell}{\ell} t^{2\ell}\right).$$

△

Remarque 3.10. Une matrice de taille 2×2 à coefficients réels qui n'est pas trigonalisable sur $\mathbb{R}$ est forcément diagonalisable sur $\mathbb{C}$. En effet, son polynôme caractéristique (de degré 2) aura deux racines conjuguées non réelles, et donc distinctes.

Exemple 3.11. La matrice de permutation circulaire M de l'Exemple 2.4 n'est pas non plus trigonalisable sur $\mathbb{R}$, puisque son polynôme caractéristique $\chi_M(X) = X^3 - 1 = (X - 1)(X^2 + X + 1)$ n'est pas scindé sur $\mathbb{R}$. Sur $\mathbb{C}$ on a $\chi_M(X) = (X - 1)(X - j)(X - j^2)$, donc M est diagonalisable dans $\mathrm{Mat}_{3\times 3}(\mathbb{C})$ avec trois valeurs propres distinctes 1, j et j^2. On en déduit que pour tout $k \in \mathbb{N}$,

$$\mathrm{tr}(M^k) = 1 + j^k + j^{2k} = \begin{cases} 3 & \text{si } k \equiv 0 \bmod 3\,; \\ 0 & \text{sinon.} \end{cases}$$

On trouve aussi $\det(\mathrm{I}_2 - tB) = (1 - t)(1 - tj)(1 - tj^2) = (1 - t)(1 + t + t^2) = 1 - t^3$. La formule de la Proposition 3.5 donne ainsi

$$(1 - t^3)^{-1} = \exp\left(\sum_{\ell=1}^{+\infty} \frac{1}{\ell} t^{3\ell}\right).$$

△

Remarque 3.12. Une matrice de taille 3×3 à coefficients réels qui n'est pas trigonalisable sur $\mathbb{R}$ est forcément diagonalisable sur $\mathbb{C}$. En effet, son polynôme caractéristique (de degré 3) a une racine réelle et deux racines conjuguées non réelles (les trois valeurs propres complexes sont donc distinctes deux à deux).

Chapitre 3

Polynôme minimal et réduction

Soit E un $\mathbb{K}$-espace vectoriel. On notera $\mathbf{0}$ l'endomorphisme nul, afin de ne le confondre ni avec $0 \in \mathbb{K}$ ni avec $0_E \in E$.

1. Polynômes d'endomorphismes

1.1. Morphisme d'évaluation et polynômes annulateurs.

Pour tout endomorphisme $f \in \mathrm{End}(E)$, on rappelle (HAX202X) l'application linéaire d'**évaluation en** f

$$\begin{array}{rcl} ev_f \,:\, \mathbb{K}[X] & \longrightarrow & \mathrm{End}(E) \\ P(X) = a_0 + a_1 X + \cdots + a_d X^d & \longmapsto & P(f) := a_0 \mathrm{id}_E + a_1 f + \cdots + a_d f^{\circ d}\,. \end{array}$$

Proposition 1.1. *ev_f est un morphisme d'anneaux.*

DÉMONSTRATION. Il découle automatiquement de la définition que $ev_f(1) = \mathrm{id}_E$. Il reste donc à démontrer (en exercice) que $(PQ)(f) = P(f)Q(f)$. □

Remarque 1.2. Comme l'anneau des polynômes est commutatif, on trouve que pour tout endomorphisme $f \in \mathrm{End}(E)$, et pour tous polynômes $P, Q \in \mathbb{K}[X]$, $P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f)$. Autrement dit, $P(f)$ et $Q(f)$ commutent.

Exemple 1.3 (polynôme d'une matrice diagonale)**.** Si $E = \mathbb{K}^n$, on rappelle qu'il existe un isomorphisme canonique $\mathrm{End}(E) \cong \mathrm{Mat}_{n\times n}(\mathbb{K})$ (d'espace vectoriels et d'anneaux [1]) qui envoie f sur $M_{\mathcal{C},\mathcal{C}}(f)$, où $\mathcal{C} = (e_1, \ldots, e_n)$ est la base canonique. Pour toute matrice $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ on a donc aussi un morphisme d'évaluation $ev_A : \mathbb{K}[X] \to \mathrm{Mat}_{n\times n}(\mathbb{K})$ qui a les mêmes propriétés que ev_f. Étant donnée une matrice diagonale $A = diag(\lambda_1, \ldots, \lambda_n)$ on sait que pour tout $k \in \mathbb{N}$, $A^k = diag(\lambda_1^k, \ldots, \lambda_n^k)$, de sorte que $P(A) = diag\big(P(\lambda_1), \ldots, P(\lambda_n)\big)$. △

Proposition 1.4 (équivariance par conjugaison)**.** *Si $\phi : E \to F$ est un isomorphisme d'espaces vectoriels, que $f \in \mathrm{End}(E)$, et que $g := \phi \circ f \circ \phi^{-1} \in \mathrm{End}(F)$, alors pour tout polynôme $P \in \mathbb{K}[X]$, $P(g) = \phi \circ P(f) \circ \phi^{-1}$.*

Comme conséquence immédiate on obtient que si deux matrices sont conjuguées alors les évaluation d'un polynôme en ces deux matrices sont conjugués aussi, par la même matrice de passage. Autrement dit, si $B = QAQ^{-1}$ alors $P(B) = QP(A)Q^{-1}$.

DÉMONSTRATION. Il suffit de vérifier cette identitée pour $P(X) = X^n$ quel que soit $n \in \mathbb{N}$. Or on sait que $(\phi \circ f \circ \phi^{-1})^{\circ n} = \phi \circ f^{\circ n} \circ \phi^{-1}$. □

Définition 1.5 (polynôme annulateur)**.** On dit que $P \in \mathbb{K}[X]$ est un **polynôme annulateur** de $f \in \mathrm{End}(E)$ si $P(f) = \mathbf{0}$.

Exemples 1.6. Par définition :

1. L'année prochaine, vous parlerez de $\mathbb{K}$-**algèbre**.

(1) si $p \in \mathrm{End}(E)$ est une projection alors $X^2 - X$ est un polynôme annulateur de p ;

(2) si $s \in \mathrm{End}(E)$ est une symétrie alors $X^2 - 1$ est un polynôme annulateur de s ;

(3) si $f \in \mathrm{End}(E)$ est niloptent (ce qui signifie qu'il existe $n \in \mathbb{N}^*$ tel que $f^{\circ n} = \mathbf{0}$) alors il existe $n \in \mathbb{N}^*$ tel que X^n est un polynôme annulateur de f.

(4) $X - 1$ est un polynôme annulateur de id_E. △

Proposition 1.7 (spectre ⊂ racines)**.** *Si $P \in \mathbb{K}[X]$ est un polynôme annulateur de $f \in \mathrm{End}(E)$ alors toute valeur propre de f est racine de P.*

Attention, la réciproque n'est pas vraie ! Par exemple, le polynôme $(X-1)(X-2)$ est annulateur de l'endomorphisme id_E, dont 2 n'est pas une valeur propre.

DÉMONSTRATION. Soit $v \in E \backslash \{0_E\}$ un vecteur propre associé à la valeur propre λ. Alors pour tout $n \in \mathbb{N}$, $f^{\circ n}(v) = \lambda^n v$. Par conséquent $P(f)(v) = P(\lambda)(v)$. Ainsi, si $P(f) = \mathbf{0}$ alors $P(\lambda)v = 0_E$ et donc $P(\lambda) = 0$ (car $v \neq 0_E$). □

Théorème 1.8 (théorème de Cayley–Hamilton)**.** *Supposons que E est de dimension finie. Alors le polynôme caractéristique χ_f d'un endomorphisme $f \in \mathrm{End}(E)$ est annulateur de f.*

DÉMONSTRATION. Pour tout $v \in E$, notre objectif est de montrer que $\chi_f(f)(v) = 0_E$. Si $v = 0_E$ c'est évident, on suppose donc que $v \neq 0_E$ et on considère le plus grand entier $1 \leq k \leq n$ tel que $\mathcal{U} := \big(v, f(v), f^{\circ 2}(v), \dots, f^{\circ(k-1)}(v)\big)$ est une famille libre. Notons $F = \mathrm{Vect}(\mathcal{U}) \subset E$ le sous-espace vectoriel engendré par cette famille. Comme $\big(v, f(v), \dots, f^{\circ k}(v)\big)$ est liée, $f^{\circ k}(v) \in F$ et on en déduit deux choses :

— F est stable par f, de sorte qu'on peut définir un endomorphisme $g := f_{|F} \in \mathrm{End}(F)$;

— il existe des scalaires $a_0, \dots, a_{k-1} \in \mathbb{K}$ tels que
$$f^{\circ k}(v) + a_{k-1} f^{\circ(k-1)}(v) + \cdots + a_2 f^{\circ 2}(v) + a_1 f(v) + a_0 v = 0_E\,.$$

La matrice de g dans base $\mathcal{U}$ de F est une matrice dite compagnon :
$$M_{\mathcal{U},\mathcal{U}}(g) = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

dont le polynôme caractéristique est $\chi_g(X) = X^k + a_{k-1}X^{k-1} + \cdots + a_1 X + a_0$ (à démontrer en exercice). On observe que $\chi_g(f)(v) = 0$. Or nous savons d'après le Lemme 2.16 que χ_g divise χ_f : il existe un polynôme P tel que $\chi_f = P\chi_g$. Ainsi $\chi_f(f)(v) = P(f)\big(\chi_g(f)(v)\big) = 0_E$. □

UNE AUTRE DÉMONSTRATION DU THÉORÈME DE CAYLEY–HAMILTON. Choisissons une base $\mathcal{B}$ de E, notons $A := M_{\mathcal{B},\mathcal{B}}(f) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$, et montrons que $\chi_A(A) = \mathbf{0}$.

On **admettra** le point (hors-programme) qui suit : pour tout polynôme $P \in \mathbb{K}[X]$, il existe un corps $\mathbb{L}$ contenant $\mathbb{K}$ comme sous-corps (on dit que $\mathbb{L}$ est une **extension de corps** de $\mathbb{K}$) et tel que P est scindé dans $\mathbb{L}[X]$.

On choisit une extension de corps $\mathbb{L}$ de $\mathbb{K}$ dans laquelle χ_A est scindé, ce qui signifie que A est trigonalisable dans $\mathrm{Mat}_{n\times n}(\mathbb{L})$. Autrement dit, A est conjuguée par une matrice inversible Q à une matrice triangulaire supérieur T dans $\mathrm{Mat}_{n\times n}(\mathbb{L})$: $A = QTQ^{-1}$. On a donc

$\chi_A(A) = \chi_T(A) = Q\chi_T(T)Q^{-1}$. Il nous suffit ainsi de démontrer que $\chi_T(T) = \mathbf{0}$. On rappelle que, comme T est triangulaire, $\chi_T(X) = \prod_{i=1}^n (X - t_i)$ où $t_1, \ldots, t_n$ sont les éléments diagonaux de T.
Pour tout $0 \leq k \leq n$, notons $V_k := \mathrm{Vect}(e_1, \ldots e_k) \subset \mathbb{L}^n$ le sous-espace vectoriel engendré par les k premiers vecteurs de la base canonique (en particulier, $V_0 = \{0_{\mathbb{L}^n}\}$). On a $(T - t_k \cdot \mathrm{I}_n)(V_k) \subset V_{k-1}$: en effet, comme T est triangulaire supérieure, $T(e_k) \in t_k e_k + V_{k-1}$. On conclut par un calcul récursif :

$$\begin{aligned} \chi_T(T)(\mathbb{L}^n) &= (T - t_1 \cdot \mathrm{I}_n) \cdots (T - t_{n-1} \cdot \mathrm{I}_n)(T - t_n \cdot \mathrm{I}_n)(V_n) \\ &\subset (T - t_1 \cdot \mathrm{I}_n) \cdots (T - t_{n-1} \cdot \mathrm{I}_n)(V_{n-1}) \\ &\vdots \\ &\subset (T - t_1 \cdot \mathrm{I}_n)(V_1) \subset V_0 = \{0_{\mathbb{L}^n}\}. \end{aligned}$$

Donc $\chi_T(T) = \mathbf{0}$. □

1.2. Polynôme minimal.

Définition 1.9 (polynôme minimal)**.** Un dit qu'un polynôme annulateur d'un endomorphisme $f \in \mathrm{End}(E)$ est **minimal** si son degré est minimal parmi les polynômes annulateurs de f qui sont non nuls.

Proposition 1.10 (minimalité des polynômes minimaux)**.** *Si P est un polynôme minimal de $f \in \mathrm{End}(E)$ alors P divise tout polynôme annulateur de f.*

DÉMONSTRATION. Soit Q un polynôme annulateur de f. On note R le reste de la division euclidienne de Q par P : $R = Q - AP$ et $deg(R) < deg(P)$. Ainsi $R(f) = Q(f) - A(f) \circ P(f) = \mathbf{0}$ et donc R est un polynôme annulateur de f de degré strictement plus petit que P. Par conséquent $R = 0$ et P divise Q. □

On en déduit que si P_1 et P_2 sont deux polynômes minimaux de f alors il existe $\lambda \in \mathbb{K}^\times$ tel que $P = \lambda Q$. Par conséquent, il existe un unique polynôme minimal de f qui soit **unitaire** (c'est-à-dire dont le coefficient dominant est 1). On l'appelle **le polynôme minimal de** f et on le note μ_f.

Proposition 1.11 (spectre = racines, le retour)**.** *Les valeurs propres de f son exactement les racines de son polynôme minimal μ_f.*

DÉMONSTRATION. Par définition, μ_f est un polynôme annulateur de f. Donc d'après la Proposition 1.7 toute valeur propre de f est racine de μ_f. Par ailleurs, la Proposition 1.10 et le Théorème de Cayley–Hamilton (Théorème 1.8) nous disent que μ_f divise le polynôme caractéristique χ_f. Donc toute racine de μ_f est une racine de χ_f, et toute racine de χ_f est une valeur propre (c'est le Théorème 2.8). □

Exemple 1.12. Le polynôme caractéristique de $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est $(X-1)^2$. Il coïncide donc avec le polynôme caractéristique de la matrice identité I_2. Néanmoins, leurs polynômes minimaux ne coïncdent pas :
— $X - 1$ annule I_2, de sorte que $\mu_{\mathrm{I}_2}(X) = X - 1$;
— $X - 1$ n'annule pas A, de sorte que $\mu_A(X) = \chi_A(X) = (X-1)^2$. △

Exemple 1.13. Supposons que $E = F_1 \oplus F_2$ et considérons $f = \lambda_1 \mathrm{id}_{F_1} + \lambda_2 \mathrm{id}_{F_2}$. Vérifions que $P(X) = (X - \lambda_1)(X - \lambda_2)$ est un polynôme annulateur de f : on choisit une base $(v_1, \ldots, v_k)$

de F_1 et une base $(v_{k+1}, \ldots, v_n)$ de F_2, de sorte que dans la matrice $\mathcal{B} = (v_1, \ldots, v_n)$ de E la matrice de f est diagonale par blocs de la forme

$$M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 \cdot \mathrm{I}_k & \mathbf{0} \\ \mathbf{0} & \lambda_2 \cdot \mathrm{I}_{n-k} \end{pmatrix},$$

et ainsi

$$P\big(M_{\mathcal{B},\mathcal{B}}(f)\big) = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\lambda_2 - \lambda_1) \cdot \mathrm{I}_{n-k} \end{pmatrix} \begin{pmatrix} (\lambda_1 - \lambda_2) \cdot \mathrm{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \mathbf{0}.$$

Si $\lambda_1 \neq \lambda_2$ alors ni $(X - \lambda_1)$ ni $(X - \lambda_2)$ n'annule f, et dans ce cas $\mu_f(X) = (X - \lambda_1)(X - \lambda_2)$. [On rappelle au passage que le polynôme caractéristique est dans ce cas $\chi_f(X) = (X - \lambda_1)^k (X - \lambda_2)^{n-k}$.] $\triangle$

Exercice 1.14. *Généraliser l'exemple qui précède à la situation suivante. Si $E = F_1 \oplus \cdots \oplus F_m$ et $f = \sum_{i=1}^m \lambda_i \mathrm{id}_{F_i}$ alors montrer que le polynôme $P(X) = \prod_{i=1}^m (X - \lambda_i)$ annule f. Si de plus les λ_i sont distincts deux à deux alors montrer que $P(X)$ est le polynôme minimal de f. En déduire que le polynôme minimal d'une matrice diagonale est scindé à racines simples (les racines étant les valeurs propres de la matrice). Comparer avec le polynôme caractéristique.*

Théorème 1.15 (encore un critère de diagonalisabilité)**.** *Un endomorphisme $f \in \mathrm{End}(E)$ est diagnalisable si et seulement si son polynôme minimal est scindé à racines simples.*

DÉMONSTRATION. $\boxed{\Rightarrow}$ Si f est diagonalisable alors il existe une base de E dans laquelle la matrice de f est diagonale. D'après l'exercice 1.14 le polynôme minimal d'un telle matrice est scindé à racines simples.
$\boxed{\Leftarrow}$ Supposons réciproquement que μ_f est scindé à racines imples : $\mu_f(X) = \prod_{i=1}^m (X - \lambda_i)$, avec des λ_i deux à deux distincts. Nous allons avoir besoin du résultat suivant :

LEMME 1.16 (lemme des noyaux, aussi appelé “théorème de décomposition des noyaux”). *Soient $P_1, \ldots, P_k \in \mathbb{K}[X]$ des polynômes deux à deux premiers entre eux. Alors*

$$\ker\big((P_1 \cdots P_k)(f)\big) = \bigoplus_{i=1}^k \ker\big(P_i(f)\big)$$

DÉMONSTRATION DU LEMME. Commençons par démontrer le lemme dans le cas de deux polynômes $P, Q \in \mathbb{K}[X]$ premiers entre eux. On rappelle ce que ça signifie : les seuls diviseurs communs de P et Q sont les polynômes constants (non nul). D'après le théorème de Bézout, c'est équivalent à l'affirmation suivant : il existe $A, B \in \mathbb{K}[X]$ tels que $AP + BQ = 1$. Il est clair que $\ker\big(P(f)\big) \subset \ker\big((PQ)(f)\big)$: si $P(f)(v) = 0_E$ alors

$$(PQ)(f)(v) = (QP)(f)(v) = Q(f)\big(P(f)(v)\big) = Q(f)(0_E) = 0_E.$$

On a la même chose avec Q, de sorte que $\ker\big(P(f)\big) + \ker\big(Q(f)\big) \subset \ker\big((PQ)(f)\big)$. Montrons l'inclusion réciproque. Pour tout $v \in E$ on a $v = v_1 + v_2$ avec

$$v_1 = (AP)(f)(v) \qquad \text{et} \qquad v_2 = (BQ)(f)(v).$$

Si $(PQ)(f)(v) = 0_E$ alors $Q(f)(v_1) = 0$ et $P(f)(v_2) = 0_E$. Donc

$$\ker\big((PQ)(f)\big) \subset \ker\big(P(f)\big) + \ker\big(Q(f)\big).$$

Montrons enfin que la somme est directe. Soit $v \in \ker\big(P(f)\big) \cap \ker\big(Q(f)\big)$. Alors

$$v = (AP)(f)(v) + (BQ)(f)(v) = 0_E + 0_E = 0_E.$$

Pour terminer la démonstration, on remarque que $P_1, \dots P_k \in \mathbb{K}[X]$ sont 2 à 2 premiers entre eux si et seulement si P_1 est premier avec $P_2 \cdots P_k$ et $P_2, \dots, P_k$ sont 2 à 2 premiers entre eux. On peut alors démontrer le résultat par récurrence en utilisant à chaque étape le résultat que l'on vient de démontrer pour deux polynômes. □

Continuons maintenant la démonstration du Théorème 1.15. Comme μ_f annule f (par définition), et que les $X - \lambda_i$ sont deux à deux premiers entre eux, le lemme des noyaux nous dit que

$$E = \ker(\mathbf{0}) = \ker\big(\mu_f(f)\big) = \oplus_{i=1}^m \ker(f - \lambda_i \mathrm{id}_E)\,.$$

En choisissant une base $\mathcal{B}_i$ pour chaque i et en posant $\mathcal{B} = \mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_m$ on trouve que la matrice de f dans la base $\mathcal{B}$ de E est $diag(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots\dots, \lambda_m, \dots, \lambda_m)$, où chaque λ_i est répété $\dim\big(\ker(f - \lambda_i \mathrm{id}_E)\big)$ fois. □

Corollaire 1.17. *Un endomorphisme $f \in \mathrm{End}(E)$ est diagonalisable si et seulement si il admet un polynôme annulateur scindé à racines simples.*

DÉMONSTRATION. Tout diviseur d'un polynôme scindé à racines simples est lui-même scindé à racines simples. □

Corollaire 1.18. *Si $f \in \mathrm{End}(E)$ et $F \subset E$ est stable par f alors $g := f_{|F} \in \mathrm{End}(F)$ est diagonalisable aussi.*

DÉMONSTRATION. On sait que $\mu_f(f) = \mathbf{0}$ (par définition), de sorte que $\mu_f(g) = \mathbf{0}$. Or si f est diagonalisable alors μ_f est scindé à racines simples (c'est le Théorème 1.15) et donc g est diagonalisable (c'est le Corollaire 1.17). □

1.3. Sous-espaces caractéristiques.

On a vu dans la démonstration du Théorème 2.18 que pour un endomorphisme diagonalisable $f \in \mathrm{End}(E)$, on a une décomposition

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) \qquad \text{avec} \quad f_{|E_\lambda(f)} = \lambda \mathrm{id}_{E_\lambda(f)}\,.$$

L'objectif de cette partie est de montrer que dans le cas d'un endomorphisme trigonalisable $f \in \mathrm{End}(E)$ on a une décomposition

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} K_\lambda(f) \qquad \text{avec} \quad f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)} \text{ nilpotent}\,.$$

Pour tout le reste de cette partie, on fixe un endomorphisme $f \in \mathrm{End}(E)$, et on ne suppose pour l'instant plus que E est de dimension finie (mais cette hypothèse reviendra vite).

Définition 1.19 (vecteur propre généralisé)**.** On dit que $v \in E$ est un **vecteur propre généralisé** de f de poids $\lambda \in \mathbb{K}$ si il existe un entier $m \in \mathbb{N}$ tel que $(f - \lambda \mathrm{id}_E)^{\circ m}(v) = 0_E$. Le plus petit entier m satisfaisant la condition est appelé **hauteur** de v.

Les vecteurs propres généralisés de hauteur 1 sont exactement les vecteurs propres au sens usuel.

Exemples 1.20.

(1) Le vecteur nul est un vecteur propre généralisé de hauteur 0 de tout endomorphisme f, de n'importe quel poids $\lambda \in \mathbb{K}$.

(2) Tous les vecteurs de $\mathbb{K}^2$ sont des vecteurs propres généralisés de poids λ de la matrice $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Le vecteur nul est de hauteur 0, les vecteurs non nuls colinéaires à e_1 sont de hauteur 1, et les autres vecteurs sont de hauteur de 2. $\triangle$

Exercice 1.21. *Montrer que si $v \in E$ est un vecteur propre généralisé de poids λ et de hauteur $m \geq 1$ pour f alors $w := (f - \lambda \mathrm{id}_E)^{\circ(m-1)}(v)$ est un vecteur propre de f de valeur propre associée λ. En particulier, si E est de dimension finie alors λ est une racine du polynôme caractéristique (et du polynôme minimal) de f.*

Définition 1.22 (sous-espace caractéristique)**.** Le sous-espace caractéristique de f associé à $\lambda \in \mathrm{Sp}(f)$ est l'ensemble

$$K_\lambda(f) := \bigcup_{m \geq 0} \ker\left((f - \lambda \mathrm{id}_E)^{\circ m}\right)$$

des vecteurs propres généralisés de poids λ de f.

Proposition 1.23. *Pour tout $\lambda \in \mathrm{Sp}(f)$, $K_\lambda(f)$ est un sous-espace vectoriel de E qui est stable par f. De plus, si E est de dimension finie alors $f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)}$ est nilpotent (d'ordre de nilpotence $q(\lambda) \leq k(\lambda) := \dim\big(K_\lambda(f)\big)$).*

Pour démontrer cette Proposition, nous allons utiliser le résultat qui suit :

LEMME 1.24 (lemme des noyaux emboités). *Si $\phi \in \mathrm{End}(E)$ et $P, Q \in \mathbb{K}[X]$ alors $\ker\big(Q(\phi)\big) \subset \ker\big((PQ)(\phi)\big)$. De plus, si $\ker\big(Q(\phi)\big) = \ker\big((PQ)(\phi)\big)$ alors $\ker\big(Q(\phi)\big) = \ker\big((P^kQ)(\phi)\big)$ quel que soit $k \in \mathbb{N}$. En particulier :*

- *Si $P(X) = X$ on obtient que $\ker\big(Q(\phi)\big)$ est stable par ϕ.*
- *Si $P(X) = X$ et $Q(X) = X^m$ on obtient que $\ker(\phi^{\circ m}) \subset \ker(\phi^{\circ(m+1)})$. De plus, si $\ker(\phi^{\circ m}) = \ker(\phi^{\circ(m+1)})$ alors $\ker(\phi^{\circ m}) = \ker(\phi^{\circ(m+k)})$ quel que soit $k \in \mathbb{N}$. Une telle égalité se produit forcément pour $m \leq n = \dim(E)$ si E est de dimension finie.*

DÉMONSTRATION DU LEMME. Il n'y a pas grand chose à démontrer. Nous avons déjà vu que $\ker\big(Q(\phi)\big) \subset \ker\big((PQ)(\phi)\big)$ lors de la démonstration du lemme des noyaux (Lemme 1.16). Les affirmations qui restent sont des conséquences immédiates de ce fait (à vérifier en exercice). $\square$

DÉMONSTRATION DE LA PROPOSITION. En appliquant d'abord le lemme à $\phi = f$ et $Q(X) = (X - \lambda)^m$ on trouve que chaque sous-espace vectoriel $\ker\big((f - \lambda \mathrm{id}_E)^{\circ m}\big)$ est stable par f. Leur réunion l'est donc aussi.
En appliquant ensuite le lemme à $\phi = f - \lambda \mathrm{id}_E$, on trouve que

$$\ker\left((f - \lambda \mathrm{id}_E)^{\circ m}\right) \subset \ker\left((f - \lambda \mathrm{id}_E)^{\circ(m+1)}\right),$$

d'où l'on déduit que $K_\lambda(f)$ est un sous-espace vectoriel car réunion croissante de sous-espaces vectoriels [2].
Enfin, pour tout $v \in K_\lambda(f)$, par définition il existe $m_v \in \mathbb{N}$ tel que $v \in \ker\big((f - \lambda \mathrm{id}_E)^{\circ m_v}\big)$. Si E de dimension finie alors on choisit une base $\mathcal{B} = (v_1, \ldots, v_{k(\lambda)})$ de $K_\lambda(f)$ et on pose $m = \max(m_{v_i} \,|\, 1 \leq i \leq n)$. On a ainsi

$$\ker\left((f - \lambda \mathrm{id}_E)^{\circ m}\right) = K_\lambda(f).$$

2. On rappelle (HAX202X) que si $F, G \subset E$ sont des sous-espaces vectoriels, $F \cup G$ n'est est un que si $F \subset G$ ou $G \subset F$.

Comme $K_\lambda(f)$ est stable par f on a aussi

$$\ker\left(f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)})^{\circ m}\right) = K_\lambda(f)\,,$$

de sorte qu'en appliquant le lemme à $\phi = f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)}$ on trouve $q(\lambda) \leq k(\lambda)$. [On rappelle que $q(\lambda)$ est l'ordre de nilpotence de $f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)}$; c'est donc le plus petit m tel que $\ker\left(f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)})^{\circ m}\right) = K_\lambda(f)$.] □

On suppose désormais que E est de dimension finie.

Corollaire 1.25. *Pour tout $\lambda \in \mathrm{Sp}(f)$:*

(1) L'ordre de nilpotence $q(\lambda)$ est la multiplicité de λ dans le polynôme minimal μ_f ;

(2) La dimension $k(\lambda)$ de $K_\lambda(f)$ est la multiplicité algébrique de λ.

DÉMONSTRATION. On commence par démontrer le point (1). Notons $p(\lambda)$ l'ordre de multiplicité de λ dans le polynôme minimal : $\mu_f(X) = (X-\lambda)^{p(\lambda)}Q(X)$, où $Q(X)$ est premier avec $(X-\lambda)^{p(\lambda)}$, et donc avec toutes les puissances non nuls de $X-\lambda$. D'après le lemme des noyaux (Lemme 1.16) on donc

$$E = \ker(\mathbf{0}) = \ker\left(\mu_f(f)\right) = \ker\left((f-\lambda \mathrm{id}_E)^{\circ p(\lambda)}\right) \oplus \ker\left(Q(f)\right) \subset K_\lambda(f) \oplus \ker\left(Q(f)\right).$$

Par conséquent $\ker\left((f-\lambda \mathrm{id}_E)^{\circ p(\lambda)}\right) = K_\lambda(f)$ et donc $q(\lambda) \leq p(\lambda)$. On choisit une base $\mathcal{B}_1 = (v_1, \dots, v_{k(\lambda)})$ de $K_\lambda(f)$ et une base une base $\mathcal{B}_2 = (v_{k(\lambda)+1}, \dots, v_n)$ de $\ker\left(Q(f)\right)$, ce qui nous donne une base $\mathcal{B} = \mathcal{B}_1 \sqcup \mathcal{B}_2 = (v_1, \dots, v_n)$ de E. Comme $K_\lambda(f)$ et $\ker\left(Q(f)\right)$ sont stables par f, la matrice $M := M_{\mathcal{B},\mathcal{B}}(f)$ de f dans cette base est diagonale par blocs : $M = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$. On remarque que $(X-\lambda)^{q(\lambda)}Q(X)$ annule M, car $(X-\lambda)^{q(\lambda)}$ annule A et Q annule D. Par minimalité du polynôme minimal, on obtient que $p(\lambda) = q(\lambda)$.
Démontrons maintenant le point (2). En utilisant la formule du polynôme caractéristique pour les matrices diagonales par blocs on trouve que $\chi_f(X) = \chi_M(X) = \chi_A(X)\chi_D(X)$. Comme A est annulée par une puissance de $X-\lambda$, son polynôme polynôme minimal est une puissance de $X-\lambda$, et il en va de même pour son polynôme caractéristique[3] qui est donc $\chi_A(X) = (X-\lambda)^{k(\lambda)}$. Comme λ n'est pas une racine de Q, qui annule D, λ n'est pas une valeur propre de D et donc pas une racine de χ_D. Par conséquent $k(\lambda)$ est bien la multiplicité de λ dans $\chi_f(X)$. □

Remarque 1.26. Il découle immédiatement du lemme des noyaux que les sous-espaces caractéristiques $\left(K_\lambda(f)\right)_{\lambda \in \mathrm{Sp}(f)}$ sont en somme directe. En effet, les polynômes de la famille $\left((X-\lambda)^{q(\lambda)}\right)_{\lambda \in \mathrm{Sp}(f)}$ sont deux à deux premiers entre eux.

Exercice 1.27. *Pour les matrices qui suivent (qui ont toutes 0 comme unique valeur propre), calculez $m_a(0)$, $q(0)$ et $m_g(0)$:*

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

3. On reviendra sur ce point dans la partie suivante.

Théorème 1.28 (décomposition en sous-espaces caractéristiques). *Si $f \in \mathrm{End}(E)$ est trigonalisable alors*

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} K_\lambda(f).$$

DÉMONSTRATION. On sait déjà que la somme est directe (cf. Remarque précédente). Si $f \in \mathrm{End}(E)$ est trigonalisable alors son polynôme caractéristique est scindé : $\chi_f(X) = \prod_{\lambda \in \mathrm{Sp}(f)}(X-\lambda)^{k(\lambda)}$. Comme χ_f annule f (c'est le théorème de Cayley–Hamilton),

$$E = \ker\big(\chi_f(f)\big) = \sum_{\lambda \in \mathrm{Sp}(f)} \ker\big((f-\lambda \mathrm{id}_E)^{\circ k(\lambda)}\big).$$

Or nous avons vu dans la Proposition 1.23 et son Corollaire 1.25 que $\ker\big((f-\lambda \mathrm{id}_E)^{\circ k(\lambda)}\big) = K_\lambda(f)$. □

2. Réduction

2.1. Décomposition de Dunford.

Ce paragraphe est dévolu à la démonstration du résultat qui suit :

Théorème 2.1 (décomposition de Dunford). *Si $f \in \mathrm{End}(E)$ est trigonalisable, alors il existe une unique paire $(d,n) \in \mathrm{End}(E) \times \mathrm{End}(E)$ satisfaisant les propriétés qui suivent :*

(1) $f = d + n$;

(2) d est diagonalisable ;

(3) n est nilpotent ;

(4) $d \circ n = n \circ d$ (on dit que d et n commutent).

De plus, d (et donc n) est un polynôme en f ; c'est-à-dire qu'il existe $P \in \mathbb{K}[X]$ tel que $d = P(f)$.

DÉMONSTRATION DE L'EXISTENCE. D'après le Théorème 1.28, $E = \oplus_{\lambda \in \mathrm{Sp}(f)} K_\lambda(f)$. De plus, on a vu dans la Proposition 1.23 que $K_\lambda(f)$ stable par f. On définit $d : E \to E$ de la manière suivante : pour tout $v = \sum_{\lambda \in \mathrm{Sp}(f)} v_\lambda \in E$, avec $v_\lambda \in K_\lambda(f)$ (on rappelle qu'une telle décomposition est unique puisque la somme est directe),

$$d(v) := \sum_{\lambda \in \mathrm{Sp}(f)} \lambda v_\lambda.$$

On pose alors $n := f - d$, de sorte que $f = d + n$. La propriété (1) est vérifiée.
Montrons que la propriété (2) est vérifiée : d est diagonalisable. On remarque que $K_\lambda(f)$ est stable par d quel que soit $\lambda \in \mathrm{Sp}(f)$: en effet, quel que soit $v \in K_\lambda(f)$, $d(v) = \lambda v \in K_\lambda(f)$. De plus $d_{|K_\lambda(f)} = \lambda \mathrm{id}_{K_\lambda(f)}$ est évidemment diagonalisable. Donc d est diagonalisable.
Montrons que la propriété (3) est vérifiée : n est nilpotent. Pour tout $\lambda \in \mathrm{Sp}(f)$, $K_\lambda(f)$ est stable par f et d, donc par $n = f - d$, et $n_{|K_\lambda(f)} = f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)}$ est nilpotent (d'après la Proposition 1.23) d'ordre de nilpotence $q(\lambda)$. Posons $q := \max\big(q(\lambda) \,|\, \lambda \in \mathrm{Sp}(f)\big)$. Alors $n^{\circ q} = \mathbf{0}$.
Montrons finalement que la propriété (4) est vérifiée : $d \circ n = n \circ d$. Si $v \in K_\lambda(f)$, alors

$$d\big(n(v)\big) = \lambda n(v) = n(\lambda v) = n\big(d(v)\big).$$

La première égalité vient du fait que $n(v) \in K_\lambda(f)$, puisque $K_\lambda(f)$ est stable par n. □

Exemple 2.2. La décomposition de Dunford $M = D + N$ de la matrice $M = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ est donnée par $D = \mathrm{I}_2$ et $N = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$. On a bien $M = D + N$, avec D diagonalisable (ici, elle est même diagonale) et N nilpotente (d'ordre de nilpotence 2), et $DN = N = ND$. △

Exemple 2.3 (attention, c'est un piège)**.** La décomposition de Dunford $M = D + N$ de la matrice $M = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ est donnée par $D = M$ et $N = \mathbf{0}$. En effet, D est diagonalisable (puisque son polynôme caractéristique $\chi_M(X) = (X-1)(X-2)$ est scindé à racines simples. Vérifier que $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$ est une base de vecteurs propres de M, et que $P_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $P_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. △

DÉMONSTRATION DE L'UNICITÉ DE LA DÉCOMPOSITION DE DUNFORD. Commençons par montrer que dans la démonstration de l'existence de la décomposition $f = d + n$, l'endomorphisme diagonalisable d est un polyôme en f. Pour cela, on observe dans un premier temps que $d = \sum_{\lambda \in \mathrm{Sp}(f)} \lambda \pi_\lambda$, où $\pi_\lambda : E \to E$ est la projection sur $K_\lambda(f)$ définie comme suit :

$$\pi_\lambda \left(\sum_{\alpha \in \mathrm{Sp}(f)} v_\alpha \right) := v_\lambda \quad \text{si } v_\alpha \in K_\alpha(f)\,.$$

Il suffit donc de démontrer que pour tout $\lambda \in \mathrm{Sp}(f)$, π_λ est un polynôme en f. On considère le polynôme

$$Q_\lambda(X) := \prod_{\alpha \mathrm{Sp}(f) \setminus \{\lambda\}} (X - \alpha)^{q(\alpha)}\,.$$

Comme λ n'est pas une racine de ce polynôme, $Q_\lambda(\lambda) \neq 0$, et ce polynôme admet donc un inverse modulo $(X - \lambda)^k$ quel que soit $k \geq 1$[4]. En prenant $k = q(\lambda)$, on obtient donc un polynôme $R_\lambda(X)$ tel que $P_\lambda(X) := R_\lambda(X) Q_\lambda(X) = 1 \bmod (X - \lambda)^{q(\lambda)}$. Si $v \in K_\alpha(f)$, alors

$$P_\lambda(f)(v) = \begin{cases} 0_E & \text{si } \alpha \neq \lambda \text{ (car } Q_\lambda(f)(v) = 0_E) \\ v & \text{si } \alpha = \lambda \text{ (car } (f - \lambda \mathrm{id})^{\circ q(\lambda)}(v) = 0_E)\,. \end{cases}$$

Autrement dit, $P_\lambda(f) = \pi_\lambda$.

Nous pouvons maintenant passer à la démonstration de l'unicité. On considère la décomposition $f = d + n$ construite lors de la preuve de l'existence. Supposons maintenant que $f = d' + n'$, avec (d', n') vérifiant les propriétés (1) à (4) du théorème de décomposition de Dunford. Comme d' commute avec n', alors d' commute avec f : en effet,

$$d' \circ f = d' \circ d' + d' \circ n' = d' \circ d' + n' \circ d' = f \circ d'\,.$$

Nous avons vu que $d = P(f)$, de sorte que d' commute aussi avec d (à démontrer en exercice). Nous avons deux endomorphismes diagonalisables, d et d', qui commutent. Ils admettent

4. Pour le démontrer on écrit $Q_\lambda(X) = a + (X - \lambda) T_\lambda(X)$, avec $a = Q_\lambda(\lambda) \neq 0$. On choisit alors comme inverse modulo $(X - \lambda)^k$ de $Q_\lambda(X)$ le polynôme

$$\sum_{j=0}^{k-1} (-1)^j a^{-j-1} (X - \lambda)^j T_\lambda(X)^j\,.$$

donc une base commune $\mathcal{B}$ de diagonalisation (il s'agit d'un exercice traité en TD) ; en particulier, dans cette base $\mathcal{B}$, la matrice $d - d'$ est diagonale. Finalement, on observe que n et n' commutent : en effet,

$$\begin{aligned} n \circ n' &= (f-d) \circ (f-d') = f^{\circ 2} - d \circ f - f \circ d' + d \circ d' \\ &= f^{\circ 2} - f \circ d - d' \circ f + d' \circ d = (f-d') \circ (f-d) = n' \circ n \,. \end{aligned}$$

on en déduit que $n' - n$ est également nilpotent (à faire en exercice : la somme de deux endomorphismes nilpotents qui commutent est un endomorphisme nilpotent). Ainsi la matrice de $d - d' = n' - n$ dans la base $\mathcal{B}$ est diagonale et nilpotente. Elle est donc nulle ; on trouve que $d = d'$ et $n = n'$. □

2.2. Réduction des endomorphismes nilpotents.

Soit $f \in \mathrm{End}(E)$ un endomorphisme nilpotent. Tout vecteur $v \in E$ est un vecteur propre généralisé pour la valeur propre 0, dont on note q_v la hauteur (on rappelle que c'est le plus petit entier $k \in \mathbb{N}$ tel que $f^{\circ k}(v) = 0_E$).

LEMME 2.4. *Sous les hypothèses qui précèdent, si $v \neq 0_E$, alors $\big(v, f(v), \ldots, f^{\circ(q_v-1)}(v)\big)$ est une famille libre.*

DÉMONSTRATION. Laissée en exercice : montrer par récurrence sur k que si $k < q_v$ alors $\big(v, f(v), \ldots, f^{\circ k}(v)\big)$ est une famille libre. □

On note $H_v := \mathrm{Vect}\big(v, f(v), \ldots, f^{\circ(q_v-1)}(v)\big)$, qu'on appelle le *sous-espace cyclique* engendré par v. Il est évidemment stable par f. La matrice de $f_{|H_v}$ dans la base $\big(f^{\circ(q_v-1)}(v), \ldots, f(v), v\big)$ est un *bloc de Jordan*

$$J_{0,q_v} := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Théorème 2.5. *Si E est de dimension finie, alors il existe une famille $(v_1, \ldots, v_r)$ de vecteurs de E, où $r = \dim\big(\ker(f)\big)$, telle que*

$$E = \bigoplus_{i=1}^{r} H_{v_i} \,.$$

Ainsi il existe une base de E dans laquelle la matrice de f est constituée de r blocs diagonaux de Jordan $J_{0,q_{v_1}}$, ..., $J_{0,q_{v_r}}$.

DÉMONSTRATION. Admise. □

2.3. Réduction de Jordan.

Soit $f \in \mathrm{End}(E)$, avec E de dimension finie. On peut appliquer ce qui a été fait précédemment sur chacun des sous-espaces caractéristiques $K_\lambda(f)$ de E, à l'endomorphisme nilpotent $f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)}$. Il existe donc une base de $K_\lambda(f)$ dans laquelle la matrice de cet endomorphisme nilpotent est constituté de blocs diagonaux de Jordan de la forme J_{0,q_v}. Dans cette base, la matrice de $f_{|K_\lambda(f)} = \lambda \mathrm{id}_{K_\lambda(f)} + (f_{|K_\lambda(f)} - \lambda \mathrm{id}_{K_\lambda(f)})$ est donc constituée de blocs diagonaux de Jordan de la forme

$$J_{\lambda,q_v} := \lambda \mathrm{I}_{q_v} + J_{0,q_v} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda \end{pmatrix}$$

Si f est trigonalisable, alors E est la somme directe des sous-espaces caractéristiques de f, de sorte qu'il existe une base de E dans laquelle la matrice de f est constitutée de blocs diagonaux de Jordan de la forme J_{λ,q_v}, avec $\lambda \in \mathrm{Sp}(f)$ et $v \in K_\lambda(f)$.

3. Applications

3.1. Suites récurrentes linéaires.

On a déjà vu en introduction de ce cours qu'il est relativement facile de calculer les puissances d'une matrice que l'ont sait diagonaliser.

Exemple 3.1. Revenons à l'Exemple 2.3. On a

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

Ainsi, pour tout $n \in \mathbb{N}$,

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}^n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2^n - 1 \\ 0 & 2^n \end{pmatrix} .$$

△

Pour les matrices trigonalisables, la décomposition de Dunford permet de simplifier les calculs de puissances de matrices. En effet, si $M = D + N$, avec D diagonalisable, N nilpotente d'ordre de nilpotence q, et $DN = ND$, alors on a

$$M^k = (D+N)^k = \sum_{j=0}^{k} \binom{k}{j} N^j D^{k-j} \left(= \sum_{j=0}^{q-1} \binom{k}{j} N^j D^{k-j} \text{ si } q \leq k \right) .$$

Nous avons pu utiliser la formule du binôme car N et D commutent.

Exemple 3.2. Revenons à l'Exemple 2.2, où la décomposition de Dunford est

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2 + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} .$$

Pour tout $k \geq 1$,

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}^k = \binom{k}{0} \mathrm{I}_2 + \binom{k}{1} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2k \\ 0 & 1 \end{pmatrix} .$$

△

On a une variante de la formule, qui nous indique qu'on peut en particulier calculer toute puissance d'un endomorphisme trigonalisable indépendamment sur chaque facteur de la décomposition en sous-espaces caractéristiques (puisque $f = \sum_{\lambda \in \mathrm{Sp}(f)} f \circ \pi_\lambda$) :

Proposition 3.3. *Si $f \in \mathrm{End}(E)$ est trigonalisable, alors pour tout $k \in \mathbb{N}$,*

$$f^{\circ k} = \sum_{\lambda \in \mathrm{Sp}(f)} \left(\sum_{j=0}^{\min\left(k, q(\lambda)-1\right)} \binom{k}{j} \lambda^{k-j} (f - \lambda \mathrm{id})^{\circ j} \right) \circ \pi_\lambda \,.$$

DÉMONSTRATION. En effet, il suffit de vérifier cette identité sur chaque sous-espace caractéristique : pour $v \in K_\lambda(f)$, on doit donc vérifier que

$$f^{\circ k}(v) = \sum_{j=0}^{\min\left(k, q(\lambda-1)\right)} \binom{k}{j} \lambda^{k-j} (f - \lambda \mathrm{id})^{\circ j}(v) \,.$$

Cela découle de la formule du binôme pour la somme $f = \lambda \mathrm{id} + (f - \lambda \mathrm{id})$, et du fait que $(f - \lambda \mathrm{id})^{\circ q(\lambda)}(v) = 0_E$ lorsque $v \in K_\lambda(f)$. □

Remarque 3.4. En particulier, si $k \geq \dim(E)$, alors $q(\lambda) \leq k$ et la formule devient

$$f^{\circ k} = \sum_{\lambda \in \mathrm{Sp}(f)} \left(\sum_{j=0}^{q(\lambda)-1} \binom{k}{j} \lambda^{k-j} (f - \lambda \mathrm{id})^{\circ j} \right) \circ \pi_\lambda \,.$$

Si de plus 0 n'est pas une valeur propre de f, alors

$$f^{\circ k} = \sum_{\lambda \in \mathrm{Sp}(f)} \lambda^k \left(\sum_{j=0}^{q(\lambda)-1} \frac{\binom{k}{j}}{\lambda^j} (f - \lambda \mathrm{id})^{\circ j} \circ \pi_\lambda \right) \,.$$

Dans cette dernière formule, l'expression à l'intérieur des parenthèses ne dépend pas de k, sauf pour le terme $\binom{k}{j} = \frac{k(k-1)\cdots(k-j+1)}{j!}$ qui est un polynôme en k de degré $j < q(\lambda)$.

Théorème 3.5. *On considère le polynôme $P(X) := X^N - a_{N-1}X^{N-1} - \cdots - a_1 X - a_0 \in \mathbb{K}[X]$. On suppose que $P(X)$ est scindé et que $a_0 \neq 0$, et on note $k(\lambda)$ la multiplicité d'une racine λ de $P(X)$. Alors l'ensemble des suites $(u_n)_{n \geq 0} \in \mathbb{K}^{\mathbb{N}}$ satisfaisant la relation de récurrence*

$$u_{n+N} = a_0 u_n + a_1 u_{n+1} + \cdots + a_{N-1} u_{n+N-1} \tag{$\star$}$$

est égal à l'ensemble des suites dont le terme général est de la forme

$$\sum_{\lambda \text{ racine de } P} P_\lambda(n) \lambda^n \,,$$

avec $P_\lambda(X) \in \mathbb{K}[X]_{<k(\lambda)}$.

On remarque que, comme $a_0 \neq 0$, 0 n'est pas une racine de P.

DÉMONSTRATION. On procède par double inclusion.
$\boxed{\supset}$ Supposons que $u_n = n^k \lambda^n$, avec $\lambda \in \mathbb{K}$ racine de P. Montrons par récurrence sur $k \in \mathbb{N}$ qu'il existe des polynômes $Q_{0,n,k}, \ldots, Q_{k,n,k} \in \mathbb{K}[X]$ tels que

$$u_{n+N} - \sum_{j=0}^{N-1} a_j u_{n+j} = \sum_{i=0}^{k} Q_{i,n,k}(\lambda) P^{(i)}(\lambda) \,.$$

<u>Initialisation</u> : pour $k = 0$,

$$\lambda^{n+N} - \sum_{j=0}^{N-1} a_j \lambda^{n+j} = P(\lambda) \qquad (Q_{0,n,0}(X) = \lambda^n) \,.$$

Hérédité : supposons la propriété vérifiée au rang k, et calculons :

$$\begin{aligned}(n+N)^{k+1}\lambda^{n+N}-\sum_{j=0}^{N-1}a_j(n+j)^{k+1}\lambda^{n+j} &= \lambda\left((n+N)^k\lambda^{n+N}-\sum_{j=0}^{N-1}a_j(n+j)^k\lambda^{n+j}\right)' \\ &= \lambda\left(\sum_{i=0}^{k}Q_{i,n,k}(\lambda)P^{(i)}(\lambda)\right)' \\ &= \sum_{i=0}^{k}\left(Q'_{i,n,k}(\lambda)P^{(i)}(\lambda)+Q_{i,n,k}(\lambda)P^{(i+1)}(\lambda)\right) \\ &= \sum_{i=0}^{k+1}Q_{i,n,k+1}(\lambda)P^{(i)}(\lambda)\,,\end{aligned}$$

avec $Q_{0,n,k+1}=Q'_{0,n,k}$, $Q_{1,n,k+1}=Q'_{1,n,k}+Q_{0,n,k}$, ..., $Q_{k,n,k+1}=Q'_{k,n,k}+Q_{k-1,n,k}$, et $Q_{k+1,n,k+1}=Q_{k,n,k}$.
On en déduit que si $k<k(\lambda)$, alors

$$(n+N)^{k+1}\lambda^{n+N}-\sum_{j=0}^{N-1}a_j(n+j)^{k+1}\lambda^{n+j}=\sum_{i=0}^{k}Q_{i,n,k}(\lambda)\underbrace{P^{(i)}(\lambda)}_{=0}=0\,,$$

de sorte que $(n^k\lambda^n)_{n\geq 0}$ satisfait la relation de récurrence $(\star)$. L'ensemble des solutions de la relation de récurrence $(\star)$ étant un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$, toute combinaison linéaire

$$\sum_{\lambda \text{ racine de } P}\ \sum_{i=0}^{k(\lambda)-1}a_{i,\lambda}n^i\lambda^n$$

de telles solutions est encore une solution.
$\boxed{\subset}$ On sait que la suite scalaire $(u_n)_{n\geq 0}\in\mathbb{K}^{\mathbb{N}}$ satisfait la relation de récurrence $(\star)$ si et seulement si la suite vectorielle $(X_n)_{n\geq 0}\in(\mathbb{K}^N)^{\mathbb{N}}$ définie par $X_n:=\begin{pmatrix}u_n\\u_{n+1}\\\vdots\\u_{n+N-1}\end{pmatrix}$ satisfait la relation de récurrence $X_{n+1}=CX_n$, avec

$$C=\begin{pmatrix}0&1&0&\cdots&0\\\vdots&0&\ddots&\ddots&\vdots\\\vdots&\vdots&\ddots&\ddots&0\\0&0&\cdots&0&1\\a_0&a_1&\cdots&a_{N-2}&a_{N-1}\end{pmatrix}$$

(il s'agit d'un exercice traité en TD). Les valeurs propres de C sont les racines de $P(X)$.
Posons $V:=C^{-N}X_0$ (on rappelle que C est inversible car $a_0\neq 0$), de sorte que $X_n=C^{n+N}V\in\mathbb{K}^N$, et considérons la décomposition (unique) de V dans les sous-espaces caractéristique : $V=\sum_{\lambda\in\mathrm{Sp}(C)}V_\lambda$, avec $V_\lambda\in K_\lambda(C)$. D'après la Proposition 3.3 et la Remarque

3.4, on a

$$X_n = C^{n+N}V = \sum_{\lambda \in \mathrm{Sp}(C)} \lambda^{n+N} \left(\sum_{j=0}^{q(\lambda)-1} \frac{\binom{n+N}{j}}{\lambda^j} (C - \lambda \mathrm{I}_N)^j V_\lambda \right) .$$

Ainsi, u_n étant la première coordonnée de X_n, on trouve que u_n est bien de la forme souhaitée (toujours d'après la Remarque 3.4). □

Exemple 3.6 (suite de Fibonacci)**.** On a vu dans l'introduction que la suite de Fibonacci, définie par la relation de récurrence $u_{n+2} = u_n + u_{n+1}$, $u_0 = 0$ et $u_1 = 1$, a pour terme général

$$u_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n .$$

Ici, $\frac{1\pm\sqrt{5}}{2}$ sont les racines du trinôme $X^2 - X - 1$. △

Exemple 3.7 (récurrence d'ordre 1)**.** Si $u_{n+1} = au_n$, avec $a \neq 0$, alors le polynôme est $P(X) = X - a$, ayant pour unique racine a, qui est d'ordre 1. Dans ce cas $(u_n)_{n\geq 0}$ est une suite géométrique de raison a et son terme général est $u_n = u_0 a^n$. △

Exemple 3.8 (récurrence d'ordre 2)**.** Si $u_{n+2} = au_n + bu_{n+1}$, alors le polynôme est $P(X) = X^2 - bX - a$. On suppose P scindé et $a \neq 0$. Il y a deux possibilités :

(1) Soit il y a deux racines distinctes $\lambda_1 \neq \lambda_2$. Dans ce cas il existe des constantes $\alpha_1, \alpha_2 \in \mathbb{K}$ telles que

$$u_n = \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n .$$

Les constantes sont déterminées par u_0 et u_1 : $\alpha_1 + \alpha_2 = u_0$ et $\alpha\lambda_1 + \alpha_2\lambda_2 = u_1$. À titre d'exercice, déterminer les constantes pour la variante de la suite de Fibonacci où $u_0 = 2$ et $u_1 = 1$.

(2) Soit il y a une racine double λ. Dans ce cas il existe des constantes $\alpha, \beta \in \mathbb{K}$ telles que

$$u_n = (\alpha + n\beta)\lambda^n .$$

Ici encore les constantes sont déterminées par u_0 et u_1 : $\alpha = u_0$ et $(\alpha + \beta)\lambda = u_1$. Comme $\lambda \neq 0$ (puique $a \neq 0$), alors $\alpha = u_1$ et $\beta = u_1/\lambda - u_0$. △

3.2. Exponentielle matricielle.

Plaçons-nous sur $\mathbb{C}$ et commençons par quelques faits que nous admettrons (dans la suite, toutes les matrices sont supposées carrées pour plus de simplicité) :

— Par définition, une suite de matrices $(A_n)_{n\geq 0}$ converge vers une matrice A (quand n tend vers $+\infty$) si pour tout (i, j) la suite des coefficients $a_{n,ij}$ converge vers a_{ij}.

— On a une norme sur l'espace vectoriel $\mathrm{Mat}_{N\times N}(\mathbb{C})$, définie par

$$\|A\| := \max\Big(\sum_{j=1}^{N} |a_{ij}| \, , \, 1 \leq i \leq N \Big) .$$

Elle satisfait les propriétés qui suivent :

(1) $\|A\| = 0$ si et seulement si $A = \mathbf{0}$.

(2) $\|A + B\| \leq \|A\| + \|B\|$.

(3) $\|AB\| \leq \|A\| \, \|B\|$.

(4) Si λ est un scalaire alors $\|\lambda A\| = |\lambda| \, \|A\|$.

- Une suite de matrice $(A_n)_{n\geq 0}$ converge vers une matrice A si et seulement si la suite numérique $\big(\|A_n - A\|\big)_{n\geq 0}$ converge vers 0.
- Si la série numérique $\sum_{k=0}^{+\infty} \|A_k\|$ converge alors la série $\sum_{k=0}^{+\infty} A_k$ converge[5].

Proposition 3.9. *Pour toute matrice $A \in \mathrm{Mat}_{N\times N}(\mathbb{C})$, la série*

$$\sum_{k=0}^{+\infty} \frac{A^k}{k!}$$

converge. On note $\exp(A)$ la limite de cette série.

DÉMONSTRATION. Pour tout nombre réel x, la série $\sum_{k=0}^{+\infty} \frac{x^k}{k!}$ converge (vers e^x). On conclut en posant $x = \|A\|$ (et en utilisant que $\|A^k\| \leq \|A\|^k$). $\square$

Exercice 3.10. *Vérifier que si $D = diag(\lambda_1, \dots, \lambda_N)$ alors $\exp(D) = diag(e^{\lambda_1}, \dots, e^{\lambda_N})$. En particulier, $\exp(\mathbf{0}) = \mathrm{I}_N$.*

On admettra que la propriétés qui suit est satisfaite par l'exponentielle :
- Si A et B commutent, alors $\exp(A)\exp(B) = \exp(A+B)$.

Par conséquent $\exp(A)$ est inversible, d'inverse $\exp(-A)$.

Exemple 3.11. Considérons la matrice $M = \begin{pmatrix} \lambda & z \\ 0 & \lambda \end{pmatrix} = \lambda \mathrm{I}_2 + \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}$, où $\lambda, z \in \mathbb{C}$. On trouve donc que

$$\exp(M) = \exp(\lambda \mathrm{I}_2)\exp\left(\begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}\right) = e^{\lambda} \cdot \exp\left(\begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}\right).$$

Or $\exp\left(\begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}\right) = \mathrm{I}_2 + \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}$, et ainsi $\exp(M) = \begin{pmatrix} e^{\lambda} & ze^{\lambda} \\ 0 & e^{\lambda} \end{pmatrix}$. $\triangle$

Proposition 3.12. *Si $P \in GL_N(\mathbb{C})$ et $A \in \mathrm{Mat}_{N\times N}(\mathbb{C})$, alors $\exp(P^{-1}AP) = P^{-1}\exp(A)P$. On en déduit que $\det\big(\exp(A)\big) = e^{\mathrm{tr}(A)}$.*

DÉMONSTRATION. La première affirmation est une conséquence immédiate du fait que $(P^{-1}AP)^n = P^{-1}A^nP$ quel que soit $n \in \mathbb{N}$.
Pour la seconde affirmation, on trigonalise A : il existe une matrice inversible P et une matrice triangluaire supérieure T telles que $P^{-1}AP = T$. Ainsi $\det\big(\exp(T)\big) = \det\big(P^{-1}\exp(A)P\big) = \det\big(\exp(A)\big)$. Comme $\mathrm{tr}(T) = \mathrm{tr}(A)$, il suffit de démontrer la seconde affirmation pour une matrice triangulaire supérieure T. Montrer en exercice que $\exp(T)$ est triangulaire supérieure, et que si λ_i est le i-ème élément diagonal de T alors le i-ème élément diagonal de $\exp(T)$ est e^{λ_i}. On en déduit que

$$\det\big(\exp(T)\big) = \prod_{i=1}^{N} e^{\lambda_i} = e^{\sum_{i=1}^{N}\lambda_i} = e^{\mathrm{tr}(T)}.$$

$\square$

Exemple 3.13. Revenons à l'Exemple 2.3. On a

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

5. En effet, pour tout (i,j), on a $|a_{k,i,j}| \leq \|A_k\|$. Ainsi, si $\sum_{k=0}^{+\infty} \|A_k\|$ converge, alors $\sum_{k=0}^{+\infty} |a_{k,i,j}|$ converge, et il en va de même pour $\sum_{k=0}^{+\infty} |a_{k,i,j}|$.

Ainsi, $\exp\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} e & 0 \\ 0 & e^2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e & e^2 \\ 0 & e^2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e & e^2-e \\ 0 & e^2 \end{pmatrix}$. △

Exemple 3.14. On considère la matrice $M_t := \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = t\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, où $t \in \mathbb{R}$. Elle est diagonalisable sur $\mathbb{C}$:

$$M_t = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\begin{pmatrix} it & 0 \\ 0 & -it \end{pmatrix}\begin{pmatrix} -i/2 & 1/2 \\ 1/2 & -i/2 \end{pmatrix}$$

Par conséquent

$$\begin{aligned} \exp(M_t) &= \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\exp\left(\begin{pmatrix} it & 0 \\ 0 & -it \end{pmatrix}\right)\begin{pmatrix} -i/2 & 1/2 \\ 1/2 & -i/2 \end{pmatrix} \\ &= \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}\begin{pmatrix} -i/2 & 1/2 \\ 1/2 & -i/2 \end{pmatrix} \\ &= \begin{pmatrix} (e^{it}+e^{-it})/2 & i(e^{it}-e^{-it})/2 \\ i(e^{-it}-e^{it})/2 & (e^{it}+e^{-it})/2 \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \end{aligned}$$

On a donc $\det\big(\exp(M_t)\big) = \cos(t)^2 + \sin(t)^2 = 1 = e^0 = e^{\mathrm{tr}(M_t)}$. Vérifier en exercice qu'en appliquant directement la définition de l'exponentielle matricielle, on retrouve le développement en série de $\cos(t)$ et $\sin(t)$.[6] △

Nous allons maintenant voir une formule générale pour l'exponentielle d'une matrice $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, qui s'appuie sur la décomposition de Dunford et les projections sur les sous-espaces caractéristiques. Pour toute valeur propre $\lambda \in \mathrm{Sp}(A)$, on note Π_λ la matrice (dans la base canonique) de la projection

$$\begin{aligned} \pi_\lambda : \mathbb{C}^n = \bigoplus_{\mu\in\mathrm{Sp}(A)} K_\mu(A) &\longrightarrow \mathbb{C}^n \\ \sum_\mu \underbrace{v_\mu}_{\in K_\mu(A)} &\longmapsto v_\lambda \end{aligned}$$

sur le sous-espace caractéristique $K_\lambda(A)$. On rappelle que Π_λ est un polynôme en A, et qu'on a la décomposition de Dunford

$$A = \underbrace{\sum_{\lambda\in\mathrm{Sp}(A)} \lambda\Pi_\lambda}_{=D} + \underbrace{\sum_{\lambda\in\mathrm{Sp}(A)} (A-\lambda\mathrm{I}_n)\Pi_\lambda}_{=N}.$$

Proposition 3.15. *Pour tout $t \in \mathbb{C}$,*

$$\exp(tA) = \sum_{\lambda\in\mathrm{Sp}(A)} e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)\Pi_\lambda.$$

DÉMONSTRATION. On procède par étapes :

▷ *Étape 1* : calculons dans un premier temps $\exp(tD)$. Les observations suivantes sont des conséquences directes de la définition de Π_λ :

— $\mathrm{I}_n = \sum_{\lambda\in\mathrm{Sp}(A)} \Pi_\lambda$;

6. Indication : on peut observer que $M_t^2 = -t^2\mathrm{I}_2$, et en déduire une formule pour M_t^n lorsque n est pair et lorsque n est impair.

— $\Pi_\lambda^2 = \Pi_\lambda$;
— si $\lambda \neq \mu$ alors $\Pi_\lambda \Pi_\mu = \mathbf{0} = \Pi_\mu \Pi_\lambda$.

Ainsi, d'une part

$$\exp(tD) = \exp\left(\sum_{\lambda \in \mathrm{Sp}(A)} t\lambda \Pi_\lambda\right) = \prod_{\lambda \in \mathrm{Sp}(A)} \exp(t\lambda\Pi_\lambda)\,,$$

et d'autre part

$$\exp(t\lambda\Pi_\lambda) = \mathrm{I_n} + \sum_{k=1}^{+\infty} \frac{t^k\lambda^k}{k!}\Pi_\lambda^k = \mathrm{I_n} + \sum_{k=1}^{+\infty} \frac{t^k\lambda^k}{k!}\Pi_\lambda = \mathrm{I_n} + (e^{t\lambda}-1)\Pi_\lambda\,.$$

Par conséquent

$$\exp(tD) = \prod_{\lambda \in \mathrm{Sp}(A)} \big(\mathrm{I_n} + (e^{t\lambda}-1)\Pi_\lambda\big) = \mathrm{I_n} + \sum_{\lambda \in \mathrm{Sp}(A)} (e^{t\lambda}-1)\Pi_\lambda = \sum_{\lambda \in \mathrm{Sp}(A)} e^{t\lambda}\Pi_\lambda\,.$$

$\triangleright$ *Étape 2* : passons maintenant au calcul de $\exp(tN)$. Comme $K_\lambda(A)$ est stable par l'endomorphisme canoniquement associé à A, $A\Pi_\lambda = \Pi_\lambda A$. Par conséquent, si $\lambda \neq \mu$, alors

$$(A-\lambda \mathrm{I}_n)\Pi_\lambda(A-\mu\mathrm{I}_n)\Pi_\mu = (A-\lambda\mathrm{I}_n)(A-\mu\mathrm{I}_n)\Pi_\lambda\Pi_\mu = \mathbf{0} = (A-\mu\mathrm{I}_n)\Pi_\mu(A-\lambda\mathrm{I}_n)\Pi_\lambda\,.$$

Ainsi, d'une part

$$\exp(tN) = \exp\left(\sum_{\lambda \in \mathrm{Sp}(A)} t(A-\lambda\mathrm{I}_n)\Pi_\lambda\right) = \prod_{\lambda \in \mathrm{Sp}(A)} \exp\big(t(A-\lambda\mathrm{I}_n)\Pi_\lambda\big)\,,$$

et d'autre part

$$\exp\big(t(A-\lambda\mathrm{I}_n)\Pi_\lambda\big) = \mathrm{I_n} + \sum_{k=1}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda\,.$$

Par conséquent

$$\begin{aligned}
\exp(tN) &= \prod_{\lambda \in \mathrm{Sp}(A)} \left(\mathrm{I_n} + \sum_{k=1}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda\right) = \mathrm{I_n} + \sum_{\lambda \in \mathrm{Sp}(A)} \sum_{k=1}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda \\
&= \sum_{\lambda \in \mathrm{Sp}(A)} \sum_{k=0}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda \qquad \left(\text{car } \mathrm{I}_n = \sum_{\lambda \in \mathrm{Sp}(A)} \Pi_\lambda\right)\,.
\end{aligned}$$

$\triangleright$ *Étape 3* : pour finir, comme D et N commutent, $\exp(tA) = \exp(tD+tN) = \exp(tD)\exp(tN)$. Ainsi

$$\begin{aligned}
\exp(tA) &= \left(\sum_{\lambda \in \mathrm{Sp}(A)} e^{t\lambda}\Pi_\lambda\right)\left(\sum_{\lambda \in \mathrm{Sp}(A)} \sum_{k=0}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda\right) \\
&= \sum_{\lambda \in \mathrm{Sp}(A)} e^{t\lambda} \sum_{k=0}^{q(\lambda)-1} \frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\Pi_\lambda\,.
\end{aligned}$$

$\square$

3.3. Équations différentielles linéaires (à coefficients constants).

Pour une fonction à valeur matricielle $A = (a_{ij})_{i,j} : \mathbb{R} \to \mathrm{Mat}_{p\times q}(\mathbb{C})$ dont chaque coefficient est dérivable, on note $A' = (a'_{ij})_{i,j} : \mathbb{R} \to \mathrm{Mat}_{p\times q}(\mathbb{C})$.

Exercice 3.16.

(1) *Vérifier que si $A, B : \mathbb{R} \to \mathrm{Mat}_{p\times q}(\mathbb{C})$ sont dérivables (au sens où chaque coefficient est dérivable), alors $A + B$ l'est aussi et $(A + B)' = A' + B'$.*

(2) *Vérifier que si $A : \mathbb{R} \to \mathrm{Mat}_{p\times q}(\mathbb{C})$ et $B : \mathbb{R} \to \mathrm{Mat}_{q\times r}(\mathbb{C})$ sont dérivables, alors $AB : \mathbb{R} \to \mathrm{Mat}_{p\times r}(\mathbb{C})$ l'est aussi et $(AB)' = A'B + AB'$.*

Proposition 3.17. *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. La fonction*

$$\begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathrm{Mat}_{n\times n}(\mathbb{C}) \\ t & \longmapsto & \exp(tA) \end{array}$$

est dérivable et $\exp(tA)' = A\exp(tA) = \exp(tA)A$.

DÉMONSTRATION. D'après la Proposition 3.15,

$$\exp(tA) = \sum_{\lambda\in\mathrm{Sp}(A)} e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)\Pi_\lambda$$

est dérivable (car tous ses facteurs le sont). Ainsi, on calcule :

$$\begin{aligned}
\exp(tA)' &= \sum_{\lambda\in\mathrm{Sp}(A)}\left(e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)\right)'\Pi_\lambda \\
&= \sum_{\lambda\in\mathrm{Sp}(A)}\left(\lambda e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)+e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-2}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^{k+1}\right)\right)\Pi_\lambda \\
&= \sum_{\lambda\in\mathrm{Sp}(A)}\left(\lambda e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)+e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^{k+1}\right)\right)\Pi_\lambda \\
&= \sum_{\lambda\in\mathrm{Sp}(A)}\left(\lambda e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)+(A-\lambda\mathrm{I}_n)e^{t\lambda}\left(\sum_{k=0}^{q(\lambda)-1}\frac{t^k}{k!}(A-\lambda\mathrm{I}_n)^k\right)\right)\Pi_\lambda \\
&= A\exp(tA)\,.
\end{aligned}$$

La troisième égalité vient du fait que $(A - \lambda\mathrm{I}_n)^{q(\lambda)} = \mathbf{0}$. Par ailleurs, toujours d'après la Proposition 3.15, $\exp(tA)$ est un polynôme en A, de sorte que $A\exp(tA) = \exp(tA)A$. □

Théorème 3.18. *Soit $Y : \mathbb{R} \to \mathbb{C}^n$ une application dérivable, et soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Alors $Y' = AY$ si et seulement si $Y(t) = \exp(tA)Y(0)$.*

DÉMONSTRATION. On procède par double implication.

$\boxed{\Leftarrow}$ C'est un simple calcul, qui utilise la Proposition 3.17 :

$$\Big(\exp(tA)Y(0)\Big)' = \exp(tA)'Y(0) = A\exp(tA)Y(0)\,.$$

$\boxed{\Rightarrow}$ Supposons que $Y' = AY$ et posons $Z(t) := \exp(-tA)Y(t)$. Ici encore, la Proposition 3.17 nous dit que $Z'(t) = -\exp(-tA)AY(t) + \exp(-tA)AY(t) = \mathbf{0}$. Par conséquent Z est constante, et alors $\exp(-tA)Y(t) = Z(t) = Z(0) = Y(0)$, de sorte que $Y(t) = \exp(tA)Y(0)$. □

Corollaire 3.19. *On considère le polynôme $P(X) := X^N - a_{N-1}X^{N-1} - \cdots - a_1 X - a_0 \in \mathbb{C}[X]$. On suppose que $a_0 \neq 0$, et on note $k(\lambda)$ la multiplicité d'une racine λ de $P(X)$. Alors l'ensemble des fonctions N fois dérivables $y : \mathbb{R} \to \mathbb{C}$ satisfaisant la relation de récurrence*

$$(\star\star) \qquad y^{(N)} = a_0 y + a_1 y' + \cdots + a_{N-1} y^{(N-1)}$$

est égal à l'ensemble des fonctions de la forme

$$t \longmapsto \sum_{\lambda \text{ racine de } P} P_\lambda(t) e^{\lambda t}$$

avec $P_\lambda(X) \in \mathbb{C}[X]_{<k(\lambda)}$.

DÉMONSTRATION. On procède par double inclusion.

$\boxed{\subset}$ On commence par remarquer que l'application qui envoie y vers $Y := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(N-1)} \end{pmatrix}$ établit une bijection entre l'ensemble des solutions de l'équation différentielle ($\star\star$) et l'ensemble des solutions de l'équation $Y' = CY$ avec

$$C = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{N-2} & a_{N-1} \end{pmatrix}$$

la matrice compagnon du polynôme P (vous pouvez le vérifier en exercice). Or d'après le Théorème 3.18, $Y' = CY$ si et seulement si $Y(t) = \exp(tC)V$ avec $V \in \mathbb{C}^n$. Par ailleurs $\chi_C = \mu_C = P$, de sorte que, d'après la Proposition 3.15,

$$Y(t) = \exp(tC)V = \sum_{\lambda \text{ racine de } P} \underbrace{e^{t\lambda} \left(\sum_{k=0}^{q(\lambda)-1} \frac{t^k}{k!} (A - \lambda \mathrm{I}_n)^k \right) \Pi_\lambda V}_{\text{Les coordonnées de ce vecteur sont des polynômes en } t \text{ de degré } <k(\lambda)} .$$

$\boxed{\supset}$ Il suffit de démontrer que si λ est racine de P et que $deg(P_\lambda) < k(\lambda)$ alors $y(t) = P_\lambda(t) e^{\lambda t}$ est solution de l'équation différentielle ($\star\star$). On observe que

$$y^{(k)}(t) = \sum_{j=0}^{k} \binom{k}{j} P_\lambda^{(j)}(t) \lambda^{k-j} e^{\lambda t} \qquad \left(\underline{\text{rappel}} : (fg)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} f^{(j)} g^{(k-j)} \right) .$$

Par conséquent, si on pose $a_N := -1$, on obtient

$$\begin{aligned} \sum_{k=0}^{N} a_k y^{(k)}(t) &= \sum_{k=0}^{N} a_k \left(\sum_{j=0}^{k} \binom{k}{j} P_\lambda^{(j)}(t) \lambda^{k-j} e^{\lambda t} \right) = \sum_{j=0}^{N} P_\lambda^{(j)}(t) e^{\lambda t} \left(\sum_{k=j}^{N} a_k \binom{k}{j} \lambda^{k-j} \right) \\ &= \sum_{j=0}^{N} \frac{1}{j!} P_\lambda^{(j)}(t) e^{\lambda t} \left(\sum_{k=j}^{N} a_k \frac{k!}{(k-j)!} \lambda^{k-j} \right) = \sum_{j=0}^{N} \frac{1}{j!} P_\lambda^{(j)}(t) e^{\lambda t} P^{(j)}(\lambda) . \end{aligned}$$

Si $j \geq k(\lambda)$ alors $P_\lambda^{(j)}(t) = 0$ (car $deg(P_\lambda) < k(\lambda)$), et si $j < k(\lambda)$ alors $P^{(j)}(\lambda) = 0$ (car $k(\lambda)$ est la multiplicité de la racine λ dans P). Donc $\sum_{k=0}^{N} a_k y^{(k)}(t) = 0$, ce qui équivaut à dire que y satisfait l'équation différentielle ($\star\star$). $\square$

Exemple 3.20 (équation différentielle linéaire d'ordre 2)**.** On cherche à résoudre l'équation différentielle $y'' = ay + by'$, avec $a, b \in \mathbb{C}$ et $a \neq 0$. Dans ce cas le polynôme à considérer est $P(X) = X^2 - bX - a$. Sur $\mathbb{C}$ ce polynôme est scindé, et il y a deux possibilités :

(1) Il y a deux racines distinctes $\lambda_1 \neq \lambda_2$. Dans ce cas la solution générale de l'équation est
$$y(t) = \alpha_1 e^{\lambda_1 t} + \alpha_2 e^{\lambda_2 t}\,, \quad \alpha_1, \alpha_2 \in \mathbb{C}\,.$$
(2) Il y a une racine double λ. Dans ce cas la solution générale de l'équation est
$$y(t) = (\alpha + \beta t)e^{\lambda t}\,, \quad \alpha, \beta \in \mathbb{C}\,.$$

Remarque 3.21 (solutions réelles)**.** Dans l'exemple qui précède, supposons que $a, b \in \mathbb{R}$, et qu'on s'intéresse aux solutions réelles de l'équation. Si les racines de $P(X)$ sont réelles, alors la solution générale est la même (avec α_1, α_2, α et β réels). Si il y a deux racines complexes conjuguées $\lambda = c + id$ et $\bar{\lambda} = c - id$, alors dans ce cas les solutions réelles de l'équation sont obtenues en prenant les parties réelles et imaginaires des solutions complexes. On trouve alors comme solution générale
$$y(t) = e^{ct}\big(\alpha \cos(dt) + \beta \sin(dt)\big)\,, \quad \alpha, \beta \in \mathbb{C}\,.$$
Il y a aussi une manière purement matricielle de le démontrer (voir l'Exercice no. 65 du livret d'exercices), sans passer par les complexes.