Université Claude Bernard Lyon 1
Année 2021-2022
Automne 2021

L2 UE Algèbre III
Séquence 3

Algèbre III

Notes du cours

version du 21 décembre 2021

Johannes Kellendonk

Table des matières

Chapitre 1

Espaces vectoriels et applications linéaires

Dans cette section on rappelle (sans preuves) les notions de l'Algèbre Linéaire vu en L1.

1.1 Espaces vectoriels

Soit $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$, où $\mathbb{C}$ (plus généralement, $\mathbb{K}$ peut être un corps). Un *espace vectoriel* E sur $\mathbb{K}$ est un ensemble muni de deux operations algébriques : la somme de deux vecteurs et la multiplication avec un scalaire (un élément de $\mathbb{K}$)

$$\begin{array}{rcl} E \times E \ni (x, y) & \mapsto & x + y \in E \\ \mathbb{K} \times E \ni (\lambda, x) & \mapsto & \lambda x \in E \end{array}$$

Ces deux opérations satisfont certaines axioms comme la commutativité de l'addtion, l'associativité et une forme de distributivité. [1] En particulier, E contient le *vecteur zero* 0_E, qui satisfait $x + 0_E = x$ et $x + (-x) = 0_E$.

En appliquant plusieurs fois ces opérations à des vecteurs $x_1, \cdots, x_n \in E$ et scalaires $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$, on obtient un nouveau vecteur $x \in E$, qui est une *combinaison linéaire* des x_i

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n.$$

On note $\text{Vect}\{x_1, \cdots, x_n\}$ l'espace de toutes les combinaisons linéaires qu'on peut fabriquer (engendrer) avec les vecteurs x_1 à x_n)

$$\text{Vect}\{x_1, \cdots, x_n\} = \{\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n | \lambda_i \in \mathbb{K}\}.$$

Exemples d'espaces vectoriels sont $\{0\}$, $\mathbb{K}^n$, $\mathbb{K}_n[X]$ (polynômes d'une variable X de degré $\leq n$ à coefficients dans $\mathbb{K}$) où $\mathcal{F}(X, \mathbb{K})$ (toutes les fonctions sur une ensemble X à valeur dans $\mathbb{K}$). Un vecteur n'est donc pas toujours une "fleche", mais peut être n'importe quoi. On dit souvent aussi *espace linéaire* à la place de espace vectoriel.

---

1. Les axioms d'un espace vectoriel E :
   1. $\forall x, y \in E : x + y = y + x$
   2. $\forall x, y, z \in E : (x + y) + z = x + (y + z)$
   3. E contient un élément 0_E t.q. $\forall x \in E : x + 0_E = x$
   4. $\forall x \in E : 0x = 0_E$ (ici $0 \in \mathbb{K}$)
   5. $\forall x \in E : 1x = x$ (ici $1 \in \mathbb{K}$)
   6. $\forall x, y \in E, \forall \lambda \in \mathbb{K} : \lambda(x + y) = \lambda x + \lambda y$
   7. $\forall x \in E, \forall \lambda, \mu \in \mathbb{K} : (\lambda + \mu)x = \lambda x + \mu x$
   8. $\forall x \in E, \forall \lambda, \mu \in \mathbb{K} : (\lambda\mu)x = \lambda(\mu x)$

1.1.1 Bases

Soit E un espace vectoriel. Une famille $\{x_1, \cdots, x_n\} \subset E$ est dite *libre* (où *linéairement indépendantes*) si l'èquation (avec les variables $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$)

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0_E$$

n'admet que la solution $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Si $\{x_1, \cdots, x_n\}$ n'est pas libre, on l'appelle aussi famille *liée*, et dans ce cas il existe i t.q. x_i est combinaison linéaire des autres x_j, $j \neq i$.

Une famille $\{x_1, \cdots, x_n\} \subset E$ est dite *génératrice* si tout élément x de E est combinaison linéaire des x_i, c.a.d. ils existent $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ t.q.

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

Autrement dit, $\{x_1, \cdots, x_n\} \subset E$ est génératrice si $\mathrm{Vect}\{x_1, \cdots, x_n\} = E$.

Définition 1.1.1. Une *base* pour l'espace vectoriel E est une famille ordonnée[2] de E, qui est *libre et génératrice.*

Étant donnée une base $\mathcal{B} = (b_1, \cdots, b_n)$ de E on peut écrire chaque $x \in E$ *d'une manière unique* comme combinaison linéaire des $b_1, \cdots, b_n$,

$$x = \lambda_1 b_1 + \cdots + \lambda_n b_n$$

où $\lambda_i \in \mathbb{K}$. Les scalaires λ_i s'appellent les *coefficients* (ou *coordonnées* ou *composantes*) de x dans la base $\mathcal{B}$ et on écrit[3]

$$[x]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Une procedure inductive de construire une base pour E est la suivante : Si E n'est pas trivial, alors il contient un élément $v_1 \neq 0_E$. $\{v_1\}$ est une famille libre. On peut agrandir une famille libre $\{v_1, \cdots, v_k\}$ qui n'engendre pas E en choisissant un vecteur $x \in E \backslash \mathrm{Vect}\{v_1, \cdots, v_k\}$. Ainsi on obtient une famille libre $\{v_1, \cdots, v_k, v_{k+1} = x\}$ avec un element de plus. Cette procédure doit s'arrêter si $E \backslash \mathrm{Vect}\{v_1, \cdots, v_k\} = \emptyset$. En particulier, toute famille libre $\{v_1, \cdots, v_k\}$, peut être completée en une base pour E.

Il est important de réaliser qu'une base pour E n'est jamais unique (sauf dans le cas ou E est trivial, c.à.d. $E = \{0_E\}$). Par contre, la taille de la base (le nombre des éléments) est uniquement déterminée par E. Elle peut d'ailleurs être infinie, notament si la precédure d'en haut ne s'arrête pas. Cette taille est appellée la *dimension* de E, notée $\dim E$.

On rappelle que $\mathbb{K}^n := \overbrace{\mathbb{K} \times \cdots \times \mathbb{K}}^{n-fois}$ et qu'on note ses vecteurs $(\lambda_1, \cdots, \lambda_n)$ et appelle les λ_i les coordonnées du vecteur. La base *canonique* de $\mathbb{K}^n$ est $(e_1, \cdots, e_n)$ avec $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0)$ etc.. Dans la base canonique $(\lambda_1, \cdots, \lambda_n)$ s'écrit comme le vecteur colonne $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$. Un vecteur colonne peut donc être vu comme un vecteur de $\mathbb{K}^n$ exprimé dans la base canonique. L'application $E \ni x \mapsto [x]_{\mathcal{B}} \in \mathbb{K}^n$ est une bijection linéaire (un *isomorphisme d'espaces vectoriels*) entre E et $\mathbb{K}^n$. Toute espace vectoriel (sur $\mathbb{K}$) de dimension n est alors isomorphe à $\mathbb{K}^n$. Il est important de se rendre compte que l'isomorphisme dépend de la base $\mathcal{B}$.

2. Une famille ordonnée est une suite déléments $(b_1, b_2, \cdots, b_n)$ (ou $(b_1, b_2, \cdots)$ si la famille est infinie). On trouve aussi la notation avec des accolades $\{b_1, b_2, \cdots, b_n\}$ pour une base.

3. On trouve aussi l'écriture en ligne $(\lambda_1, \cdots, \lambda_n)$ car elle est plus compacte dans un texte, mais pour le calcul matriciel il est plus astucieux d'écrire les éléments en colonne.

1.1.2 Somme directe

Soit E un espace vectoriel. Une partie $F \subset E$ est apellée *sous-espace* de E, si $\forall x, y \in F$ et $\forall \lambda \in \mathbb{K}$

$$x + y \in F, \qquad \lambda x \in F$$

Définition 1.1.2. Soient $F_1, \ldots, F_m$ des sous-espaces vectoriels de E.

1. On définit la somme des F_i :

$$\begin{aligned} F_1 + \cdots + F_m &= \{f_1 + \cdots + f_m \mid \forall i = 1, \ldots, m,\ f_i \in F_i\} \\ &= \{e \in E \mid \exists f_1 \in F_1, \ldots, \exists f_m \in F_m,\ e = f_1 + \cdots + f_m\} \end{aligned}$$

   Cet ensemble est un sous-espace vectoriel de E (exercice).
2. On dit que les F_i sont en somme directe et on écrit $F_1 + \cdots + F_m = F_1 \oplus \cdots \oplus F_m$ si, dans l'écriture de e comme somme $e = f_1 + \cdots + f_m$, le choix des f_i est unique. Autrement dit, l'équation $f_1 + \cdots + f_m = 0$, avec $f_i \in F_i$, n'a que la solution $f_i = 0$ pour tout $i = 1, \ldots, m$.
3. On dit que E est la somme des F_i si $E = F_1 + \cdots + F_m$.
4. On dit que E est la somme directe des F_i et on l'écrit sous la forme $E = F_1 \oplus \cdots \oplus F_m$ si d'une part $E = F_1 + \cdots + F_m$ et d'autre part les F_i sont en somme directe.
5. Soit F un sous-espace de E. Un autre sous-espace $G \subset E$ est appellé espace supplémentaire de F si E est la somme direct de F et G, c.à.d. $E = F \oplus G$.

Proposition 1. *Soient F et G deux sous-espaces vectoriels de E. Alors F et G sont en somme directe si et seulement si $F \cap G = \{0_E\}$. Ainsi G est un espace supplémentaire si et seulement si $E = F + G$ et $F \cap G = \{0_E\}$.*

Démonstration. (i) On suppose que $F \cap G \neq \{0_E\}$. Alors il existe $x \in F \cap G$, $x \neq 0_E$. On a

$$x = x + 0_E = 0_E + x$$

Ce sont deux manières différentes d'écrire x comme la somme d'un élément de F avec un élément de G. Donc F et G ne sont pas en somme directe.

(ii) On suppose que $F \cap G = \{0_E\}$. Soit

$$x = y + z = y' + z'$$

avec $y, y' \in F$ et $z, z' \in G$. Alors $y - y' = z' - z$. D'où $y - y'$ et $z - z'$ appartient à $F \cap G$. D'où $y - y' = 0_E$ et $z - z' = 0_E$. Donc F et G sont en somme directe. □

Remarque 1.1.3. *1. Il est faux de penser que $E = F \oplus G \oplus H$ si et s. si $E = F + G + H$ et $F \cap G \cap H = \{0\}$; ou même $F \cap G = \{0\}$ et $F \cap H = \{0\}$ et $G \cap H = \{0\}$ (voir le TD pour un exemple). Ainsi, l'équivalence de la proposition précédente ne se généralise pas à plus de deux sous-espaces.*

2. Il y a une similitude entre famille génératrice et somme ; et entre famille libre et être en somme directe. En effet, soit $v_1, \ldots, v_m$ une famille de vecteurs dans E. Soient $F_i = \mathrm{Vect}\{v_i\}$ pour tout $i = 1, \ldots, m$. Alors la famille des v_i est libre si et s. si les F_i sont en somme directe ; et la famille des v_i engendre E si et s. si E est la somme des F_i (exercice).

Corollaire 1. *Soient $F_1, \ldots, F_m$ des sous-espaces vectoriels de E tels que $E = F_1 \oplus \cdots \oplus F_m$. Alors $\dim E = \dim F_1 + \cdots \dim F_m$.*

Pour tout $i = 1, \ldots, m$, soit $\mathcal{B}_i$ une base de F_i et soit $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$. Alors $\mathcal{B}$ est une base de E (qu'on appellera une base adaptée à la somme directe).

Démonstration. Exercice. □

1.2 Applications linéaires

Définition 1.2.1. Une *application linéaire* entre deux espaces vectoriels E et F est un fonction $f : E \to F$ qui preserve les opérations algébriques :

$$f(x+y) = f(x) + f(y), \qquad f(\lambda x) = \lambda f(x)$$

Forcément $f(0_E) = 0_F$. On note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires entre E et F. $\mathcal{L}(E,F)$ lui-même est aussi un espace vectoriel. De plus, si $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$ alors la composition $g \circ f$ est une application linéaire entre E et G.

Définition 1.2.2. Un endomorphisme d'un espace vectoriel E est une application linéaire entre E et lui-même, c.a.d. $F = E$. On note $\mathcal{L}(E) = \mathcal{L}(E,E)$.

On peut composer deux endomorphismes d'un même espace $u_1 \circ u_2$ et le resultat est de nouveau un endomorphisme. Autrement dit,

$$\mathcal{L}(E) \times \mathcal{L}(E) \ni (u,v) \mapsto u \circ v \in \mathcal{L}(E)$$

est un produit associative et $\mathcal{L}(E)$ une algèbre associative.

On note aussi $u^0 = \mathrm{id}$, $u^2 = u \circ u$ etc.. De plus, si $P \in \mathbb{K}[X]$, $P(X) = \sum_{k=0}^n a_n X^n$ est un polynôme à coefficients dans $\mathbb{K}$ et u un endomorphisme de E, alors

$$P(u) := \sum_{k=0}^{n} a_n u^n$$

est un endomorphisme de E. Par exemple l'endomorphisme $f = u^2 + u - u^0$ est donné par

$$f(x) = u(u(x)) + u(x) - x.$$

1.2.1 Noyau, image et rang

Une application $f : E \to F$ entre deux ensembles est appellée injective si $f(x_1) = f(x_2)$ implique $x_1 = x_2$. Si E, F sont des espaces vectoriels et f est linéaire, alors f est injective si et seulement si $f(x) = 0_F$ implique $x = 0_E$. Si f n'est pas injective, l'équation $f(x) = 0_F$ admet donc des solutions $x \neq 0_E$. L'ensemble de ses solutions

$$\ker f := \{x \in E | f(x) = 0_E\}$$

est un sous-espace de E appellé le *noyau* de f et noté $\ker f$.

L'image d'une application $f : E \to F$ est l'ensemble

$$\mathrm{im} f := \{f(x) |\ x \in E\}.$$

f est appellée surjective, si son image est égale à F. Si E, F sont des espaces vectoriels et f est linéaire alors son image est un sous-espace de F. On appelle la dimension de l'image de f le *rang* de f.

Théorème 1. *Soit $f : E \to F$ une application linéaire. On suppose que la dimension de E est fini. Alors*

$$\dim E = \dim \ker f + \mathrm{rang}\ f.$$

1.2.2 Rappel du calcul matriciel

Une *matrice* $m \times n$ à coefficients dans $\mathbb{K}$ est un tableau de scalaires qui a m lignes et n colonnes. On note $M_{mn}(\mathbb{K})$ les matrices $m \times n$ à coefficients dans $\mathbb{K}$. Si $n = m$ on apelle la matrice aussi une matric carrée et note $M_{nn}(\mathbb{K}) = M_n(\mathbb{K})$.

Par exemple

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

est une matrice 2×3. Elle a deux lignes

$$L_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}, \quad L_2 = \begin{pmatrix} a_{21} & a_{22} & a_{23} \end{pmatrix}$$

et trois colonnes

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad C_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}, \quad C_3 = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$$

On peut donc aussi écrire

$$A = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} C_1 & C_2 & C_3 \end{pmatrix}$$

où encore

$$A = \left(a_{ij}\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \left(L_i\right)_{1 \le i \le m} = \left(C_j\right)_{1 \le j \le n}$$

Operations élémentaires Soit $A = \left(L_i\right)_{1 \le i \le m}$ une matrice de taille $m \times n$. On appelle operation de lignes élémentaire :

1. Multiplication d'une ligne de avec un scalaire. On se fixe i et $\lambda \in \mathbb{K}$ et multiplie chaque coefficient de la ligne L_i avec le même scalaire λ : $L_i \mapsto \lambda L_i$. Les autres lignes reste inchangées.

2. Ajouter un multiple d'une ligne à une autre ligne. On se fixe i, j et $\lambda \in \mathbb{K}$ et ajoute à L_i la ligne λL_j. Les autres lignes restent inchangées : $L_i \mapsto L_i + \lambda L_j$. Ici l'addition des lignes $L_i + \lambda L_j$ est l'addition coefficient par coefficient.

3. Échanger deux lignes : $L_i \leftrightarrow L_j$.

Une operation de colonnes élémentaire et la même chose avec lignes et colonnes interchangées.

Transposée d'une matrice La matrice transposée de $A = (a_{ij})$ est la matrice ${}^tA = (a_{ji})$. Donc les colonnes de ${}^tA$ sont les lignes de A tournées 90^0 vers la droite et les lignes ${}^tA$ sont les colonnes de A tournées 90^0 vers la gauche. Autrement dit, il s'agit d'une reflection à la diagonale partant du haut à gauche vers le bas à droite.

Structure linéaire $M_{mn}(\mathbb{K})$ est une espace vectoriel sur $\mathbb{K}$. Les operations sont coefficient par coefficient. La *matrice élémentaire* E_{ij} est la matrice dont le coefficient ij est 1 pendant que tous les autres coefficients sont 0. Une famille contenant les matrices élémentaires est libres et géneratrice. En choisissant une ordre parmi ces matrices on obtient donc une base pour $M_{mn}(\mathbb{K})$.

Structure multiplicative Soit A une matrice $m \times n$ et B une matrice $n \times p$. On définie le produit AB comme la matrice C de taille $m \times p$ t.q.

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

$M_n(\mathbb{K})$ est donc une algèbre associative avec une unité. L'unité est la matrice $\mathbf{1} = \left(\delta_{ij}\right)_{1 \leq i,j \leq n}$ où $\delta_{ij} = 1$ si $i = j$ pendant que $\delta_{ij} = 1$ si $i \neq j$. On note l'inverse de A par A^{-1}.

1.2.3 Matrice comme application linéaire

Une matrice $A \in M_{mn}(\mathbb{K})$ definit une application linéaire de $\mathbb{K}^n$ dans $\mathbb{K}^m$ de la manière suivante : Si $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $x = (x_j)_{1 \leq j \leq n}$ alors

$$Ax = b \quad \text{où} \quad b = (b_i)_{1 \leq i \leq m},\ b_i = \sum_{j=1}^{n} a_{ij} x_j$$

On note que ceci a l'air comme le produit des deux matrices, si on interprète un vecteur $x \in \mathbb{K}^n$ comme une colonne, c.à.d. une matrice $n \times 1$.

Une formule utile exprime Ax a l'aide des colonnes C_i de A. Si $A = (C_j)_{1 \leq j \leq n}$ et $x = (x_j)_{1 \leq j \leq n}$ alors

$$Ax = \sum_{j=1}^{n} x_j C_j.$$

Les notions du noyau, de l'image et du rang sont les mêmes que pour les applications linéaires. En particulier,

$$\mathrm{im} A = \mathrm{Vect}(C_1, \cdots, C_n)$$

et donc le rang de A est la dimension de l'espace vectoriel engendré par les colonnes de A.

Lemme 1. *Soit A une matrice $n \times n$. A est inversible si et seulement si ses colonnes forment une famille libre (si et seulement si ses lignes forment une famille libre). Ceci est le cas si et seulement si* rang $A = n$.

1.2.4 Matrice d'une application linéaire

Définition 1.2.3. On considère une application linéaire $f : E \to F$. Soit $\mathcal{B} = (b_1, \cdots, b_n)$ une base de E et $\mathcal{D} = (d_1, \cdots, d_m)$ une base de F. La *matrice associée à f dans les bases $\mathcal{B}$ et $\mathcal{D}$* est la matrice $m \times n$ dont la j-ième colonne est $[f(b_j)]_{\mathcal{D}}$.

On trouve des notations variées pour la matrice associée à une application linéaire. Une notation utlisée en L1 était $M_{\mathcal{BD}}(f)$. Une autre notation pour la matrice associée à f dans les bases $\mathcal{B}$ et $\mathcal{D}$, qui joue bien avec le calcul matriciel, est $[f]_{\mathcal{DB}}$ (avec les positions des bases interchangées !). Dans ce cas, on a la formule

$$[f(x)]_{\mathcal{D}} = [f]_{\mathcal{DB}}[x]_{\mathcal{B}}$$

où à droite il s'agit du produit de la matrice $[f]_{\mathcal{DB}}$ avec la matrice (d'une seule colonne) $[x]_{\mathcal{B}}$. De plus, si $g : F \to G$ est une autre application linéaire et $\mathcal{H} = (h_1, \cdots, h_k)$ une base de G, alors

$$[g \circ f]_{\mathcal{HB}} = [g]_{\mathcal{HD}}[f]_{\mathcal{DB}}$$

D'ailleurs, dans ce cours on ne travaillera principalement avec des endomorphismes sur un même espace vectoriel E. On n'aura besoin que de choisir une seule base, disons $\mathcal{B}$, et on pourra simplifier la notation en écrivant $M_{\mathcal{B}}(f)$ où $[f]_{\mathcal{B}}$ pour la matrice associée à f dans cette base.

1.3 Changement de la base

Considerons un espace vectoriel E et l'application identité $\mathrm{id} : E \to E$, $\mathrm{id}(x) = x$. Soient $\mathcal{B} = (b_1, \cdots, b_n)$ et $\mathcal{D} = (d_1, \cdots, d_n)$ deux bases pour E. Appellons $\mathcal{B}$ l'ancienne base et $\mathcal{D}$ nouvelle base. Alors $[\mathrm{id}]_{\mathcal{DB}}$ est la matrice associé à id si on prend la base $\mathcal{B}$ pour l'espace de départ et $\mathcal{D}$ pour l'espace d'arrivé. En particulier, la $j$ième colonne de $[\mathrm{id}]_{\mathcal{DB}}$ est $[b_j]_{\mathcal{D}}$, le vecteur b_j exprimé dans la base $\mathcal{D}$.

D'une manière similaire $[\mathrm{id}]_{\mathcal{BD}}$ est la matrice qui a comme $j$ième colonne le vecteur d_j exprimé dans la base $\mathcal{B}$.

Définition 1.3.1. La *matrice de passage*[4] *de la base* $\mathcal{B}$ *vers la base* $\mathcal{D}$ est la matrice

$$P_{\mathcal{BD}} := [\mathrm{id}]_{\mathcal{BD}}$$

Ses colonnes contiennent les vecteurs de la nouvelle base $\mathcal{D}$ exprimés dans l'ancienne base $\mathcal{B}$: **la j-ième colonne de la matrice de passage de $\mathcal{B}$ vers $\mathcal{D}$ est la colonne $[d_j]_{\mathcal{B}}$.**

1.3.1 Vecteur dans la nouvelle base

La matrice de passage permet de calculer les coordonnées d'un vecteur $x \in E$ dans la base $\mathcal{B}$ à l'aide de ses coordonnées dans la base $\mathcal{D}$ et vice versa.

Lemme 2. *On a*

$$[x]_{\mathcal{B}} = P_{\mathcal{BD}}[x]_{\mathcal{D}}$$

et donc aussi

$$[x]_{\mathcal{D}} = P_{\mathcal{BD}}^{-1}[x]_{\mathcal{B}} = P_{\mathcal{DB}}[x]_{\mathcal{B}}.$$

Démonstration. Exprimons l'équation $x = \mathrm{id}(x)$ en prenant la base $\mathcal{B}$ pour l'espace de départ de $\mathrm{id} : E \to E$ et $\mathcal{D}$ pour l'espace d'arrivé :

$$[x]_{\mathcal{B}} = [\mathrm{id}(x)]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{BD}}[x]_{\mathcal{D}}$$

D'où la première équation.

Comme $\mathrm{id} = \mathrm{id} \circ \mathrm{id}$, on a

$$1_n = [\mathrm{id}]_{\mathcal{BB}} = [\mathrm{id} \circ \mathrm{id}]_{\mathcal{BB}} = [\mathrm{id}]_{\mathcal{BD}}[\mathrm{id}]_{\mathcal{DB}}$$

donc

$$[\mathrm{id}]_{\mathcal{BD}} = [\mathrm{id}]_{\mathcal{DB}}{}^{-1}.$$

D'où la deuxième équation. □

Example : Soit $E = \mathbb{R}$, $\mathcal{B} = (1)$ et $\mathcal{D} = (10)$. Soit $x = 2$. Alors $[x]_{\mathcal{B}} = 2$ et $[x]_{\mathcal{D}} = \frac{1}{5}$. La matrice de passage de $\mathcal{B}$ vers $\mathcal{D}$ est $P_{\mathcal{BD}} = (10)$, et $P_{\mathcal{DB}} = (\frac{1}{10})$.

4. En anglais, on dit "change of base matrix" or "transition matrix". Attention, dans certains livres on appelle $P_{\mathcal{DB}}$ la matrice de passage de $\mathcal{B}$ vers $\mathcal{D}$, au lieu de $P_{\mathcal{BD}}$! Notre convention est la même que celle que vous trouvez sur Wikipedia (version francaise). Elle se justifie par le fait que les coefficients (p_{ij}) de $P_{\mathcal{BD}}$ satisfont $d_i = \sum_j p_{ij} b_j$.

1.3.2 Application linéaire dans la nouvelle base

Soit $f : E \to F$ une application linéaire, $\mathcal{B}$ une base pour E et $\mathcal{D}$ une base pour F. On suppose connaître les coefficients $[f]_{\mathcal{D}\mathcal{B}}$ dans la base $\mathcal{B}$ et $\mathcal{D}$. Soit $\mathcal{B}'$ une autre base de E et $\mathcal{D}'$ une base pour F. On veut calculer $[f]_{\mathcal{D}'\mathcal{B}'}$. Nous avons l'équation

$$[f]_{\mathcal{D}'\mathcal{B}'} = [\mathrm{id} \circ f \circ \mathrm{id}]_{\mathcal{D}'\mathcal{B}'} = [\mathrm{id}]_{\mathcal{D}'\mathcal{D}}[f]_{\mathcal{D}\mathcal{B}}[\mathrm{id}]_{\mathcal{B}\mathcal{B}'} = P^{-1}_{\mathcal{D}\mathcal{D}'}[f]_{\mathcal{D}\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}$$

donc on peut obtenir la matrice dans les nouvelles bases en multipliant la matrice dans les anciennes bases avec une matrice de passage et une matrice de passage inversée. Dans le cas d'un endomorphism $f : E \to E$ le changement de base se lit

$$[f]_{\mathcal{B}'} = P^{-1}_{\mathcal{B}\mathcal{B}'}[f]_{\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}$$

la nouvelle matrice est donc obtenue à partir de l'ancienne par conjugaison avec la matrice de passage.

Chapitre 2

Déterminant

2.1 Permutations

Soit X et Y deux ensembles et $f : X \to Y$ une fonction. f est bijective si f est injective et surjective. Une fonction bijective est inversible : son inverse (ou réciproque) f^{-1} est

$$f^{-1}(y) = x, \quad \text{avec } x \in X \text{ t.q. } f(x) = y.$$

Si $Y = X$ alors on peut composer deux fonctions $f_1, f_2 : X \to X$, $f_1 \circ f_2$. Cet opération est associative. De plus, si $f : X \to X$ est bijective, alors $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}$. De plus, $(f_1 \circ f_2)^{-1} = f_2^{-1} \circ f_1^{-1}$ pour deux fonctions bijectives.

L'ensemble de bijections de X forment un groupe avec élément neutre id, la multiplication étant la composition des bijections. Nous intéressons ici au cas que X est un ensemble fini. Autrement dit, il existe $n \in \mathbb{N}^*$ et une bijection entre X et l'ensemble $\{1, 2, \cdots, n\}$. Cette bijection donne une énumeration des éléments de X.

Définition 2.1.1. Une fonction bijective de X dans X est appellée une permutation de X. On note S_n le groupe des permutations de $\{1, 2, \cdots, n\}$.

Une notation effective pour les permutations de $\{1, 2, \cdots, n\}$ est la suivante

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

Dans cette écriture la permutation id a deux lignes égaux. Pour trouver l'inverse de σ il suffit d'échanger les deux lignes et de ré-ordonner les colonnes pour obénir l'ordre croissante dans la première ligne.

Proposition 2. *S_n contient $n!$ éléments.*

Démonstration. On pourrait faire une récurrence mais il suffit de compter. En effet, il y a n choix pour $\sigma(n)$. Une fois ce choix effectué, il reste $n-1$ choix pour $\sigma(n-1)$. Une fois $\sigma(n)$ et $\sigma(n-1)$ fixés, il reste $n-2$ choix pour $\sigma(n-2)$. Lorsqu'on arrive à $\sigma(2)$, il n'y a plus que deux choix possibles et pour $\sigma(1)$ il n'y a plus qu'une seule possibilité. Ainsi le nombre total de possiblités est bien $n!$. □

Définition 2.1.2. Une transposition est une permutation qui échange deux éléments et laisse les autres éléments fixe.

Dans l'écriture d'en haut, une transposition n'a donc que deux colonnes, disons la i et la j-ième, qui n'ont pas les mêmes chiffres. On note cette transposition aussi (ij). Alors le produit

de la transposition (ij) avec la permutation $\sigma = \begin{pmatrix} 1 & \cdots & n \\ k_1 & \cdots & k_n \end{pmatrix}$ est obtenue ainsi :
$(ij) \circ \sigma$ est obtenu en échangeant dans la deuxième ligne de σ les chiffres i et j,
$\sigma \circ (ij)$ est obtenu en échangeant dans la première ligne de σ les chiffres i et j et puis en ré-ordonnant les colonnes pour obtenir l'ordre croissant dans la première ligne.

Théorème 2. *Toute permutation est une composition (un produit) de transpositions.*

Démonstration. Soit $\sigma = \sigma_n = \begin{pmatrix} 1 & \cdots & n \\ k_1 & \cdots & k_n \end{pmatrix}$. Soit τ_n la permutation qui échange n avec k_n et laisse tous les autres éléments fixe (si $k_n = n$ alors $\tau_n = \mathrm{id}$, sinon τ_n est une transposition). Alors, il existe une permutation de $\{1, \cdots, n-1\}$, $\begin{pmatrix} 1 & \cdots & n-1 \\ k'_1 & \cdots & k'_{n-1} \end{pmatrix}$, tel que

$$\tau_n \circ \sigma_n = \begin{pmatrix} 1 & \cdots & n-1 & n \\ k'_1 & \cdots & k'_{n-1} & n \end{pmatrix}$$

On itère cette procedure : soit τ_{n-1} la permutation qui échange $n-1$ avec k'_{n-1}. Alors il existe une permutation de $\{1, \cdots, n-2\}$, $\begin{pmatrix} 1 & \cdots & n-2 \\ k'_2 & \cdots & k'_{n-1} \end{pmatrix}$, tel que

$$\tau_{n-1} \circ \tau_n \circ \sigma_n = \begin{pmatrix} 1 & \cdots & n-2 & n-1 & n \\ k''_1 & \cdots & k''_{n-2} & n-1 & n \end{pmatrix}$$

etc.. Ainsi on trouve n permutations τ_i t.q.

$$\tau_1 \circ \cdots \tau_n \circ \sigma_n = \mathrm{id}.$$

Plus précisement, τ_i est une transposition ou $\tau_i = \mathrm{id}$. Il en suit que

$$\sigma = \tau_n^{-1} \circ \cdots \circ \tau_1^{-1} = \tau_n \circ \cdots \circ \tau_1.$$

Les $\tau_i = \mathrm{id}$ s'annullent dans cette expression, pendant que les autres τ_i sont des transpositions. Donc σ est une composition des transpositions. □

Exemple 2.1.3. *Soit* $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 5 & 3 \end{pmatrix} \in S_6$. *Alors*

$$(36) \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 3 & 1 & 5 & 6 \end{pmatrix}$$

$$(14) \circ (36) \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 4 & 5 & 6 \end{pmatrix}$$

$$(12) \circ (14) \circ (36) \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

D'où

$$\sigma = (1\ 2) \circ (2\ 4) \circ (3\ 6)$$

Mais aussi

$$\sigma = (3\ 6) \circ (4\ 1) \circ (1\ 2)$$

De plus

$$(1\ 2) \circ (2\ 3) \circ (1\ 2) = (1\ 3).$$

Remarque 2.1.4. *Comme le montre ces exemples, la décomposition en produit de transpositions n'est pas unique (ni sur les transpositions ni sur leur nombre).*

2.1.1 Signe d'une permutation

Théorème 3. *Il existe une fonction unique $\varepsilon : S_n \to \{+1, -1\}$, appellée la fonction signature, qui satisfait*

1. $\varepsilon(\sigma) = -1$ *pour toute transposition* σ,
2. $\varepsilon(\sigma \circ \tau) = \varepsilon(\sigma)\varepsilon(\tau)$ *pour tout* $\sigma, \tau \in S_n$.

Démonstration. Unicité. On suppose qu'une telle fonction existe. Comme toute permutation peut-etre factorisée en un produit de transpositions, la formule $\varepsilon(\sigma \circ \tau) = \varepsilon(\sigma)\varepsilon(\tau)$ permet de calculer sa valeur sur une permutation quelconque, notament $\varepsilon(\sigma) = (-1)^k$ ou k est le nombre de facteurs dans la factorisation. Néanmoins ce n'est pas clair qu'une telle fonction existe, car la decomposition en transpositions n'est pas unique.

Existence. On pose

$$\varepsilon(\sigma) := \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Comme

$$\prod_{1 \le i < j \le n} |\sigma(j) - \sigma(i)| = \prod_{1 \le i < j \le n} |j - i|$$

on a $\varepsilon(\sigma) \in \{-1, 1\}$.

Soit $\sigma = (kl)$, $k < l$. Alors $\frac{\sigma(j)-\sigma(i)}{j-i} = 1$ si i, j, k, l sont deux à deux distinct. Dans le produit il suffit donc de considérer les facteurs $\frac{\sigma(j)-\sigma(i)}{j-i}$ où un des i, j coincide avec un des k, l ainsi que le facteur où $i = k$ et $j = l$. Le dernier donne $\frac{\sigma(l)-\sigma(k)}{l-k} = -1$. Donc

$$\begin{aligned}
\varepsilon(\sigma) &= -\left(\prod_{\substack{1 \le i < j \le n \\ i=k, j \ne l}} \frac{\sigma(j) - \sigma(k)}{j - k}\right)\left(\prod_{\substack{1 \le i < j \le n \\ j=k, i \ne l}} \frac{\sigma(k) - \sigma(i)}{k - i}\right)\left(\prod_{\substack{1 \le i < j \le n \\ i=l, j \ne k}} \frac{\sigma(j) - \sigma(l)}{j - l}\right)\left(\prod_{\substack{1 \le i < j \le n \\ j=l, i \ne k}} \frac{\sigma(l) - \sigma(i)}{l - i}\right) \\
&= -\left(\prod_{\substack{1 \le i < j \le n \\ i=k<j \ne l}} \frac{j - l}{j - k}\right)\left(\prod_{\substack{1 \le i < j \le n \\ i<j=k<l}} \frac{l - i}{k - i}\right)\left(\prod_{\substack{1 \le i < j \le n \\ k<i=l<j}} \frac{j - k}{j - l}\right)\left(\prod_{\substack{1 \le i < j \le n \\ j=l>i \ne k}} \frac{k - i}{l - i}\right)
\end{aligned}$$

Comme $\varepsilon(\sigma)$ est un signe il suffit maintenant de compter les facteurs negatifs. Les facteurs du deuxième et troisième produit sont tous positifs. Dans le premier produit, le facteur est negatif si $k < j < l$ et dans le quatrième si $k < i < l$. Le nombre de facteurs negatifs est donc pair. Donc $\varepsilon(\sigma) = -1$.

Soit σ et τ deux permutations. On observe que $(\sigma(j) - \sigma(i))(j - i) = (\sigma(i) - \sigma(j))(i - j)$ et donc,

$$\prod_{1 \le i < j \le n} (\sigma(\tau(j)) - \sigma(\tau(i)))(\tau(j) - \tau(i)) = \prod_{1 \le i < j \le n} (\sigma(j) - \sigma(i))(j - i)$$

. Il s'ensuit que

$$\begin{aligned}
\prod_{1 \le i < j \le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} &= \prod_{1 \le i < j \le n} \frac{(\sigma(\tau(j)) - \sigma(\tau(i)))(\tau(j) - \tau(i))}{(\tau(j) - \tau(i))(\tau(j) - \tau(i))} \\
&= \prod_{1 \le i < j \le n} \frac{(\sigma(j) - \sigma(i))(j - i)}{(j - i)(j - i)} \\
&= \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}
\end{aligned}$$

Finalement,

$$\begin{aligned}
\varepsilon(\sigma\circ\tau) &= \prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \\
&= \prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}\,\frac{\tau(j)-\tau(i)}{j-i} \\
&= \left(\prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}\right)\left(\prod_{1\le i<j\le n} \frac{\tau(j)-\tau(i)}{j-i}\right) \\
&= \varepsilon(\sigma)\ \varepsilon(\tau).
\end{aligned}$$

□

Corollaire 2. *Soit σ une permutation. Le nombre des transpositions dans une factorisation de σ en un produit de transpositions est soit pair (dans ce cas la permutation a signe $+1$) soit impair (dans ce cas la permutation a signe -1).*

A partir de cette information on trouve rapidement $\varepsilon(\mathrm{id})=1$ et, pour toute permutation σ, $\varepsilon(\sigma)=\varepsilon(\sigma^{-1})$.

2.2 Déterminant d'une matrice

Définition 2.2.1. Soit $A\in M_n(\mathbb{K})$. On note $A=(a_{ij})$ avec $1\le i,j\le n$. On définit le déterminant de A :

$$\det(A)=\sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{i=1}^{n}a_{\sigma(i),i}$$

Ainsi $\det(A)\in\mathbb{K}$.

Autres notations.

$$\det(A)=|A|=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}. \end{vmatrix}$$

Exemple 2.2.2. *1. ($n=2$) Soit $A=\begin{pmatrix} a & b \\ c & d\end{pmatrix}$. Alors $\det(A)=ad-bc$. En effet, les deux permutations de S_2 sont $\sigma=\mathrm{id}$ et $\sigma'=(12)$. On obtient alors*

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{vmatrix} &= \varepsilon(\sigma)a_{\sigma(1),1}a_{\sigma(2),2}+\varepsilon(\sigma')a_{\sigma'(1),1}a_{\sigma'(2),2} \\
&= ad-cb
\end{aligned}$$

2. ($n=3$)

$$S_3=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\right\}$$

On obtient alors

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

Proposition 3. *Soit $M = (m_{ij}) \in M_n(\mathbb{K})$ une matrice block-triagonale, c.à.d. il existe $k \leq n$ tel que $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ ou $A = (a_{ij}) \in M_k(\mathbb{K})$ et $D = (d_{ij}) \in M_{n-k}(\mathbb{K})$ sont des matrice carrées et B est une matrice k fois $n-k$. Alors*

$$\det(M) = \det(A)\det(D).$$

Démonstration. Une autre manière de caractériser une telle matrice block-triagonale $M = (m_{ij})$ est de dire que $m_{ij} = 0$ si $i > k$ et $j \leq k$. Supposons que c'est le cas. Alors le produit $\prod_{i=1}^n m_{\sigma(i)\,i}$, qui fait partie de la formule pour le déterminant, s'annulle, si $\sigma(j) > k$ pour un $j \leq k$. Donc, une condition necessaire pourque le produit est non-nul est, que σ envoie la partie $\{k+1, \cdots, n\}$ en lui-meme. Comme σ est une bijection ceci entraine que σ envoie aussi la partie $\{1, \cdots, k\}$ en lui-meme. Donc σ doit être la composition d'une permutation σ_1 de $\{1, \cdots, k\}$ avec une permutation σ_2 de $\{k+1, \cdots, n\}$. Il en suit que

$$\begin{aligned}
\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n m_{\sigma(i)\,i} &= \sum_{\sigma_1 \in S_k} \sum_{\sigma_2 \in S_{n-k}} \varepsilon(\sigma_1\sigma_2) \prod_{i=1}^k m_{\sigma_1(i)\,i} \prod_{i=k+1}^n m_{\sigma_2(i)\,i} \\
&= \left(\sum_{\sigma_1 \in S_k} \varepsilon(\sigma_1) \prod_{i=1}^k a_{\sigma_1(i)\,i} \right) \left(\sum_{\sigma_2 \in S_{n-k}} \varepsilon(\sigma_2) \prod_{i=k+1}^n d_{\sigma_2(i)-k\,i-k} \right) \\
&= \det(A)\det(D).
\end{aligned}$$

□

Corollaire 3. *Soit $A = (a_{ij}) \in M_n(\mathbb{K})$ une matrice triagonale. Alors*

$$\det(A) = \prod_{i=1}^n a_{i\,i}$$

Proposition 4. *Le déterminant est une application multilinéaire alternée en les colonnes, c'est-à-dire : Soit $A \in M_n(\mathbb{K})$ et soient $C_1, \ldots, C_n$ les colonnes de A.*

1. *Soit $i \in \{1, \ldots, n\}$. On suppose qu'il existe des colonnes C, C' et desscalaires $\lambda, \lambda' \in \mathbb{K}$ tels que $C_i = \lambda C + \lambda' C'$. Alors*

$$\det(A) = \lambda \det(C_1 \ \cdots \ \underbrace{C}_{position\ i} \ \cdots \ C_n) + \lambda' \det(C_1 \ \cdots \ \underbrace{C}_{position\ i} \ \cdots \ C_n).$$

2. *Soient $i, j \in \{1, \ldots, n\}$ avec $i \neq j$. Soit B la matrice obtenue en échangeant les colonnes C_i et C_j dans A (i.e. la colonne i de B est C_j et la colonne j de B est C_i, les autres colonnes étant les mêmes que dans A).*
   Alors $\det(B) = -\det(A)$.

Démonstration. 1. Notons $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ et $C' = \begin{pmatrix} c'_1 \\ \vdots \\ c'_n \end{pmatrix}$. Notons aussi B la matrice obtenue en remplaçant, dans A, la colonne C_i par C. De même, B' la matrice obtenue en remplaçant C_i par C'. L'hypothèse nous dit que pour tout $k = 1, \ldots, n$, $a_{ki} = \lambda c_k + \lambda' c'_k$. Par conséquent

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{\sigma(1),1}\cdots a_{\sigma(i),i}\cdots a_{\sigma(n),n} \\
&= \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{\sigma(1),1}\cdots (\lambda c_{\sigma(i)} + \lambda' c'_{\sigma(i)})\cdots a_{\sigma(n),n} \\
&= \lambda \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{\sigma(1),1}\cdots c_{\sigma(i)}\cdots a_{\sigma(n),n} + \lambda' \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{\sigma(1),1}\cdots c'_{\sigma(i)}\cdots a_{\sigma(n),n} \\
&= \lambda \det(B) + \lambda' \det(B').
\end{aligned}$$

2. Considérons la transposition $\tau = (i\ j) \in S_n$. La matrice B est alors $(C_{\tau(1)}\ \cdots\ C_{\tau(n)})$. On a :

$$\begin{aligned}
\det(B) &= \sum_{\sigma\in s_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),\tau(i)} \\
&= \sum_{\sigma\in s_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(\tau^{-1}(\tau(i))),\tau(i)} \\
&= \sum_{\sigma\in s_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(\tau^{-1}(i)),i} \\
&= \varepsilon(\tau) \sum_{\sigma\in s_n} \varepsilon(\sigma\circ\tau^{-1}) \prod_{i=1}^{n} a_{\sigma(\tau^{-1}(i)),i} \\
&= \varepsilon(\tau) \det(C_1\ \cdots\ C_n) \\
&= -\det(A).
\end{aligned}$$

□

La même démonstration que précédemment permet de prouver la propositon suivante.

Proposition 5. *Soient $C_1, \ldots, C_n$ les colonnes d'une certaine matrice $A \in M_n(\mathbb{K})$. Soit $\tau \in S_n$, alors*

$$\det(C_{\tau(1)}\ \cdots\ C_{\tau(n)}) = \varepsilon(\tau)\det(C_1\ \cdots\ C_n).$$

Remarque 2.2.3. *L'application* det *n'est pas linéaire. En fait,*

1. *pour $A \in M_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$, $\det(\lambda A) = \lambda^n \det(A)$.*
2. *pour $A, B \in M_n(\mathbb{K})$, en général on a $\det(A+B) \neq \det(A) + \det(B)$.*

Démonstration. 1. Si on note C_i les colonnes de A alors $\det(\lambda A) = \det(\lambda C_1\ \lambda C_2\ \cdots\ \lambda C_n) = \lambda \det(C_1\ \lambda C2\ \cdots\ \lambda C_n) = \lambda^2 \det(C_1\ C_2\ \cdots\ \lambda C_n) = \cdots$.

2. Voici un contre-exemple : $\det(I_n + I_n) = \det(2I_n) = 2^n \det(I_n) = 2^n$ alors que $\det(I_n) + \det(I_n) = 2$. Les deux termes sont différents si $n \geq 2$.

□

Proposition 6. *Soit $A \in M_n(\mathbb{K})$ une matrice ayant deux colonnes égales. Alors* $\det(A) = 0$.

Démonstration. Par hypothèse, il existe $i \neq j$ tel que les colonnes C_i et C_j soient égales. La matrice obtenue en échangeant ces colonnes est encore égale à A. Par la prop. 4, on obtient alors $\det(A) = -\det(A)$ ce qui donne $\det(A) = 0$. □

Proposition 7. *Soit $A \in M_n(\mathbb{K})$. Alors* $\det({}^tA) = \det(A)$.

Démonstration.

$$\begin{aligned}
\det({}^tA) &= \sum_\sigma \varepsilon(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} \\
&= \sum_\sigma \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma^{-1}(\sigma(i)),\sigma(i)} \\
&= \sum_\sigma \varepsilon(\sigma) \prod_{j=1}^n a_{\sigma^{-1}(j),j} \\
&= \sum_\sigma \varepsilon(\sigma^{-1}) \prod_{i=1}^n a_{\sigma^{-1}(j),j} \\
&= \det(A)
\end{aligned}$$

□

Corollaire 4. *Le déterminant est multilinéaire en les lignes. De plus si τ est une permutation des lignes de A alors le déterminant de la matrice obtenue en permutant les lignes à l'aide de τ est égal à $\varepsilon(\tau)\det(A)$.*

Démonstration. Ces propriétés sont vraies pour les colonnes et la prop. précédente permet de les avoir pour les lignes. □

Théorème 4. *Pour $A, B \in M_n(\mathbb{K})$, $\det(AB) = \det(A)\det(B)$.*

Démonstration. Notons $A = (a_{ij}) = (A_1 \ \cdots \ A_n)$, $B = (b_{ij})$, $C = AB = (c_{ij}) = (C_1 \ \cdots \ C_n)$ de sorte que $c_{ik} = \sum_{j=1}^n = a_{ij}b_{jk}$. On a alors

$$C_k = \begin{pmatrix} c_{1k} \\ \vdots \\ c_{nk} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}b_{jk} \\ \vdots \\ \sum_{j=1}^n a_{nj}b_{jk} \end{pmatrix} = \sum_{j=1}^n b_{jk} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \sum_{j=1}^n b_{jk}A_j.$$

D'où

$$\begin{aligned}
\det(AB) &= \det(C_1 \ C_2 \cdots \ C_n) \\
&= \det\Big(\sum_{i_1=1}^n b_{i_1,1}A_{i_1} \quad \sum_{i_2=1}^n b_{i_2,2}A_{i_2} \quad \cdots \quad \sum_{i_n=1}^n b_{i_n,n}A_{i_n}\Big) \\
&\overset{Prop.\ 4}{=} \sum_{1\le i_1,\dots,i_n\le n} b_{i_1,1}\ b_{i_2,2}\cdots b_{i_n,n}\ \det(A_{i_1}\ A_{i_2}\ \cdots\ A_{i_n}).
\end{aligned}$$

Dans cette somme, dès que deux i_j sont égaux, on obtient deux colonnes égales et le terme disparaît. Il ne reste que les termes où $i_1, \dots, i_n$ sont distincts deux à deux. Autrement dit les termes pour lesquelles $(i_1, \dots, i_n)$ est une permutation de $\{1, \dots, n\}$. On note alors $i_1 = \sigma(1) \dots, i_n = \sigma(n)$ et on peut réécrire

$$\begin{aligned}
\det(AB) &= \sum_{\sigma\in S_n} b_{\sigma(1),1}\ b_{\sigma(2),2}\cdots\ b_{\sigma(n),n}\ \det(A_{\sigma(1)}\ \cdots\ A_{\sigma(n)}) \\
&\overset{Prop.\ 5}{=} \sum_{\sigma\in S_n} b_{\sigma(1),1}\ b_{\sigma(2),2}\cdots\ b_{\sigma(n),n}\ \varepsilon(\sigma)\det(A_1\ \cdots\ A_n) \\
&= \det(B)\det(A).
\end{aligned}$$

□

Théorème 5. *Soit $A \in M_n(\mathbb{K})$, alors*

$$\det(A) \neq 0 \iff A \textit{ est inversible.}$$

Démonstration. "$\Leftarrow$" : Soit B l'inverse de A alors $AB = \mathrm{I_n}$ et $1 = \det(\mathrm{I_n}) = \det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B})$ d'où $\det(A) \neq 0$.

"$\Rightarrow$" : Par contraposée, on suppose A non inversible donc les colonnes C_i de A sont liées. Quitte à faire un échange de colonnes, on peut supposer que $C_1 = \lambda_2 C_2 + \cdots + \lambda_n C_n$ avec $\lambda_i \in \mathbb{K}$. Par conséquent $\det(A) = \det(\sum_2^n \lambda_i C_i \;\; C_2 \;\; \cdots \;\; C_n) = \sum_{i=1}^n \lambda_i \det(C_i \; C_2 \;\; \cdots \;\; C_n) = 0$. □

Corollaire 5. *Si A est une matrice inversible,* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Démonstration. Si A est inversible, alors $AA^{-1} = \mathrm{I_n}$. Donc

$$1 = \det(\mathrm{I_n}) = \det(\mathrm{A})\det(\mathrm{A^{-1}}).$$

□

2.2.1 Déterminant d'un endomorphisme

Dans ce paragraphe, E désigne un $\mathbb{K}$ espace vectoriel de dimension n et $u : E \to E$ un endomorphisme de E.

Lemme 3. *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $A = M_{\mathcal{B}}(u)$ et $A' = M_{\mathcal{B}'}(u)$. Alors* $\det(A) = \det(A')$.

Démonstration. Soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors $A' = P^{-1}AP$ et on a $\det(A') = \det(P^{-1})\det(A)\det(P) = (\det(P))^{-1}\det(P)\det(A) = \det(A)$. □

Définition 2.2.4. On définit le déterminant de u comme étant le déterminant de la matrice de u dans n'importe quelle base de E.

Le lemme précédent nous assure que cette définition ne dépend pas du choix de la base.

Nous donnons les deux propositions suivantes sans démonstrations car ces dernières sont faciles ou triviales.

Proposition 8. *On a les équivalences suivantes : u est bijectif $\iff$ u est injectif $\iff$ u est surjectif $\iff$ $\det(u) \neq 0$.*

Proposition 9. *Soient u et v deux endomorphismes de E. On a*

1. $\det(u \circ v) = \det(u)\det(v)$,
2. $\det(\mathrm{Id}_E) = 1$,
3. *Si u est inversible alors* $\det(u^{-1}) = (\det(u))^{-1}$.

2.3 Calcul pratique du déterminant

On utilise très rarement la formule dans la définition du déterminant pour le calculer. Voici quelques stratégies pour le calcul pratique d'un déterminant.

2.3.1 Par transformations élémentaires de la matrice

Une manière efficace passe par l'application des transformations élémentaires à la matrice pour la rendre triagonale. Bienque le déterminant n'est pas invariant sous transformation élémentaire, il ne peut changer que d'un signe, et ce signe peut être déterminé.

Proposition 10. *1. Le déterminant est opposé si on échange deux lignes ou deux colonnes.*

2. Si on remplace une colonne C (resp. une ligne L) par $C+$"une combinaison linéaire des autres" (resp. $L+\cdots$) alors le déterminant ne change pas.

Démonstration. 1. Déjà vu pour les colonnes et l'égalité $\det(A) = \det({}^tA)$ l'implique pour les lignes.

2. Faisons la preuve pour $C = C_1$. Remplacons alors C_1 par $C_1 + \sum_{i=2}^n \lambda_i C_i$,

$$\begin{aligned}\det(C_1 + \sum_{i=2}^n \lambda_i C_i \;\; C_2 \;\cdots\; C_n) &= \det(C_1 \; C_2 \;\cdots\; C_n) + \sum_{i=2}^n \lambda_i \det(C_i \; C_2 \;\cdots\; C_n) \\ &= \det(C_1 \; C_2 \;\cdots\; C_n).\end{aligned}$$

□

2.3.2 Developpement du déterminant lelong d'une colonne ou ligne

Une autre méthode de calcul est le developpement du déterminant lelong une colonne ou ligne. Elle répose sur le

Lemme 4.

$$\begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ 1 & a_{k2} & \cdots & a_{kn} \\ \vdots & & & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = (-1)^{k-1} \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Démonstration. Si $k = 1$ alors le résultat découle de la Prop. 3. Si $k > 1$ on echange la k-ième avec la $k-1$-ième ligne ; en conséquence le déterminant prend un signe $-$. On itère jusqu'en arrivant à une matrice où les 0 de la première colonne sont tous en bas. Ca fait $k-1$ opérations d'échange. D'où le resultat. □

Corollaire 6.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{k=1}^n (-1)^{k-1} a_{k1} \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,2} & \cdots & a_{k-1,n} \\ a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Démonstration. Comme le déterminant est linéaire dans les colonnes on a

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{k=1}^{n} a_{k1} \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ 1 & a_{k2} & \cdots & a_{kn} \\ \vdots & & & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

et le résultat découle du dernier lemme. □

Grâce à la transposition, on a un résultat similaire avec les lignes. De plus, en tenant compte d'un signe (notament $(-1)^{j-1}$) on a un resultat similaire en developpant lelong la j-ième colonne. Relié à cela est la notion du cofacteur d'une matrice :

Soit $M = (m_{ij}) \in M_n(\mathbb{K})$ et $k, l \leq n$. Si on enlève la $k$ième ligne et la $l$ième colonne on obtient une matrice noté $\tilde{M}_{kl}$ de taille $n-1 \times n-1$.

$$M = \begin{array}{c} \quad\quad l \\ \begin{pmatrix} A & \vdots & B \\ \cdots & \cdot & \cdots \\ C & \vdots & D \end{pmatrix} k, \end{array} \quad \tilde{M}_{kl} := \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

ou
$A = (a_{ij}) \in M_{k-1\,l-1}(\mathbb{K})$ avec $a_{ij} = m_{ij}$, $i < k$, $j < l$,
$B = (b_{ij}) \in M_{k-1\,n-l}(\mathbb{K})$ avec $b_{ij} = m_{ij}$, $i < k$, $j > l$
$C = (c_{ij}) \in M_{n-k\,l-1}(\mathbb{K})$ avec $c_{ij} = m_{ij}$, $i > k$, $j < l$
$D = (d_{ij}) \in M_{n-k\,n-l}(\mathbb{K})$ avec $c_{ij} = m_{ij}$, $i > k$, $j > l$

Définition 2.3.1. Le nombre

$$M_{kl} := (-1)^{k+l} \det(\tilde{M}_{kl})$$

est appelé le cofacteur d'indice (k, l) de M.[1]

Théorème 6. *Soit $A \in M_n(\mathbb{K})$. Pour tout $k = 1, \ldots, n$,*

$$\det(A) = a_{k1}A_{k1} + \cdots + a_{kn}A_{kn}. \tag{2.1}$$

Pour tout $j = 1, \ldots, n$,

$$\det(A) = a_{1j}A_{1j} + \cdots + a_{nj}A_{nj}. \tag{2.2}$$

Démonstration. La première expression est le developpement du déterminant lelong la $k$ième ligne. La deuxième expression est le developpement du déterminant lelong la $j$ième colonne.

D'abord on observe que la deuxième formule correspond, si $j = 1$, au cas du Cor. 6. Pour j quelconque on obtient la deuxième formule par application des échanges de colonnes : d'abord j avec $j-1$, puis $j-1$ avec $j-2$, etc.. Après $j-1$ échanges on obtient de nouveau le cas du Cor. 6. Chaque échange amène à un signe -1. De plus, la $j$ième colonne est devenu la première et les autres colonnes ont gardées leur ordre. D'où la formule (2.2) (le signe est déjà pris en charge par le signe dans la définition du cofacteur).

La formule (2.1) peut être obtenu par transposition de (2.2). □

1. Le nombre $\det(\tilde{M}_{kl})$ est aussi appelé mineur de d'ordre $n-1$ d'indice (k, l). On ne parlera pas de mineurs d'ordre $< n-1$ dans ce cours.

2.4 Une formule pour l'inverse d'une matrice

On peut exprimer l'inverse A^{-1}, sous l'hypothèse que $\det(A) \neq 0$, à l'aide des cofacteurs. Bienque cette formule devient lourde si la dimension augmente, elle est utile théoriquement.

Définition 2.4.1. La comatrice de la matrice $A \in M_n(\mathbb{K})$ est la matrice $\mathrm{co}(A) \in M_n(\mathbb{K})$, qui a comme coefficient ij le cofacteur de A d'indice (i,j),

$$\mathrm{co}(A) := (A_{ij}).$$

Le transpośe de la comatrice est alors ${}^t\mathrm{co}(A) = (A_{ji})$.

Théorème 7. *Soit $A \in M_n(\mathbb{K})$. Alors*

$$A\,{}^t\mathrm{co}(A) = {}^t\mathrm{co}(A)A = \det(A)I_n.$$

Corollaire 7. *Si A est inversible alors $A^{-1} = \det(A)^{-1}\ {}^t\mathrm{co}(A)$.*

Démonstration du théorème. Comme d'habitude on note a_{ij} les coefficients de A et A_{ij} les cofacteurs associés à A. Pour $i,j \in \{1,\ldots,n\}$, on considère

$$\Gamma_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

— Si $i = j$ alors $\Gamma_{ij} = \det(A)$: en effet, c'est le développement du déterminant par rapport à la ligne i.

— Si $i \neq j$ alors $\Gamma_{ij} = 0$. En effet, soit B la matrice obtenue à partir de A en remplaçant la ligne j par la ligne i (les lignes autres que la ligne j étant celles de A). Alors d'une part $\det(B) = 0$ et d'autre part si on développe $\det(B)$ par rapport à la ligne j, on obtient Γ_{ij}.

Par conséquent : $\Gamma_{ij} = \delta_{ij}\det(A)$.
Le membre de gauche est le coefficient (i,j) de $A\ {}^t\mathrm{co}(A)$ et le membre de droite est le coefficient (i,j) de $\det(A)\ I_n$. On a donc démontré que $A\ {}^t\mathrm{co}(A) = \det(A)\ I_n$.
Pour obtenir l'égalité ${}^t\mathrm{co}(A)\ A = \det(A)\ I_n$ on fait la même chose en utilisant $\Gamma'_{ij} = \sum_{k=1}^n a_{kj}A_{ik}$. □

2.4.1 Formules de Cramer

Théorème 8. *Soient $a_{ij} \in \mathbb{K}$ avec $i,j \in \{1,\ldots,n\}$. Soient $b_1,\ldots,b_n \in \mathbb{K}$. On considère le système suivant*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Notons $A = (a_{ij}) \in M_n(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in M_{n\times 1}(\mathbb{K})$. Si $\det(A) \neq 0$ alors l'unique solution du système est donnée par

$$x_i = \frac{\det(C_1 \cdots C_{i-1}\ B\ C_{i+1} \cdots C_n)}{\det(A)}$$

où C_i désigne la colonne i de A.

Démonstration. Notons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}(\mathbb{K})$ l'unique solution du système $AX = B$. On a donc $B = x_1C_1 + \cdots + x_nC_n$ d'où

$$\begin{aligned} \det(C_1 \cdots C_{i-1}\ B\ C_{i+1} \cdots C_n) &= \sum_{j=1}^{n} x_j \det(C_1 \cdots C_{i-1}\ C_j\ C_{i+1} \cdots C_n) \\ &= x_i \det(C_1 \cdots C_{i-1}\ C_i\ C_{i+1} \cdots C_n) \\ &= x_i \det(A). \end{aligned}$$

□

Chapitre 3

Equation propre et spectre d'un endomorphisme

Plusieurs questions trouvent des réponses à travers la notion des valeurs propres (spectre) d'un endomorphisme. Par exemple :

1. Quelles propriétés d'une matrice carrée restent inchangées si on conjuge la matrice par une matrice inversible ?
2. Quelles propriétés d'un endomorphisme peuvent se déduire de l'expression de la matrice associée à l'endomorphisme dans une base ?
3. Quelle est la description la plus simple d'un endomorphisme ?

De plus, dans des applications diverses, le spectre d'un endomorphisme a très souvent une interprétation importante (instrument de musique).

3.1 Valeur, vecteur, espace propre

Définition 3.1.1. Soit u un endomorphisme d'un espace vectoriel E. L'équation

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

où $\lambda \in \mathbb{K}$ et $x \in E$, est appellée l'équation propre pour u.

Définition 3.1.2. Soit u un endomorphisme d'un espace vectoriel E. On appelle $\lambda \in \mathbb{K}$ *valeur propre* pour u si l'équation propre $u(x) = \lambda x$ admet une la solution $x \neq 0_E$.

Une telle solution x est appellée *vecteur propre* de u (pour la valeur propre λ).

On note que, pour tout $\lambda \in \mathbb{K}$, l'équation propre a toujours la solution $x = 0_E$. Celle-ci n'est donc pas interéssante. C'est pour cela on n'appelle 0_E pas vecteur propre.

Si $x \neq 0_E$ est vecteur propre pour u alors il existe $\lambda \in \mathbb{K}$ t.q. $(u - \lambda \mathrm{id})(x) = 0$. On appelle

$$E_\lambda := \ker(u - \lambda \mathrm{id})$$

l'espace propre pour λ (de u). E_λ contient toujours 0_E, mais est un sous-espace non-trivial seulement de E si λ est une valeur propre pour u. On appelle l'ensemble des valuers propres le *spectre* de u. On va voir plus bas que, si E est de dimension fini, on peut déterminer les valeurs propres en passant par le polynôme caractéristique.

Proposition 11. *Les espaces propres d'un endomorphisme sont en somme directe, c.à.d.*

$$E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}$$

où $\lambda_1, \cdots, \lambda_k$ sont les valeurs propres (distinctes) de u.

Démonstration. Par récurrence sur k. Pour $k = 1$ il n'y a rien à prouver.

Soit $\lambda_1, \cdots, \lambda_k$ des valeurs propres distinctes de u. Soit $x_i \in E_{\lambda_i}$ et

$$x_1 + \cdots + x_k = 0_E.$$

On a alors

$$0_E = (u - \lambda_k \mathrm{id})(0_E) = \sum_{i=1}^{k}(u - \lambda_k \mathrm{id})(x_i) = \sum_{i=1}^{k}(\lambda_i - \lambda_k)x_i = \sum_{i=1}^{k-1}(\lambda_i - \lambda_k)x_i$$

Par hypothèse de récurrence on a $y_1 + \cdots + y_{k-1} = 0$, $y_i \in E_{\lambda_i}$, implique $y_i = 0$ pour $i = 1, \cdots, k-1$. D'ou $(\lambda_i - \lambda_k)x_i = 0$ pour tout $i \leq k-1$. Comme $\lambda_i \neq \lambda_k$ pour $i \leq k-1$ ceci implique $x_i = 0$ pour tout $i \leq k-1$. Donc aussi $x_k = 0$. La somme $E_{\lambda_1}, \cdots, E_{\lambda_k}$ est alors directe. $\square$

3.2 Endomorphismes diagonalisables

Une matrice *diagonale* est une matrice carrée qui n'a que des coefficients non-nulles sur la diagonale, c.à.d.

$$a_{ij} = 0, \quad \text{si } i \neq j$$

On va utiliser la notation $\mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ pour la matrice diagonale avec $a_{ii} = \lambda_i$.

Définition 3.2.1. Soit $A \in M_n(\mathbb{K})$. On dit que A est diagonalisable (sur $\mathbb{K}$) s'il existe une matrice inversible $P \in M_n(\mathbb{K})$ t.q. $D = P^{-1}AP$ est une matrice diagonale. Autrement dit, A est conjugé à une matrice diagonale.

Soit u un endomorphisme d'un espace vectoriel E sur $\mathbb{K}$, de dimension fini. u est diagonalisable si E admet une base $\mathcal{B}$, t.q. $[u]_{\mathcal{B}}$ est une matrice diagonale.

Comme on peut interprêter une matrice $n \times n$ comme la matrice d'un endomorphisme de $\mathbb{K}^n$ dans la base canonique, la première définition est un cas particulier de la deuxieme. En effet, P joue le role de la matrice de passage de la base canonique vers une base dans laquelle A est diagonale.

La notion de matrice diagonale dépend du corps $\mathbb{K}$. Il se peut qu'une matrice $A \in M_n(\mathbb{C})$ n'a que des coefficients *réels*. Si A est diagonalisable et il existe une matrice inversible *réelle* P t.q. $D = P^{-1}AP$ est diagonale, alors A est même diagonalisable dans $M_n(\mathbb{R})$ et on spécifie que A est diagonalisable sur $\mathbb{R}$.

De l'autre côté, il se peut qu'une matrice $A \in M_n(\mathbb{R})$ n'est pas diagonalisable sur $\mathbb{R}$, c.à.d. qu'il n'existe pas de matrice inversible *réelle* P t.q. $D = P^{-1}AP$ est diagonale, mais qu'il existe une matrice inversible *complexe* P t.q. $D = P^{-1}AP$ est diagonale. Dans ce cas, A est diagonalisable dans $M_n(\mathbb{C})$ et on dit qu'elle est diagonalisable sur $\mathbb{C}$.

Une matrice qui est diagonalisable sur $\mathbb{R}$ est alors aussi diagonalisable sur $\mathbb{C}$, mais une matrice réelle qui est diagonalisable sur $\mathbb{C}$ n'est pas forcement diagonalisable sur $\mathbb{R}$.

Théorème 9 (Première critère de diagonalisation)**.** *Soit u un endomorphisme sur E, de dimension n. u est diagonalisable si et seulement si E admet une base $\mathcal{B} = \{b_1, \cdots, b_n\}$ de vecteurs propres de u. Dans ce cas $[u]_{\mathcal{B}} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ où λ_i est la valeur propre pour b_i.*

Démonstration. Soit $\mathcal{B} = \{b_1, \cdots, b_n\}$ une base de vecteurs propres de u. Il existe alors λ_i t.q. $u(b_i) = \lambda_i b_i$. Alors par calcul directe

$$[u]_{\mathcal{B}} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n).$$

Soit maintenant u diagonalisable. Il existe donc une base $\mathcal{B} = \{b_1, \cdots, b_n\}$ t.q. $[u]_\mathcal{B}$ est une matrice diagonale, disons $[u]_\mathcal{B} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ avec des scalaires $\lambda_j \in \mathbb{K}$. On trouve

$$[u(b_j)]_\mathcal{B} = [u]_\mathcal{B}[b_j]_\mathcal{B} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)e_j = \lambda_j e_j = [\lambda_j b_j]_\mathcal{B}$$

Donc $[u(b_j) - \lambda_j b_j]_\mathcal{B} = 0_{\mathbb{K}^n}$. Donc $u(b_j) - \lambda_j b_j = 0_E$. Comme $b_j \neq 0_E$ c'est un vecteur propre pour la valeur propre λ_j. □

Tout vecteur propre appartient à un espace propre. Donc si E admet une base de vecteurs propres on a $E \subset E_{\lambda_1} + \cdots + E_{\lambda_k}$, ce qui entraine

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \tag{3.1}$$

par Prop. 11. De l'autre côté, si $\mathcal{B}_i$ est une base the E_{λ_i} (et une telle base exists toujours) alors (3.1) implique que $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_n$ est une base pour E. Donc (3.1) est une condition necessaire et suffisant pour que E admet une base de vecteurs propres. On peut aussi formuler ca comme ca : Soit $g_\lambda = \dim E_\lambda$, dite la *multiplicité géométrique* de la valeur propre λ. Par définition d'une valeur propre, $g_\lambda \geq 1$. On a alors

Lemme 5. *Soit u un endomorphisme sur E, de dimension n. u est diagonalisable si et seulement si la somme des multiplicités géométriques vaut n.*

Démonstration. On a vu que u est diagonalisable si et seulement si (3.1) est satisfait. Comme l'inclusion $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \subset E$ est toujours vrai, la somme des dimensions des espaces propres doit être $\leq n$ et est égal à n si et seulement si l'inclusion est un égalité. □

Exemple 3.2.2.

1. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *est diagonalisable sur* $\mathbb{Q}$.
2. $\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ *n'est pas diagonalisable sur* $\mathbb{Q}$, *mais sur* $\mathbb{R}$.
3. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *n'est pas diagonalisable sur* $\mathbb{R}$, *mais sur* $\mathbb{C}$.
4. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *n'est pas diagonalisable sur* $\mathbb{C}$.

3.3 Polynôme caractéristique

Dans l'exemple en haut on était capable de calculer facilement les valeurs propres et les vecteurs propres car la dimension n'était pas très élevée. Le polynôme caractéristique est l'outil qui permet (en principe) de déterminer tous les valeurs propres en toute dimension.

Définition 3.3.1. Soit $A = (a_{ij})$ une matrice $n \times n$. On appelle

$$P_A(\lambda) := \det(A - \lambda 1_n)$$

le polynôme caractéristique de A.

Soit u un endomorphisme d'un espace vectoriel de dimension finie. On appelle

$$P_u(\lambda) := \det(u - \lambda \mathrm{id})$$

le polynôme caractéristique de u.

Les deux définitions an haut sont bien sûr reliées, car le detérminant d'un endomorphisme est défini à l'aide d'une matrice pour l'endomorphisme. Ainsi $\det(u-\lambda\mathrm{id}) = \det([u-\lambda\mathrm{id}]_{\mathcal{B}}) = \det([u]_{\mathcal{B}} - \lambda 1_n)$ où $\mathcal{B}$ est n'importe quel base pour E et $n = \dim E$. Dans ce qui suit nous étudions le polynôme caractéristique des matrices, mais les énoncés peuvent facilement être reformulées en termes d'endomorphismes.

On note que $P_A(\lambda)$ est bien un polynôme en λ, son degré est n.

Lemme 6. *Soit A une matrice $n \times n$. Alors*

$$P_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\mathrm{Tr}(A)\ \lambda^{n-1} + \cdots + \det(A)$$

où $\mathrm{Tr}(A) = \sum_{i=1}^n a_{ii}$ est la trace de A.

Démonstration. On rappelle que $\det(A - \lambda 1_n)$ est une somme de termes de la forme

$$Q_\sigma(\lambda) = \varepsilon(\sigma) \prod_{i=1}^n (a_{\sigma(i)i} - \lambda \delta_{\sigma(i)i}). \tag{3.2}$$

Chacun des Q_σ est un polynôme en λ. Pourqu'une puissance λ^n apparaisse, il faut que tous les $\delta_{\sigma(i)i}$ soient non-nuls. Ceci est le cas seulement si $\sigma = \mathrm{id}$. Dans ce cas on obtient de (3.2)

$$Q_{\mathrm{id}}(\lambda) = \prod_{i=1}^n (a_{ii} - \lambda) = (-\lambda)^n + \sum_{i=1}^n a_{ii}(-\lambda)^{n-1} + \cdots$$

où les $\cdots$ sont des termes en λ^m avec $m < n-1$. Si $\sigma \neq \mathrm{id}$ alors au moins deux des $\delta_{\sigma(i)i}$ sont nuls. Dans ce cas Q_σ est un polynôme de λ de degré $\leq n-2$.

Finalement le terme constant est $P_A(0) = \det(A)$. □

Théorème 10. *λ est une valeur propre de u si et seulement si $P_u(\lambda) = 0$.*

Démonstration. On a la chaine d'équivalences suivante :

λ est valeur propre de u $\Leftrightarrow$ $\ker(u - \lambda\mathrm{id}) \neq \{0_E\}$ $\Leftrightarrow$ $u - \lambda\mathrm{id}$ n'est pas injective $\Leftrightarrow$ $u - \lambda\mathrm{id}$ n'est pas inversible $\Leftrightarrow$ $\det(u - \lambda\mathrm{id}) = 0$. □

Corollaire 8. *Une matrice $A \in M_n(\mathbb{K})$ possède au plus n valeurs propres.*

Démonstration. $P_A(\lambda)$ est un polynôme de degré n. Un polynôme de degré n admet au plus n racines. □

Soit $P \in \mathbb{K}[X]$ un polynôme et λ une racine de P. On dit que λ a multiplicité m si $P(X)$ est divisible par $(X-\lambda)^m$ mais pas divisible par $(X-\lambda)^{m+1}$.

Définition 3.3.2. Soit λ une valeur propre de A. On dit que la *multiplicité algébrique* de λ est m si, en tant que racine du polynôme caractéristique P_A, λ a multiplicité m. On note la multiplicité algébrique de λ par m_λ.

Une valeur propre simple est donc une valeur propre de multiplicité algébrique 1.

Définition 3.3.3. Un polynôme $P \in \mathbb{K}_n[X]$ est scindé si il se factorise en facteurs linéaires, c.à.d. il existe $\lambda_i \in \mathbb{K}$ (pas forcement distinct) et $c \in \mathbb{K}$ t.q.

$$P[X] = c \prod_{i=1}^n (X - \lambda_i)$$

On rappelle le théorème fondamental de l'Algèbre :

Théorème 11. *Tout polynôme $P \in \mathbb{C}[X]$ est scindé.*

Ce résultat n'est pas vrai si $\mathbb{K} = \mathbb{R}$, comme le montre l'exemple du polynôme $P[X] = X^2+1$. Bienque $X^2 + 1 = (X + i)(X - i)$, i n'est pas réel.

Corollaire 9.

1. *Une matrice $A \in M_n(\mathbb{C})$ possède au moins une valeur propre (complexe).*
2. *Une matrice $A \in M_{2n+1}(\mathbb{R})$ possède au moins une valeur propre réelle.*

Démonstration. Si $A \in M_n(\mathbb{C})$ alors le polynôme caractéristique est scindé (sur $\mathbb{C}$). Un polynôme scindé admet au moins une racine. D'où le premier resultat.

Si $A \in M_n(\mathbb{R})$, alors A est aussi une matrice à coefficient complexe, c.à.d. $A \in M_n(\mathbb{C})$. Le polynôme caractéristique de A est donc scindé sur $\mathbb{C}$, mais pas forcément sur $\mathbb{R}$. Or, comme A est une matrice réelle, son polynôme caractéristique P_A a des coefficients réels. Donc

$$P_A(\bar{\lambda}) = \overline{P_A(\lambda)}.$$

Ainsi, si λ est un racine de multiplicité m_λ alors $\bar{\lambda}$ est un racine de multiplicité $m_{\bar{\lambda}} = m_\lambda$. Il en suit que, si P_A n'a pas de racine réelle, alors son degré est pair. □

Une matrice $A \in M_{2n}(\mathbb{R})$ possède une valeur propre complexe, mais pas forcément une valeur propre réelle. Il faut faire attention à ca.

Lemme 7. *On a $m_\lambda \geq g_\lambda = \dim E_\lambda$.*

Démonstration. Soit $\mathcal{B}_\lambda$ une base pour E_λ. On peut la compléter en une base $\mathcal{B}$ pour E. Si $\mathcal{B}_\lambda$ sont les premiers éléments de la base alors

$$[u]_{\mathcal{B}} = \begin{pmatrix} \lambda 1_{g_\lambda} & B \\ 0 & D \end{pmatrix},$$

une forme block triangulaire où le block en haut à gauge est la matrice diagonale de taille $g_\lambda = \dim E_\lambda$, qui a partout λ sur la diagonale. On calcule rapidement

$$P_u(\lambda') = (\lambda - \lambda')^{g_\lambda} P_D(\lambda')$$

ce qui montre que m_λ est au moins aussi grand que g_λ. □

Théorème 12 (Deuxième critère de diagonalisation)**.** *Soit u un endomorophism d'un espace vectoriel de dimension n. u est diagonalisable si et seulement si son polynôme caractéristique est scindé et pour toute valeur propre λ la multiplicité algébrique et la multiplicité géométrique coincident, c.à.d. $m_\lambda = g_\lambda$.*

Démonstration. Supposons que le polynôme caractéristique de u est scindé. Alors la somme des multiplicités algébriques m_λ est n. Si, de plus, $m_\lambda = g_\lambda$ pour toute valeur propre, alors la somme des g_λ est aussi n. Donc, par Lemma 5, u est diagonalisable.

Si u est diagonalisable on choisit une base $\mathcal{B}$ de vecteurs propres. Alors $[u]_{\mathcal{B}}$ est diagonal, disons $[u]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ (ici les λ_i ne sont pas forcément distincts). On calcule rapidement que

$$P_u(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda).$$

Donc P_u est scindé. Par Lemma 5 la somme des g_λ est n et comme $m_\lambda \geq g_\lambda$ et la somme des m_λ est le degré de P_A on doit avoir $g_\lambda = m_\lambda$ pour tout valeur propre λ. □

Il en suit une condition suffisante pourque u soit diagonalisable, mais cette condition n'est pas necessaire.

Corollaire 10. *Si u admet n valeurs propres disctinctes (ou de façon équivalente, si P_u est scindé à racines simples) alors u est diagonalisable.*

Démonstration. L'hypothèse nous dit que $1 = m_\lambda \geq g_\lambda \geq 1$ pour toute valeur propre λ. □

3.3.1 Algorithme de diagonalisation

Étant donné un endomorphisme u diagonalisable, on se pose le problème de trouver une base $\mathcal{B}$ t.q. $[u]_\mathcal{B}$ soit diagonale et puis d'expliciter la forme de $[u]_\mathcal{B}$. Voici les étappes :

1. Trouver les racines du polynôme caractéristique P_u. Chaque racine λ_i est une valeur propre.
2. Résoudre l'équation $(u - \lambda_i \mathrm{id})(x) = 0_E$ pour chaque i. Les solutions forment l'espace propre E_{λ_i} de λ_i. Choisir une base $\mathcal{B}_i$ pour E_{λ_i}.
3. La matrice $[u]_\mathcal{B}$ associée à u dans la base $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ est alors une matrice diagonale, qui contient les valeurs propres sur la diagonale, chaqu'une autant de fois que sa multiplicité algébrique.

Des bases différentes pour E peuvent amener à une matrice diagonale. Néanmoins, à part de l'ordre des éléments sur la diagonale, la forme diagonale de u est unique.

Étant donnée une matrice A, donc un endomorphisme de $\mathbb{K}^n$ ou la matrice d'un endomorphism $A = [u]_\mathcal{B}$ dans une base $\mathcal{B}$, diagonaliser A veut dire de trouver une autre base dans laquelle elle est diagonale. Déterminer une forme diagonale de A correspond donc à un changement de base. Voici les étappes :

1. Trouver les racines du polynômes caractéristiques P_A. Chaque racine λ_i est une valeur propre de A.
2. Résoudre l'équation $(A - \lambda_i 1_n)x = 0$ pour chaque i. Ici $A - \lambda_i 1_n$ est une matrice $n \times n$ et x est un vecteur colonne de taille n. Les solutions forment l'espace propre E_{λ_i} de λ_i exprimé dans la base canonique. Ainsi on trouve g_{λ_i} vecteurs propres linéairement indépendents pour chaque i et, mise ensemble, on obtient n vecteurs propres linéairement indépendents.
3. La forme diagonale D pour A est alors

$$D = P^{-1}AP$$

   où P est la matrice dont la $j$ième colonne correspond au $j$ième vecteur propre.

   D est alors une matrice diagonale qui contient chaque valeur propre autant de fois que sa multiplicité algébrique.

On remarque que la matrice P est la matrice de passage de la base canonique vers la base donnée par les vecteurs propres qu'on a déterminé. La base des vecteurs propres n'est pas unique donc P dépendent des choix. Pourtant, à part de l'ordre des éléments sur la diagonale, D est unique.

On pourrait donc répondre à la première question du chapitre : le spectre d'un endomorphisme avec multiplicité ne dépend pas de la base. Il en suit que toute quantité qui s'exprime avec le spectre (le déterminant, la trace,...) ne dépendent pas de la base.

Chapitre 4

Application aux équations linéaires d'évolution : principe de découplage

Dans ce chapitre on va voir que la diagonalisation des matrices correspond au découplage d'une équation linéaire qui décrit l'évolution temporaire d'un système de n degrés de libertés. On va considérer deux cas : les récurrences linéaires de plusieurs variables et les équations différentiels linéaires de plusieurs variables.

Bien sûr, pas toute matrice est diagonalisable, mais si elle est, les calculs sont plus simples. On reviendra aux cas plus généraux plus bas.

4.1 Puissances et fonctions des matrices diagonalisables

Soit A une matrice $n \times n$. Dans ce qui suit, il est utile de voir A comme élément de $M_n(\mathbb{C})$, même si ses coefficients sont réels, et de travailler avec la diagonalisation sur $\mathbb{C}$. On va supposer dans cette section que A soit diagonalisable (sur $\mathbb{C}$), c.à.d. qu'il existe une matrice $n \times n$ inversible P t.q.

$$D = P^{-1}AP$$

est une matrice diagonale, $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. On rappelle que les λ_i sont les valeurs propres de A et la $i$ième colonne de P est un vecteur propre pour λ_i.

Lemme 8. *Soit A diagonalisable et $k \in \mathbb{N}$. Alors, avec la notation d'en haut*

$$A^k = P \ \mathrm{diag}(\lambda_1^k, \cdots, \lambda_n^k) \ P^{-1}.$$

De plus, si A est inversible alors

$$A^{-1} = P \ \mathrm{diag}(\lambda_1^{-1}, \cdots, \lambda_n^{-1}) \ P^{-1}.$$

Démonstration. Si $k = 0$ on a $A^0 = 1_n$ et $\lambda_i^0 = 1$ par définition, ce qui implique le resultat. Si $k = 1$ alors $PDP^{-1} = PP^{-1}APP^{-1} = A$, et si $k > 1$

$$A^k = \overbrace{(PDP^{-1}) \cdots (PDP^{-1})}^{k-fois} = PD^kP^{-1}$$

car les PP^{-1} au milieu s'annullent. Clairement, $D^k = \mathrm{diag}(\lambda_1^k, \cdots, \lambda_n^k)$ pour tout $k \in \mathbb{N}$.

A est inversible si et seulement si tous les valeurs propres sont non-nulles. Dans ce cas D est inversible et

$$A^{-1} = (PDP^{-1})^{-1} = PD^{-1}P^{-1} = P \ \mathrm{diag}(\lambda_1^{-1}, \cdots, \lambda_n^{-1}) \ P^{-1}.$$

$\square$

Corollaire 11. *Soit A diagonalisable et $Q \in \mathbb{C}[X]$ un polynôme. Alors, avec la notation d'en haut*

$$Q(A) = P \ \mathrm{diag}(Q(\lambda_1), \cdots, Q(\lambda_n)) \ P^{-1}.$$

Remarque 4.1.1. *Soit A diagonalisable et $f : \mathbb{C} \to \mathbb{C}$ une fonction. Alors, avec la notation d'en haut on peut définir*

$$f(A) := P \ \mathrm{diag}(f(\lambda_1), \cdots, f(\lambda_n)) \ P^{-1}.$$

On obtient ainsi la possibilité de travailler avec des fonctions comme $\sin(A)$ *où* $\exp(A)$*. Ce* calcul fonctionelle *repose sur le fait que A est diagonalisable.*

4.2 Systèmes récurrents

Voici un example d'un système récurrent de premier ordre

$$\begin{aligned} u_{k+1} &= au_k + bv_k \\ v_{k+1} &= cu_k + dv_k \end{aligned}$$

où $a, b, c, d \in \mathbb{K}$ et $k \in \mathbb{N}$. Une solution d'un tel système consiste en deux suites $(u_k)_{k\in\mathbb{N}}$ et $(v_k)_{k\in\mathbb{N}}$ à valeur dans $\mathbb{K}$, qui satisfont ces deux équations. Si on fixe les valeurs u_0 et v_0 on rend la solution unique. u_0 et v_0 sont appelés conditions initiales.

Les deux équations du système sont couplés, car la valeur de u_{k+1} dépend de v_k et la valeur de v_{k+1} dépend de u_k. L'astuce pour résoudre le système est de prendre des combinaisons linéaires des suites (u_k) et (v_k) pour obténir un système découplé. Ceci est relié à la diagonalisation de la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. En effet, si $X_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix} \in \mathbb{K}^2$, alors $(X_k)_{k\in\mathbb{N}}$ est une suite de vecteurs de $\mathbb{K}^2$ et le système peut s'écrire $X_{k+1} = AX_k$ et les conditions initiales correspondent à la préscription du vecteur X_0.

Plus généralement un système récurrent et un système d'équations du type

$$X_{k+1} = AX_k \qquad k \in \mathbb{N} \tag{4.1}$$

dont l'inconnue est $X = (X_k)_{k\in\mathbb{N}}$ avec $X_k \in \mathbb{K}^n = M_{n\times 1}(\mathbb{K})$ et où $A \in M_n(\mathbb{K})$.

On suppose maintenant que A est diagonalisable, c.à.d. il existe n vecteurs propres linéairement indépendants. La matrice P qui a comme colonnes ces vecteurs propres diagonalise A : $D = P^{-1}AP = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. Alors avec $Y_k = P^{-1}X_k$ on obtient de (4.1)

$$Y_{k+1} = P^{-1}X_{k+1} = P^{-1}AX_k = DY_k \qquad k \in \mathbb{N} \tag{4.2}$$

Comme D est diagonale, le système (4.2) consistent à n récurrences découplées, la $i$ième composante de Y_{k+1} ne dépend que de la $i$ième composante de Y_k et nous pouvons résoudre chaqu'une de ces récurrences individuellement. Notons y_{ik} la $i$ième composante de Y_k on obtient alors, par récurrence $y_{ik} = \lambda_i^k y_{i0}$. Ceci revient à la solution

$$X_k = A^k X_0 = P\mathrm{diag}(\lambda_1^k, \cdots, \lambda_n^k)P^{-1}X_0. \tag{4.3}$$

du système d'origine.

4.3 Systèmes d'équations différentielles de premier ordre

Voici un example d'un système d'équations différentielles linéaires de premier ordre

$$\begin{aligned} u'(t) &= au(t) + bv(t) \\ v'(t) &= cu(t) + dv(t) \end{aligned}$$

où $a, b, c, d \in \mathbb{K}$ et $t \in \mathbb{R}^{\geq 0}$. Une solution d'un tel système consiste en deux fonctions dérivables $u(t)$ et $v(t)$ à valeur dans $\mathbb{K}$, qui dans ce contexte est $\mathbb{R}$ où $\mathbb{C}$. Si on fixe les valeurs $u(0)$ et $v(0)$ (où de leurs dérivés) on rend la solution unique. $u(0)$ et $v(0)$ sont appelés conditions initiales.

Plus généralement, un système d'équations différentielles linéaires de premier ordre est donné par

$$X'(t) = AX(t) \tag{4.4}$$

dont l'inconnue est une fonction vectorielle dérivable $X(t) \in \mathbb{K}^n = M_{n\times 1}(\mathbb{K})$ et où $A \in M_n(\mathbb{K})$. Les conditions initiales fixent $X(0)$ (où $X'(0)$). Vu que A ne dépend pas de t, on dit que ce système est à coefficients constants.

On suppose que A est diagonalisable, avec matrice P qui diagonalise A comme en haut, et introduit une nouvelle fonction vectorielle $Y(t) = P^{-1}X(t)$. On obtient alors un système découplé d'équations différentielles pour les composantes $y_i(t)$ de $Y(t)$

$$y_i'(t) = \lambda_i y_i(t) \tag{4.5}$$

où λ_i est la valeur propre de A associée à la i-ième colonne de P (qui est, on rappelle, un vecteur propre de A). On obtient alors les solutions

$$y_i(t) = e^{\lambda_i t} y_i(0) \tag{4.6}$$

Le système d'origine (4.7) a donc la solution

$$X(t) = P\mathrm{diag}(e^{\lambda_1 t}, \cdots, e^{\lambda_n t})P^{-1}X_0 \tag{4.7}$$

qui peut aussi être écrite sous la forme $X(t) = e^{At}X(0)$.

4.4 Équation différentielle linéaire d'ordre k

En s'intéresse maintenant à une équation différentielle du type

$$f^{(k)}(t) = a_{k-1}f^{(k)}(t) + \cdots + a_0 f(t) \tag{4.8}$$

où $f : \mathbb{R} \to \mathbb{C}$ est une fonction k-fois dérivable (on note la k-ième dérivée $f^{(k)}(t)$). Pour résoudre cette équation on définit la fonction vectorielle

$$X(t) = \begin{pmatrix} f^{(k-1)}(t) \\ \vdots \\ f(t) \end{pmatrix}.$$

Alors (4.8) s'écrit

$$X'(t) = \begin{pmatrix} f^{(k)}(t) \\ \vdots \\ f^{(0)}(t) \end{pmatrix} = AX(t) = \begin{pmatrix} a_{k-1} & a_{k-2} & \cdots & a_0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} f^{(k-1)}(t) \\ \vdots \\ f(t) \end{pmatrix}. \tag{4.9}$$

Si A est diagonalisable, cet équation devient

$$X'(t) = P \begin{pmatrix} \lambda_1 & & 0 \\ \vdots & \ddots & \\ 0 & \cdots & \lambda_n \end{pmatrix} P^{-1} X(t)$$

où $\lambda_1, \cdots, \lambda_k$ sont les valeurs propores de A et $P = (v_1 \cdots v_k)$ est la matrice dont la i-ième colonne v_i correspond au vecteur propre associé à λ_i. Si une valeur propre à une multiplicité algébrique m, alors il faut trouver m vecteurs propres indépendants pour cette valeur propre. La solution est alors

$$X(t) = P \begin{pmatrix} e_1^\lambda & & 0 \\ \vdots & \ddots & \\ 0 & \cdots & e_n^\lambda \end{pmatrix} P^{-1} X(0)$$

On va voir plutard comment résoudre l'équation si A n'est pas diagonalisable.

Chapitre 5

Sous-espaces stables

Dans ce chapitre E désigne un espace vectoriel sur $\mathbb{K}$ et u un endomorphisme de E.

5.1 Notion de sous-espace stable

Définition 5.1.1. Soit u un endomorphisme de E et F un sous-espace vectoriel de E. On dit que F est stable par u si $u(F) \subset F$, c.à.d. pour tout $x \in F$, $u(x) \in F$.

Une autre terminologie qu'on trouve est sous-espace invariant pour sous-espace stable.

Proposition 12. *Soient F_1 et F_2 deux sous-espaces stables pour u alors $F_1 + F_2$ et $F_1 \cap F_2$ sont aussi stable pour u.*

Démonstration. Soit $x = x_1 + x_2$ avec $x_i \in F_i$ alors $u(x) = u(x_1) + u(x_2) \in F_1 + F_2$.
Soit $f \in F_1 \cap F_2$. Alors $u(x) \in F_1$ et $u(x) \in F_2$, d'où le resultat. □

Voici quelques exemples.

1. Clairement, $F = \{0_E\}$ et $F = E$ sont des sous-espaces stables par n'importe quel endomorphisme.
2. Tout sous-espace de E est stable pour id.
3. Soit $x \in E$, $x \neq 0_E$, et $F = \text{Vect}(x) = \{\lambda x : \lambda \in \mathbb{K}\}$. F est un sous-espace de dimension 1. Clairement F est stable par u si et seulement si il existe λ t.q. $u(x) = \lambda x$, c.à.d. x est un vecteur propre de u. Donc les seuls sous-espaces stables pour u, qui sont de dimension 1, sont les espaces engendrés par ses vecteurs propres.
4. Un espace propre $E_\lambda(u)$ est stable par u (Prop. 12).

Proposition 13. *Soient u et v deux endomorphismes de E tels que $u \circ v = v \circ u$ alors* $\ker(v)$ *et* $\text{Im}(v)$ *sont stables par u.*

Démonstration. Soit $x \in \ker(v)$. Alors $v(u(x)) = u(v(x)) = u(0) = 0$ donc $u(x) \in \ker(v)$. Soit $y \in \text{Im}(v)$. Alors il existe $x \in E$ tel que $v(x) = y$. Alors $u(y) = u(v(x)) = v(u(x)) \in \text{Im}(v)$. □

5.1.1 Espace caractéristique

Lemme 9. *Soit $k \in \mathbb{N}$.* $\text{im}\ u^k$ *est un sous-espace stable par u et*

$$\text{im}\ u^{k+1} \subset \text{im}\ u^k.$$

Si E est de dimension n, alors il existe $q \leq n$ t.q. pour tout $k \geq q$ on a $\text{im}\ u^k = \text{im}\ u^q$.

Démonstration. On prenant $v = u^k$ dans Prop. 13 on trouve que im u^k est stable. Soit $y \in$ im u^{k+1}. Il existe donc $x \in E$ t.q. $y = u^{k+1}(x)$. Donc $y = u^k(z)$ avec $z = u(x) \in E$.

Soit $r_k =$ rang u^k. L'inclusion im $u^{k+1} \subset$ im u^k montre que $r_{k+1} \leq r_k$ et on a égalité si et seulement si im $u^{k+1} =$ im u^k. Supposons que im $u^{k+1} =$ im u^k. Alors im $u^{k+2} = u(\text{im } u^{k+1}) = u(\text{im } u^k) = \text{im } u^{k+1} = \text{im } u^k$. Donc si $r_{k+1} = r_k$ alors $r_m = r_k$ pour tout $m \geq k$. La suite $(r_k)_k$ est donc strictement décroissante jusqu'à ce quelle atteint sa valeur minimale. Comme $r_0 = n$ le premier k pour lequel r_k est minimal doit être plus petit ou égal à n. □

Lemme 10. *Soit* $k \in \mathbb{N}$. $\ker u^k$ *est un sous-espace stable par* u *et*

$$\ker u^k \subset \ker u^{k+1}.$$

Si E *est de dimension* n*, alors il existe* $q \leq n$ *t.q. pour tout* $k \geq q$ *on a* $\ker u^k = \ker u^q$.

Démonstration. On prenant $v = u^k$ dans Prop. 13 on trouve que $\ker u^k$ est stable.

L'inclusion $\ker u^k \subset \ker u^{k+1}$ montre que $\dim \ker u^k \leq \dim \ker u^{k+1}$ et on a égalité si et seulement si $\ker u^{k+1} = \ker u^k$. Par le théorème du rang on a $\dim \ker u^k = n - r_k$. Le deuxième resultat découle alors du dernier lemme. □

Définition 5.1.2. Soit λ une valeur propre de u. Soit q_λ le plus petit entier naturel t.q. pour tout $k \geq q$ on a $\ker(u - \lambda \text{id})^k = \ker(u - \lambda \text{id})^{q_\lambda}$. On appelle $F_\lambda(u) := \ker(u - \lambda \text{id})^{q_\lambda}$ l'espace caractéristique pour la valeur propre λ de u (où simplement espace caractéristique de λ).

Corollaire 12. $F_\lambda(u)$ *est un sous-espace stable pour* u*. Il coincide avec l'espace propre de* λ, $F_\lambda(u) = E_\lambda(u)$*, si et seulement si* $q_\lambda = 1$.

Démonstration. Si $q_\lambda = 1$ alors $F_\lambda(u) = E_\lambda(u)$ par définition. Si $F_\lambda(u) = E_\lambda(u)$ alors $\ker(u - \lambda \text{id})^k = \ker(u - \lambda \text{id})$ pour tout $k \geq 1$, alors $q_\lambda = 1$. □

5.1.2 Espace cyclique

Si $x \neq 0_E$ n'est pas un vecteur propre de u alors $\text{Vect}(x)$ n'est pas stable par u. Le plus petit sous-espace de E qui contient x est appellé *l'espace cyclique* engendré par u. On le note $\langle x \rangle_u$.

Lemme 11. *Soit* $0_E \neq x \in E$ *et* q *le plus petit entier naturel t.q. la famille* $\{x, u(x), \cdots, u^q(x)\}$ *est liée (n'est pas libre). Alors*

$$\langle x \rangle_u = \text{Vect}(x, u(x), \cdots, u^{q-1}(x))$$

En particulier, $q = \dim \langle x \rangle_u$.

Démonstration. Par définition de q la famille $\{x, u(x), \cdots, u^q(x)\}$ est liée. Donc $u^q(x)$ est combinaison linéaire des $x, u(x), \cdots, u^{q-1}(x)$. Donc $\text{Vect}(x, u(x), \cdots, u^{q-1}(x))$ contient $u^k(x)$ pour tout $k \geq 0$. $\text{Vect}(x, u(x), \cdots, u^{q-1}(x))$ est alors stable par u et contient $\langle x \rangle_u$. De l'autre coté, $\langle x \rangle_u$ doit contenir la famille $\{x, u(x), \cdots, u^{q-1}(x)\}$. Donc $\text{Vect}(x, u(x), \cdots, u^{q-1}(x)) \subset \langle x \rangle_u$. Par définition de q, $\{x, u(x), \cdots, u^{q-1}(x)\}$ est libre et donc une base pour $\text{Vect}(x, u(x), \cdots, u^{q-1}(x))$. Par conséquence, $q = \dim \langle x \rangle_u$. □

5.1.3 Sous-espace irréductible

Définition 5.1.3. Un sous espace irréductible de u est un sous-espace stable (non-trivial) pour u qui n'est pas somme directe de deux sous-espaces stables (non-triviaux) pour u.

Donc si $F \subset E$ est un sous-espace stable pour u et $F = F_1 \oplus F_2$ où F_1 et F_2 sont des sous-espaces (different de $\{0_E\}$) stable pour u, alors F n'est pas irréductible.

Le lemme suivant, qu'on va montrer plus bas, dit en particulier, que pour les endomorphismes diagonalisables, les sous-espaces irreducibles sont les plus petits sous-espaces stables (non-triviaux).

Lemme 12. *Soit u un endomorphisme diagonalisable, alors tout sous-espace irreductible est de dimension 1 et engendré par un vecteur propre de u.*

5.1.4 Endomorphisme induit

Définition 5.1.4. On suppose que F est stable par u. On définit alors $u_F : F \to F$ l'application qui envoie $f \in F$ sur $u(f)$. On appelle u_F l'endomorphisme induit sur F par u.

Attention : Il ne faut pas confondre la notion de l'endomorphisme induit sur F, u_F, avec sa restriction sur F, $u|_F$. La restriction de u sur F est l'application linéaire $u|_F : F \to E$ donné par la même formule $u|_F(x) = u(x)$ mais l'espace d'arrivé est différent. Par exemple, id_F est une bijection de F vers F, pendant que $\mathrm{id}|_F$ est l'inclusion de F dans E. $\mathrm{id}|_F$ n'est pas bijective si $F \neq E$.

Exemple 5.1.5. *Soit E un espace vectoriel avec une base $\mathcal{B} = \{b_1, b_2, b_3, b_4\}$. Soit u un endomorphisme de E dont la matrice dans la base $\mathcal{B}$ est $A = \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ -3 & 3 & -1 & 2 \end{pmatrix}$. Cette matrice nous indique que $u(b_1) = b_1 - 3b_4$ et $u(b_4) = b_1 + 2b_4$. Ainsi, si on note $F = \mathrm{Vect}\{b_1, b_4\}$ alors F est stable par u.*

L'endomorphisme induit u_F est l'application $u_F : \mathrm{Vect}\{b_1, b_4\} \to \mathrm{Vect}\{b_1, b_4\}$ donné par $u_F(b_1) = b_1 - 3b_4$ et $u_F(b_4) = b_1 + 2b_4$ et sa matrice dans la base $\mathcal{B}' = \{b_1, b_4\}$ est

$$[u_F]_{\mathcal{B}'} = \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}.$$

La restriction $u|_F$ est l'application linéaire $u|_F : \mathrm{Vect}\{b_1, b_4\} \to \mathrm{Vect}\{b_1, b_2, b_3, b_4\}$ donné par $u|_F(b_1) = b_1 - 3b_4$ et $u(b_4) = b_1 + 2b_4$ et sa matrice dans les bases $\mathcal{B}'$, $\mathcal{B}$ est

$$[u|_F]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ -3 & 2 \end{pmatrix}.$$

Proposition 14. *Soient u et v deux endomorphismes de E et F un sous-espace stable par u et v et soit $\lambda \in \mathbb{K}$. Alors F est stable par $\lambda u + v$ et par $u \circ v$. De plus, $(\lambda u + v)_F = \lambda u_F + v_F$ et $(u \circ v)_F = u_F \circ v_F$.*

Démonstration. Soit $x \in F$. Alors $(\lambda u + v)(x) = \lambda u(x) + v(x) \in F$ car F est un sous-esp. vect. de E. D'autre part, $x \in F$ donc $v(x) \in F$ et donc $u(v(x)) \in F$. Les deux égalités restantes sont directes. $\square$

Proposition 15. *On a* $\ker(u_F) = \ker(u) \cap F$*. D'autre part* $\mathrm{im}\ (u_F) \subset \mathrm{im}\ (u) \cap F$ *mais cette inclusion est stricte en générale.*

Démonstration. Soit $x \in \ker(u_F)$. Alors $x \in F$ et $u_F(x) = 0$, i.e. $u(x) = 0$. D'où $x \in \ker(u) \cap F$. La réciproque est directe.
L'inclusion sur l'image est directe. Montrons, à l'aide d'un exemple, que ce n'est pas une égalité en général. Soient E un espace vectoriel de dimension 3, $\mathcal{B} = (b_1, b_2, b_3)$ une base de E, $F = \mathrm{Vect}\{b_1, b_2\}$ et u l'endomorphisme de E tel que $u(b_1) = u(b_2) = b_1$ et $u(b_3) = b_2$. On voit directement que F est stable par u. De plus, $\mathrm{im}\ (u_F) = \mathrm{Vect}\{b_1\}$ et $\mathrm{im}\ (u) \cap F = F$. □

Corollaire 13. *Soit F un sous-espace vectoriel de E stable par u. Tout vecteur propre et toute valeur propre de u_F est vecteur propre et valeur propre de u. En particulier, si u est injectif alors u_F l'est aussi.*

Démonstration. Soit $0_F \neq x \in F$, $\lambda \in \mathbb{K}$ solution de l'équation propre pour u_F, $u_F(x) = \lambda x$. Alors $0_E \neq x \in E$ et $u(x) = \lambda x$, c. a.d. x est vecteur propre pour la valeur propre λ pour u. En particulier, si u_F n'est pas injective 0 est une valeur propre pour u_F, donc aussi pour u, donc u n'est pas injective. □

5.1.5 Sous espaces stable et forme triangulaire par blocs

Proposition 16. *Soit F un sous-espace vectoriel de E de dimension p. Soit $\mathcal{B} = (b_1, \ldots, b_p, \ldots, b_n)$ une base de E telle que $\mathcal{B}_F = (b_1, \ldots, b_p)$ soit une base de F. Alors F stable par u si et seulement si $[u]_{\mathcal{B}}$ est triangulaire par blocs, i.e. de la forme*

$$[u]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

où A est une matrice $p \times p$. Dans ce cas,

$$A = [u_F]_{\mathcal{B}_F}.$$

De plus, si $G = \mathrm{Vect}(b_{p+1}, \cdots, b_n)$ est aussi stable par u alors $[u]_{\mathcal{B}}$ est diagonale en blocs

$$[u]_{\mathcal{B}} = \begin{pmatrix} [u_F]_{\mathcal{B}_F} & 0 \\ 0 & [u_G]_{\mathcal{B}_G} \end{pmatrix}.$$

où $\mathcal{B}_G = (b_{p+1}, \cdots, b_n)$.

Démonstration. On suppose d'abord que F est stable pour u. Soit $i \leq p$. La i-ième colonne de u est $[u(b_i)]_{\mathcal{B}}$. Comme $u(b_i) \in F = \mathrm{Vect}(b_1, \cdots, b_p)$ les seulement les p premiers coefficients de cette colonne sont non-nuls. De plus, comme $u(b_i) = u_F(b_i)$ ces premiers coefficients sont donnés par la i-ième colonne de $[u_F]_{\mathcal{B}_F}$. Ceci montre la forme triangulaire en bloque. Si G est aussi stable pour u le même raisonnement s'applique au $n - p$ derniers coefficients de la j-ième colonne si $j > p$. D'où la forme diagonale en bloque.

Supposons maintenant que $[u]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$. Ceci dit que, pour $i \leq p$, dans la i-ième colonne $[u(b_i)]_{\mathcal{B}}$ seulement les p premiers coefficients sont non-nuls. Donc $u(b_i) \in \mathrm{Vect}(b_1, \cdots, b_p) = F$. Donc $u(F) \subset F$. □

Remarque 5.1.6. *La forme triangulaire est supérieure, car ce sont les p premiers vecteurs de la base $\mathcal{B}$ qui forment une base pour le sous-espace F, qui est stable pour u. Si on prendrait une base pour E t.q. les p derniers vecteurs forment une base pour F, alors la forme triangulaire serait inférieure.*

5.2 Polynômes annulateurs

5.2.1 Généralités sur les polynômes

Un polynôme non-nul est de la forme $P(X) = a_n X^n + \cdots + a_0$ avec $a_n \neq 0$. Le degré de ce polynôme, qu'on note $\deg P$, est n, et si $a_n = 1$ alors on dit que P est *unitaire*. Si $F \subset \mathbb{K}[X]$ nous posons $F^{unit} = \{P \in F : P \text{ est unitaire}\}$. Un polynôme $P \in F^{unit}$ est de degré minimal si pour tout autre polynôme $Q \in F^{unit}$ on a $\deg P \leq \deg Q$.

$\mathbb{K}[X]$ est une algèbre, c.à.d. un espace vectoriel avec un produit, le produit des polynômes. On dit qu'un sous-espace $I \subset \mathbb{K}[X]$ est un idéal si $I \neq \{0\}$ et pour tout $P \in I$ et $Q \in \mathbb{K}[X]$ on a $PQ \in I$.

Exemple 5.2.1. *Soit $\lambda \in \mathbb{K}$ et $I = \{P \in \mathbb{K}[X] : P(\lambda) = 0\}$. Alors I est un idéal.*

Proposition 17. *Soit F un sous-espace non-trivial de $\mathbb{K}[X]$. Alors F^{unit} contient un unique élément P de degré minimal. On appelle cet élément l'élément minimale de F.*

Démonstration. Tout sous-espace F non-trivial de $\mathbb{K}[X]$ contient un polynôme unitaire. Donc F^{unit} n'est pas vide. F^{unit} admet alors un polynôme de degré minimal. En effet, choisissons $P_0 \in F^{unit}$. Soit $n_0 = \deg P_0$. Si P_0 n'est pas de degré minimal, alors il existe P_1 avec $n_1 = \deg P_1 < n_0$. Si P_1 n'est pas de degré minimal, alors il existe P_2 avec $n_2 = \deg P_2 < n_2$. On peut continuer ainsi au plus n_0 fois, car le degré est un nombre positif. Le dernier polynôme ainsi choisi est alors de degré minimal. Montrons l'unicité de ce polynôme.

Soient $P, Q \in F^{unit}$ de degré minimal n. On pose $R = P - Q$. Si $R \neq 0$ il existe $0 \neq \lambda \in \mathbb{K}$ t.q. $\lambda R \in F^{unit}$. D'où $\deg R = \deg \lambda R \geq n$. Mais le coefficient devant X^n du polynôme $P - Q$ est 0, donc $\deg R < \deg P = n$. Ceci étant une contradiction on doit avoir $R = 0$. □

Proposition 18. *Soit I un idéal de $\mathbb{K}[X]$. L'élément minimale de I divise tout autre élément non-nul de I.*

Démonstration. Soit $m \in I^{unit}$ l'élément minimal de I (Prop. 17). Soit $0 \neq P \in I$. Alors $\deg P \geq \deg m$. Il existe alors des polynômes Q, R t.q.

$$P = mQ + R, \quad \deg(R) < \deg(m)$$

(algorithme d'Euclid). Comme I est un idéal on a $mQ \in I$, donc $R = P - mQ \in I$. Si $R \neq 0$ il existe $\lambda \in \mathbb{K}$ t.q. $\lambda R \in I^{unit}$. D'où $\deg R = \deg \lambda R \geq \deg m$, une contradiction. Donc $R = 0$. □

Le plus grand diviseur commun $R = \text{pgcd}(P_1, P_2)$ de deux polynômes $P_1, P_2 \in \mathbb{K}[X]$ est le polynôme unitaire de degré maximal, qui satisfait $P_1 = RQ_1$, $P_2 = RQ_2$, pour $Q_1, Q_2 \in \mathbb{K}[X]$.

Théorème 13 (Bézout)**.** *Soient $P_1, P_2 \in \mathbb{K}[X]$ deux polynômes. Alors ils existent deux polynômes $Q_1, Q_2 \in \mathbb{K}[X]$ t.q.*

$$P_1 Q_1 + P_2 Q_2 = \text{pgcd}(P_1, P_2).$$

Démonstration. Soient $P_1, P_2 \in \mathbb{K}[X]$. Soit $\mathcal{S} : \mathbb{K}[X] \times \mathbb{K}[X] \to \mathbb{K}[X]$ l'application donnée par

$$\mathcal{S}(Q_1, Q_2) = P_1 Q_1 + P_2 Q_2.$$

Montrons d'abord que l'image de $\mathcal{S}$, $I := \text{im}\, \mathcal{S}$ est un idéal. En effet, I est un sous-espace, car

$$\mathcal{S}(Q_1, Q_2) + \lambda \mathcal{S}(Q_1', Q_2') = \mathcal{S}(Q_1 + \lambda Q_1', Q_2 + \lambda Q_2'),$$

et

$$\mathcal{S}(Q_1, Q_2) P = P_1 Q_1 P + P_2 Q_2 P = \mathcal{S}(Q_1 P, Q_2 P).$$

Soit m l'élément minimal de I^{unit}. Par Prop. 18 m divise $\mathcal{S}(Q_1, Q_2)$ pour tout choix de Q_1, Q_2 t.q. au moins un des deux n'est pas nul. m divise en particulier $P_1 = \mathcal{S}(1, 0)$ et $P_2 = \mathcal{S}(0, 1)$ et donc aussi $\mathrm{pgcd}(P_1, P_2)$.

Comme $\mathrm{pgcd}(P_1, P_2)$ divise P_1 et P_2 il divise tout élément non-nul de I. En particulier $\mathrm{pgcd}(P_1, P_2)$ divise m. Ceci montre que $m = \mathrm{pgcd}(P_1, P_2)$. Donc $\mathrm{pgcd}(P_1, P_2) \in \mathrm{im}\,\mathcal{S}$. Donc il existe $Q_1, Q_2 \in \mathbb{K}[X]$ t.q. $\mathrm{pgcd}(P_1, P_2) = \mathcal{S}(Q_1, Q_2)$. □

5.2.2 Polynôme d'endomorphisme

On rappelle que deux endomorphismes u, v d'un même espace vectoriel peuvent être composés, $u \circ v$, on dit aussi que $u \circ v$ est le produit de u avec v. On note

$$u^k = \overbrace{u \circ \cdots \circ u}^{k-fois}, \quad k \geq 1, \quad \text{et} \quad u^0 = \mathrm{id}_E.$$

Définition 5.2.2. Soit $P = a_d X^d + \cdots + a_1 X + a_0 \in \mathbb{K}[X]$ un polynôme. On définit

$$P(u) = a_d u^d + \cdots a_1 u + a_0 \mathrm{id}_E.$$

$P(u)$ est un endomorphisme de E. On dit qu'un endomorphisme v de E est un polynôme en u si il existe $P \in \mathbb{K}[X]$ tel que $v = P(u)$.

Lemme 13. *Soient u, v, w des endomorphismes de E et $\lambda \in \mathbb{K}$. Alors*

$$u \circ (v + \lambda w) = u \circ v + \lambda u \circ w. \tag{5.1}$$

Démonstration. Soit $x \in E$. Alors

$$u \circ (v + \lambda w)(x) = u((v + \lambda w)(x)) = u(v(x) + \lambda w(x)) = u(v(x)) + \lambda u(w(x)).$$

□

Proposition 19. *Soient $P, Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. Alors*

1. $(\lambda P)(u) = \lambda P(u)$,
2. $(P + Q)(u) = P(u) + Q(u)$,
3. $(PQ)(u) = P(u) \circ Q(u)$.

Démonstration. (1) et (2) sont assez clair.

(3) On suppose d'abord que $P[X] = X^n$ et $Q(X) = X^m$. Alors $PQ(u) = u^{n+m} = u^n \circ u^m = P(u) \circ Q(u)$. Soient maintenant $P(X) = \sum_{i=0}^{n} a_i X^i$ et $Q(X) = \sum_{j=0}^{m} b_j X^j$, alors

$$PQ(X) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j X^{i+j}.$$

Donc

$$P(u) \circ Q(u) = \left(\sum_{i=0}^{n} a_i u^i \right) \circ \left(\sum_{j=0}^{m} b_j u^j \right) \overset{(5.1)}{=} \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j u^i \circ u^j = PQ(u).$$

□

Une matrice A de taille $n \times n$ étant un endomorphisme de $\mathbb{K}^n$ on peut aussi évaluer un polynôme en A.

5.2.3 Polynôme annulateur

Définition 5.2.3. Soit u un endomorphisme de E. Un polynôme P est appellé annulateur de u si

$$P(u) = 0_{\mathrm{End}(E)}.$$

Ici $0_{\mathrm{End}(E)}$ est l'endomorphisme nul sur E, c.à.d. pour tout $x \in E$, $0_{\mathrm{End}(E)}(x) = 0_E$.

Exemple 5.2.4. *1. Le polynôme nul est annulateur de tout endomorphisme.*

2. Si u est un projecteur, i.e. $u^2 = u$, alors $X^2 - X$ est annulateur de u.

3. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Alors le polynôme $P = X^2 - (a+d)X + ad - bc = X^2 - \mathrm{Tr}(A)X + \det(A)$ est annulateur de A.

Lemme 14. *Soit A, Q deux matrices $n \times n$ et $P \in \mathbb{K}[X]$ un polynôme. On suppose que Q est inversible. Alors $P(Q^{-1}AQ) = Q^{-1}P(A)Q$. En particulier, si P est annulateur de A il est aussi annulateur de $Q^{-1}AQ$.*

Démonstration. On suppose d'abord que $P[X] = X^k$. Alors $(Q^{-1}AQ)^k = Q^{-1}A^kQ$. Maintenant le resultat découle du fait que $Q^{-1}(A+B)Q = Q^{-1}AQ + Q^{-1}BQ$. □

Théorème 14 (Théorème de Cayley-Hamilton)**.** *Soit u un endomorphism de l'espace vectoriel E. On suppose que la dimension de E est finie égale à n. Le polynôme caractéristique de u, P_u, est annulateur de u, c.à.d. $P_u(u) = 0$.*

Pour la preuve de ce théorème nous avons besoin d'un lemme.

Lemme 15. *Soit $p \geq 1$, $a_1, \cdots, a_p \in \mathbb{K}$ et*

$$A_p = \begin{pmatrix} 0 & 0 & \cdots & a_p \\ 1 & 0 & \cdots & a_{p-1} \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & a_1 \end{pmatrix}.$$

Alors le polynôme caractéristique de A_p est $P_{A_p}(\lambda) = (-1)^p \lambda^p + (-1)^{p-1}\left(a_1\lambda^{p-1} + \cdots + a_p\right)$.

Démonstration. On developpe selon la première colonne :

$$P_{A_p}(\lambda) = \begin{vmatrix} -\lambda & 0 & \cdots & a_p \\ 1 & -\lambda & \cdots & a_{p-1} \\ 0 & \ddots & -\lambda & \vdots \\ 0 & \cdots & 1 & a_1 - \lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 0 & \cdots & a_{p-1} \\ 1 & -\lambda & \cdots & a_{p-2} \\ 0 & \ddots & -\lambda & \vdots \\ 0 & \cdots & 1 & a_1 - \lambda \end{vmatrix} - \begin{vmatrix} 0 & 0 & \cdots & a_p \\ 1 & -\lambda & \cdots & a_{p-2} \\ 0 & \ddots & -\lambda & \vdots \\ 0 & \cdots & 1 & a_1 - \lambda \end{vmatrix}.$$

Le dernier déterminant est $(-1)^p a_p$. D'où

$$\begin{aligned} P_{A_p}(\lambda) &= -\lambda P_{A_{p-1}}(\lambda) + (-1)^{p-1} a_p \\ &= \lambda^2 P_{A_{p-2}}(\lambda) + (-1)^{p-1}(\lambda a_{p-1} + a_p) \end{aligned}$$

et le resultat suit par récurrence. □

Démonstration du Théorème 14. On va montrer que, pour tout $x \in E$, $P_u(u)(x) = 0$. Soit donc $x \in E$. Si $x = 0$ alors c'est trivial ; on suppose donc $x \neq 0$. Soit $F = \langle x \rangle_u$ l'espace cyclique engendré par x. Si $p = \dim F$ alors $\mathcal{B} = \{x, u(x), u^2(x), \ldots, u^{p-1}(x)\}$ est une base pour F. u^p est donc combinaison linéaire des u^k avec $k < p$, c.à.d. ils existent $\lambda_0, \cdots, \lambda_{p-1} \in \mathbb{K}$ t.q.

$$u^p(x) = \lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{p-1} u^{p-1}(x). \tag{5.2}$$

On complète $\mathcal{B}$ en une base $\mathcal{D}$ de E. Comme F est un sous-espace stable de u la matrice de u dans la base $\mathcal{D}$ est triangulaire par blocs

$$[u]_{\mathcal{D}} = \begin{pmatrix} [u_F]_{\mathcal{B}} & B \\ 0 & D \end{pmatrix}.$$

La i-ième colonne de la matrice $[u_F]_{\mathcal{B}}$ est $[u(u^{i-1}(x))]_{\mathcal{B}}$. On trouve

$$[u^i(x)]_{\mathcal{B}} = e_{i+1}, \text{ si } i < p \quad \text{et} \quad [u^p(x)]_{\mathcal{B}} = \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{p-1} \end{pmatrix}$$

Dóù

$$[u_F]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & \lambda_0 \\ 1 & 0 & \cdots & \lambda_1 \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & \lambda_{p-1} \end{pmatrix}$$

matrice qui a exactement la forme de A_p dans Lemme 15 avec $\lambda_i = a_{p-i}$. Donc $P_{u_F}(X) = (-1)^p(X^p - (\lambda_{p-1}X^{p-1} + \cdots + \lambda_0))$. Eq. 5.2 implique alors que $P_{u_F}(u) = 0_{End(E)}$. De plus, $P_u = P_{u_F} P_D$. Donc, pour tout $x \in E$,

$$P_u(u)(x) = P_{u_F}(u)(P_D(u)(x)) = 0_E.$$

□

5.2.4 Polynôme minimal

Lemme 16. *Soit u un endomorphisme d'un espace vectoriel. On suppose que u admet un polynôme annulateur non-nul. L'ensemble $\mathcal{A}_u$ des polynômes annulateur de u est un idéal.*

Démonstration. Soit $P, Q \in \mathbb{K}[X]$. Si $P(u) = 0_{\mathrm{End}(E)}$ et $Q(u) = 0_{\mathrm{End}(E)}$ alors $(P + \lambda Q)(u) = P(u) + \lambda Q(u) = 0_{\mathrm{End}(E)}$ d'où $\mathcal{A}_u$ est un sous-espace de $\mathbb{K}[X]$. Si $P(u) = 0_{\mathrm{End}(E)}$ et Q est quelconque $PQ(u) = P(u) \circ Q(u) = 0_{\mathrm{End}(E)}$, d'où $\mathcal{A}_u$ est un idéal. □

Sous l'hypothèse que u admet un polynôme annulateur non-nul $\mathcal{A}_u$ contient donc un élément minimal. Le théorème de Cayley-Hamilton implique que cette hypothése est satisfaite, si l'espace vectoriel est de dimension finie.

Définition 5.2.5. Soit u un endomorphisme d'un espace vectoriel. On suppose que u admet un polynôme annulateur non-nul. Le *polynôme minimal de u* est l'élément minimal de $\mathcal{A}_u$. On le note m_u.

Le polynôme minimal m_u est caractérisé par les propriétées suivantes :

1. m_u est unitaire (par définition),
2. $m_u(u) = 0$ (par définition),
3. Pour tout polynôme P annulateur de u, on a $\deg(m_u) \leq \deg(P)$.

Par Prop. 18, m_u divise tout annulateur de u.

Exemple 5.2.6. *Le polynôme minimal de $u = \mathrm{id}$ est $m_{\mathrm{id}}(X) = X - 1$. Le polynôme minimal d'un projecteur $u \neq \mathrm{id}_E, 0_{End(E)}$ est $m_u(X) = X^2 - X$.*

Théorème 15. *Soit P un annulateur de u. Alors $P(\lambda) = 0$ pour toute valeur propre λ de u.*

Démonstration. Soit λ une valeur propre de u avec vecteur propre $x \neq 0_E$. On écrit $P = \sum_{i=0}^{k} a_i X^i$. Alors

$$0_E = P(u)(x) = \sum_{i=0}^{k} a_i u^i(x) = \sum_{i=0}^{k} a_i \lambda^i x = P(\lambda)x.$$

Comme $x \neq 0_E$ ceci implique $P(\lambda) = 0$. □

Ce théorème peut être utile pour chercher les valeurs propres. En effect, il se peut que c'est plus facile de trouver un polynôme annulateur de u que de déterminer son polynôme caractéristique (qui n'existe que dans le cas de dimension finie). Par exemple, si u est un projecteur, alors $P(X) = X^2 - X$ est un annulateur de u. En en déduit que les seules valeurs propres possible sont 0 et 1.

Théorème 16. *Soit u un endomorphisme d'un espace vectoriel de dimension finie. Les valeurs propres de u sont exactement les racines de m_u.*

Ainsi le polyôme caractéristique P_u et le polynôme minimal m_u ont exactement les mêmes racines.

Démonstration. Comme P_u est annulateur de u, m_u divise P_u, c.à.d. il existe un polynôme Q t.q. $P_u = Qm_u$. Soit λ une racine de m_u. Alors $P_u(\lambda) = Q(\lambda)m_u(\lambda) = 0$. D'où λ est une racine de P_u et donc valeur propre de u.

La contraposée suit du dernier théorème car m_u est un annulateur de u. □

Proposition 20. *Soit u un endomorphisme de E. Soit F un sous-espace vectoriel de E stable par u et u_F l'endomorphisme induit.*

1. *F est stable par tout polynôme en u et pour tout $Q \in \mathbb{K}[X]$, $Q(u)_F = Q(u_F)$.*
2. *Le polynôme minimal de u_F, m_{u_F}, divise m_u.*

Démonstration.

1. F est stable par u donc par ses puissances u^k et donc par tout polynôme en u.
2. Pour tout $x \in F$, $m_u(u_F)(x) = m_u(u)(x) = 0$. Le polynôme minimal de u est donc annulateur de u_F. Le polynôme minimal de u_F divise alors le polynôme minimal de u.

□

5.2.5 Lemme des noyaux

Proposition 21. *Soit u un endomorphisme. Soient $P_1, P_2 \in \mathbb{K}[X]$ deux polynômes premiers entre eux (c.à.d.* $\operatorname{pgcd}(P_1, P_2) = 1$*). Alors*

$$\ker(P_1P_2(u)) = \ker(P_1(u)) \oplus \ker(P_2(u)).$$

Démonstration. Soient $P_1, P_2 \in \mathbb{K}[X]$ deux polynômes t.q. $\operatorname{pgcd}(P_1, P_2) = 1$. Par le théorème de Bézout il existe $Q_1, Q_2 \in \mathbb{K}[X]$ tels que $Q_1P_1 + Q_2P_2 = 1$. On peut donc écrire, pour tout $x \in E$,

$$x = \operatorname{id}(x) = (Q_1P_1 + Q_2P_2)(u)(x) = (Q_1P_1(u)(x)) + (Q_2P_2(u)(x)) \tag{5.3}$$

On a $\ker(P_1(u)) \subseteq \ker(P_2(u) \circ P_1(u)) = \ker((P_2P_1)(u))$ et d'une manière similaire $\ker(P_2(u)) \subseteq \ker((P_1P_2)(u))$. D'où $\ker(P_1(u)) + \ker(P_2(u)) \subseteq \ker(P_1P_2(u))$.

Soit $x \in \ker((P_1P_2)(u))$. Posons $x = a + b$ avec $a = (Q_1P_1(u))(x)$ et $b = (Q_2P_2(u))(x)$ (voir (5.3)). On a alors

$$P_1(u)(b) = P_1(u)(Q_2P_2(u))(x)) = P_1Q_2P_2(u)(x) = Q_2(u)(P_1P_2(u)(x)) = 0_E$$

et d'une manière similaire

$$P_2(u)(a) = Q_1(u)(P_1P_2(u)(x)) = 0_E.$$

Donc $\ker((P_1P_2)(u)) \subseteq \ker(P_1(u)) + \ker(P_2(u))$. Donc $\ker((P_1P_2)(u))$ est la somme $\ker(P_1(u)) + \ker(P_2(u))$. Il reste alors à montrer que la somme est directe.

Soit $x \in \ker(P_1(u)) \cap \ker(P_2(u))$. Avec (5.3) on obtient

$$x = Q_1(u)(P_1(u)(x)) + Q_2(u)(P_2(u)(x)) = 0_E$$

d'où $\ker(P_1(u)) + \ker(P_2(u)) = \ker(P_1(u)) \oplus \ker(P_2(u))$ (la somme est directe). □

Corollaire 14 (Lemme des noyaux)**.** *Soit u un endomorphisme. Soient $P_1, \dots, P_m \in \mathbb{K}[X]$ premiers entre eux deux à deux. Alors*

$$\ker(P_1 \cdots P_m(u)) = \ker(P_1(u)) \oplus \cdots \oplus \ker(P_m(u)).$$

Démonstration. On applique le dernier lemme aux polynômes $Q_1 = P_1$ et $Q_2 = P_2 \cdots P_m$. On obtient alors que

$$\ker(P_1P_2 \cdots P_m(u)) = \ker(P_1(u)) \oplus \ker(P_2 \cdots P_m(u)).$$

Le résultat suit alors par récurrence sur m. □

Chapitre 6

Trigonalisabilité

Dans ce chapitre E désigne un $\mathbb{K}$-espace vectoriel de dimension finie n.

6.1 Endomorphisme trigonalisable

Définition 6.1.1. Un endomorphisme u de E est trigonalisable si il existe une base $\mathcal{B}$ de E telle que $[u]_\mathcal{B}$ est triangulaire supérieure.

Appliqué à une matrice $A \in M_n(\mathbb{K})$, vu comme endomorphisme de $\mathbb{K}^n$, cette définition veut dire qu'il existe P inversible et T triangulaire supérieure telles que $P^{-1}AP = T$.

Soit $A = [u]_\mathcal{B}$ la matrice de u dans la base $\mathcal{B}$. Alors u est trigonalisable si et seulement si A est trigonalisable (exercice).

On peut remplacer le terme "supérieure" par "inférieure" dans la définition. En effet soit u un endomorphisme et $\mathcal{B} = (b_1, \dots, b_n)$ une base tels que $T := [u]_\mathcal{B}$ soit triangulaire supérieure. Alors si on pose $\mathcal{B}' = (b_n, \cdots, b_2, b_1)$ alors $[u]_{\mathcal{B}'}$ sera triangulaire inférieure.

De même supposons qu'on ait des matrices carrées de même taille A, P, T t.q. $P^{-1}AP = T$. Soit alors

$$Q = \begin{pmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{pmatrix}$$

de même taille que A. Alors Q est inversible (exercice) et $Q^{-1}TQ$ est triangulaire inférieure (exercice). Par conséquent $(QP)^{-1}AQP$ est triangulaire inférieure. On peut passer de la même façon d'une matrice triangulaire inférieure à une matrice triangulaire supérieure.

Proposition 22. *Soit u un endomorphisme de E et $\mathcal{B} = (b_1, \dots, b_n)$ une base pour E. Alors : $[u]_\mathcal{B}$ est triangulaire supérieure si et seulement si pour tout $k = 1, \dots, n$, $\mathrm{Vect}\{b_1, \dots, b_k\}$ est stable par u.*

Démonstration. "$\Rightarrow$" Comme $\mathrm{Vect}\{b_1, \dots, b_k\}$ est stable par u on a $u(b_k) \in \mathrm{Vect}\{b_1, \dots, b_k\}$. Donc les derniers $n-k$ coefficients de la colonne $[u(b_k)]_\mathcal{B}$ sont 0. Donc $[u]_\mathcal{B}$ est triangulaire supérieure.

"$\Leftarrow$" Si les derniers $n-k$ coefficients de la colonne $[u(b_k)]_\mathcal{B}$ sont 0 alors $u(b_i) \in \mathrm{Vect}\{b_1, \dots, b_k\}$ pour tout $i \leq k$. Donc $\mathrm{Vect}\{b_1, \dots, b_k\}$ est stable par u. $\square$

Proposition 23. *Soit $A = (a_{ij})$ une matrice triagonale. Les valeurs propres de A sont exactement les coefficients a_{ii} et la multiplicité algébrique de la valeur propre λ est le nombre de i t.q. $\lambda = a_{ii}$. En particulier, le polynôme caractéristique de A est scindé.*

Démonstration. Comme $A - \lambda \mathbf{1}_n$ est aussi triagonale on a $P_A(\lambda) = \prod_{i=1}^n (a_{ii} - \lambda)$. Donc les a_{ii} sont les racines de P_A et la multiplicité de la racine λ correspond au nombre des facteurs avec $a_{ii} = \lambda$. $\square$

6.2 Endomorphisme nilpotent

Définition 6.2.1. Un endomorphisme u de E est dit nilpotent s'il existe $p \in \mathbb{N}$ tel que $u^p = 0$. Le plus petit entier p tel que $u^p = 0$ est appelé l'indice de nilpotence de u.

Proposition 24. *Soit u nilpotent d'indice p. Alors $m_u(X) = X^p$. En particulier, 0 est une valeur propre et c'est la seule.*

Démonstration. Comme $u^p = 0$, $m_u(X)$ divise X^q. Donc $m_u(X) = X^l$ avec $l \leq p$. Or, si $l < p$ alors $u^l = 0$ ce qui contredit la définition de l'indice de nilpotence. Comme les valeurs propres sont les racines de m_u, seulement 0 est valeur propre. □

Proposition 25. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie n. Les assertions suivantes sont équivalentes.*

1. *u est nilpotent.*
2. *Il existe une base $\mathcal{B}$ pour E t.q. $[u]_{\mathcal{B}} = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}$, c.à.d. $[u]_{\mathcal{B}}$ est triangulaire supérieure avec des 0 sur la diagonale.*
3. *$P_u = (-1)^n X^n$.*

Démonstration. (2) $\Rightarrow$ (3) par calcul direct.

Par Cayley Hamilton $P_u(u) = 0$, d'où (3) $\Rightarrow$ (1).

Il reste à montrer l'implication (1) $\Rightarrow$ (2). On raisonne par récurrence sur la dimension n. Si $n = 1$, alors $u = 0$ et il n'y a rien à faire. Regardons le passage de $n-1$ à n. Comme u est nilpotent, u admet 0 comme valeur propre. Soit $b_1 \in \ker(u)$ un vecteur propre. Soit $F = \mathrm{Vect}\{b_1\}$ et G un espace supplémentaire de F dans E, c.à.d. un sous-espace t.q. $E = F \oplus G$ (on ne suppose pas que G est stable par u). Étant donné $x \in E$ il existe alors unique $x_F \in F$ et $x_G \in G$ t.q. $x = x_F + x_G$. En utilisant cette décomposition de x on definie $\pi_G : E = F \oplus G \to F \oplus G$ l'endomorphisme

$$\pi_G(x) = x_G.$$

On a $\pi_G \circ \pi_G = \pi_G$ et $\ker \pi_G = F \subset \ker u$. Soit $w : E \to E$ définie par

$$w(x) = \pi_G(u(x)).$$

Montrons que pour tout $k \geq 1$, on a $\pi_G \circ (u^k - w^k) \circ \pi_G = 0_{\mathrm{End}(E)}$. On a

$$\pi_G \circ w^k \circ \pi_G = \pi_G \circ \pi_G \circ u \circ w^{k-1} \circ \pi_G = \pi_G \circ u \circ w^{k-1} \circ \pi_G$$

d'où

$$\pi_G \circ (u^k - w^k) \circ \pi_G = \pi_G \circ u^k \circ \pi_G - \pi_G \circ w^k \circ \pi_G = \pi_G \circ u \circ (u^{k-1} - w^{k-1}) \circ \pi_G.$$

De plus $\pi_G \circ (u^{k-1} - w^{k-1}) \circ \pi_G = 0_{\mathrm{End}(E)}$ implique $(u^{k-1} - w^{k-1}) \circ \pi_G(x) \in \ker \pi_G \subset \ker u$ pour tout $x \in E$. D'où $\pi_G \circ (u^{k-1} - w^{k-1}) \circ \pi_G = 0_{\mathrm{End}(E)}$ implique $\pi_G \circ (u^k - w^k) \circ \pi_G = 0_{\mathrm{End}(E)}$. Donc le résultat suit par récurrence car $\pi_G \circ (u^{k-1} - w^{k-1}) \circ \pi_G = 0_{\mathrm{End}(E)}$ si $k = 1$.

Soit maintenant p l'indice de nilpotence de u alors $u^p = 0_{\mathrm{End}(E)}$ et donc $\pi_G \circ w^p \circ \pi_G = \pi_G \circ w^p \circ \pi_G - \pi_G \circ u^p \circ \pi_G = 0_{\mathrm{End}(E)}$. Or, $\pi_G \circ w^p \circ \pi_G = 0_{\mathrm{End}(E)}$ implique que la reduction to w sur G, w_G, satisfair $w_G^p = 0_{\mathrm{End}(G)}$. Ainsi w_G est nilpotent. Par hypothèse de récurrence et comme $\dim G = n-1$ il existe une base $\mathcal{B}' = (b_2, \ldots, b_n)$ de G telle que $[w_G]_{\mathcal{B}'}$ soit triangulaire avec 0 sur la diagonale. Calculons la matrice de u dans la base $\mathcal{B} = (b_1, \ldots, b_n)$. Pour tout $b_i \in \mathcal{B}$ on peut décomposer $u(b_i)$

$$u(b_i) = u(b_i)_F + u(b_i)_G, \quad \text{avec } u(b_i)_F \in F,\ u(b_i)_G \in G$$

De plus, $u(b_i)_G = \pi_G(u(b_i)_G) = w(b_i)$ et $u(b_i)_F = a_i b_1$ pour un $a_i \in \mathbb{K}$. Comme $u(b_1) = 0$ la première colonne de $[u]_\mathcal{B}$ ne contient que des 0 et donc

$$[u]_\mathcal{B} = \begin{pmatrix} 0 & * \\ 0 & [w_G]_{\mathcal{B}'} \end{pmatrix}$$

où $*$ représente la ligne $a_2 \cdots a_n$. Comme $[w_G]_{\mathcal{B}'}$ est triangulaire avec 0 sur la diagonale aussi la matrice de u est triangulaire avec 0 sur la diagonale. □

Corollaire 15. *Soit u un endomorphisme de E et λ une valeur propre de u. On suppose que l'espace caractéristique F_λ de λ est de dimension finie. La réduction u_{F_λ} de u sur l'espace caractéristique F_λ est trigonalisable. En particulier, il existe une base $\mathcal{B}$ de F_λ t.q.*

$$[u_{F_\lambda}]_\mathcal{B} = \begin{pmatrix} \lambda & & * \\ \vdots & \ddots & \\ 0 & \cdots & \lambda \end{pmatrix}.$$

Démonstration. On rappelle que l'espace caractéristique de λ est donné par $F_\lambda = \ker(u - \lambda \mathrm{id}_E)^{q_\lambda}$ (Déf. 5.1.2). La reduction de $u - \lambda \mathrm{id}_E$ sur F_λ, qui est $u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$, est donc nilpotent. D'après Prop. 25 $u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$ est triagonal avec 0 sur la diagonale, donc u_{F_λ} est triagonal avec λ sur la diagonale. □

6.3 Trigonalisation

On rappelle la définition de q_λ (Définition 5.1.2), notament q_λ est le plus petit entier naturel t.q. $\ker(u - \lambda \mathrm{id})^{q_\lambda} = \ker(u - \lambda \mathrm{id})^{q_\lambda + 1}$. $F_\lambda = \ker(u - \lambda \mathrm{id})^{q_\lambda}$ est l'espace caractéristique de la valeur propre λ.

Proposition 26. *Soit u un endomorphisme de E, λ une valeur propre de u et F_λ l'espace caractéristique associé. La multiplicité de λ en tant que racine du polynôme minimal m_u de u est q_λ, c.à.d.*

$$m_u(X) = (X - \lambda)^{q_\lambda} Q(X)$$

où Q est un polynôme t.q. $Q(\lambda) \neq 0$.

Démonstration. Par Lemme 10 $\ker(u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda})^l \subset F_\lambda$ et que l'inclusion est une égalité si et seulement si $l \geq q_\lambda$. Donc $u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$ est nilpotent d'indice q_λ. D'où $(X - \lambda)^{q_\lambda}$ est le polynôme minimal m_{u_F} de u_F. Comme m_{u_F} divise m_u on a $m_u(X) = (X - \lambda)^l Q(X)$ avec $l \geq q_\lambda$ and Q un polynôme qui est premier avec $(X - \lambda)$. D'après le lemme des noyaux, $E = \ker(u - \lambda)^l \oplus \ker Q(u)$. Par Lemme 10 on a $\ker(u - \lambda)^l = \ker(u - \lambda)^{q_\lambda}$. Donc $E = \ker(u - \lambda)^{q_\lambda} \oplus \ker Q(u)$, et une nouvelle application du lemme des noyaux implique que $P(X) = (X - \lambda)^{q_\lambda} Q(X)$ est annulateur de u. Comme m_u est le plus petit annulateur, m_u divise P. Donc $l \leq q_\lambda$. Il s'ensuit que $l = q_\lambda$. □

Théorème 17. *Soit u un endomorphisme de E (de dimension finie). Les assertions suivantes sont équivalentes.*

1. *u est trigonalisable.*
2. *Le polynôme caractéristique est scindé.*

3. *u est annulé par un polynôme scindé dans $\mathbb{K}[X]$.*
4. *Le polynôme minimal m_u est scindé.*
5. *E est la somme directe des espaces caractéristiques de u, $E = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_k}$.*

Démonstration. (1) $\Rightarrow$ (2) : Sous l'hypothèse (1) P_u est scindé par calcul direct.

(2) $\Rightarrow$ (3) : Par le théorème de Cayley-Hamilton le polynôme caractéristique est annulle u.

(3) $\Rightarrow$ (4) : Si on note Q le polynôme donné par l'hypothèse (3). Alors m_u divise Q. Tout diviseur d'un polynôme scindé est scindé.

(4) $\Rightarrow$ (5) : Par hypothèse $m_u = (X - \lambda_1)^{l_1} \cdots (X - \lambda_k)^{l_k}$ où $\lambda_i \in \mathbb{K}$ sont les racines de m_u et $l_i \in \mathbb{N}$ leur multiplicité. D'après Thm. 16 les λ_i sont les valeurs propres de u et par Prop. 26 $l_i = q_i$. Par le lemme des noyaux

$$E = \ker m_u(u) = \ker((u - \lambda_1 \mathrm{id})^{q_1}) \oplus \cdots \oplus \ker((u - \lambda_k \mathrm{id})^{q_k}) = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_k}.$$

(5) $\Rightarrow$ (1) : F_{λ_i} est l'espace caractérisistique de λ_i. Par Cor. 15 il existe une base $\mathcal{B}_i$ pour F_{λ_i} tel que $[u_{F_{\lambda_i}}]_{\mathcal{B}_i}$ est triangulaire. Soit $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$. Comme $E = \oplus_{i=1}^p F_{\lambda_i}$ la matrice $[u]_{\mathcal{B}}$ est diagonal. Bloc i est la matrice $[u_{F_{\lambda_i}}]_{\mathcal{B}_i}$, qui est triangulaire. Du coup $[u]_{\mathcal{B}}$ est triangulaire. □

Corollaire 16. *Soit A une matrice trigonalisable avec valeurs propres $\lambda_1, \cdots, \lambda_k$. Il existe une matrice P inversible t.q. $T = P^{-1}AP$ est triangulaire supérieure par blocs*

$$T = \begin{pmatrix} T_{\lambda_1} & 0 & \cdots & 0 \\ 0 & T_{\lambda_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & T_{\lambda_k} \end{pmatrix}$$

et T_λ est une matrice triangulaire de taille $\dim F_\lambda \times \dim F_\lambda$ qui à valeur λ sur la diagonale

$$T_\lambda = \lambda \mathbf{1}_{q_\lambda} + N_\lambda, \quad N_\lambda = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

En particulier $N_\lambda^{q_\lambda} = 0$.

Démonstration. La décomposition en blocs suit du dernier théorème. On applique Cor. 15 aux blocs associés aux valeurs propres de A. □

Corollaire 17. *Soit u un endomorphisme d'un espace vectoriel complexe de dimension finie, alors u est trigonalisable.*

Démonstration. Comme $\mathbb{K} = \mathbb{C}$ tout polynôme est scindé. En particulaire le polynôme caractéristique est scindé. Le resultat suit alors du dernier théorème. □

Théorème 18 (troisième critère de la diagonalisation)**.** *Soit u un endomorphisme de E (de dimension finie). u est diagonalisable si est seulement si son polynôme minimal est scindé et à racines simples.*

Démonstration. On rappelle que u diagonalisable si est seulement si son polynôme caractéristique est scindé et E est la somme directe de ses espaces propres E_λ. De plus, $E_\lambda \subset F_\lambda$.

Soit alors u diagonalisable. Alors u est trigonalisable. Par Thm. 17, E est la somme directe des espaces caractéristiques. Comme E est aussi la somme directe de ses espaces propres et

$E_\lambda \subset F_\lambda$ on doit avoir $E_\lambda = F_\lambda$ pour toute valeur propre. Donc $q_\lambda = 1$ pour toute valeur propre. Par Prop. 26 $q_\lambda = 1$ veut dire que λ est une racine simple de m_u.

Soit maintenant m_u scindé à racine simple. Donc

$$m_u(X) = \prod_{i=1}^{k}(X - \lambda_i)$$

ou $\lambda_1, \cdots, \lambda_k$ sont les racines (distinctes) de u. Par Prop. 26 $q_\lambda = 1$ et donc $E_\lambda = F_\lambda$ pour toute valeur propre λ. Par le lemme des noyaux,

$$E = \ker m_u(u) = \ker(u - \lambda_1 \mathrm{id}) \oplus \cdots \oplus \ker(u - \lambda_k \mathrm{id}) = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_k}.$$

Maintenant Thm. 17 implique que le polynôme caractéristique est scindé. Donc u est diagonalisable. □

Exemple 6.3.1. *Tout projecteur u est diagonalisable, car le polynôme $X^2 - X$ est scindé est à racines simples.*

Corollaire 18. *Soit u un endomorphisme de E diagonalisable ($\dim E < +\infty$). Soit F un sous espace stable pour u. Alors la reduction u_F de u à F est diagonalisable.*

Démonstration. On utilise le troisième critère de la diagonalisation. L'hypothèse implique alors que le polynôme minimal m_u est scindé est à racine simple. On a $m_u(u_F)(x) = m_u(u)(x) = 0$ pour tout $x \in F$. Donc m_u est un annullateur de u_F. Il en suit que le polynôme minimal de u_F divise m_u. Ceci implique que le polynôme minimal de u_F est scindé est à racine simple. □

6.4 Comment trouver la forme triangulaire d'un endomorphisme ?

Soit u un endomorphisme de E qui est trigonalisable. On se pose la question, comment trouver une base $\mathcal{B}$ t.q. $[u]_\mathcal{B}$ est triangulaire ?

Appliquer à une matrice A de taille $n \times n$ (un endomorphisme de $\mathbb{K}^n$) la question équivalente est, comment trouver des matrices P et T, t.q.

$$T = P^{-1}AP$$

est triangulaire ? P est alors la matrice de passage de la base canonique de $\mathbb{K}^n$ vers une base dans laquelle A est triangulaire.

La première étappe de la construction d'une forme triangulaire passe par la détermination des valeurs propres de u (ou A). Celle-ci sont les racines d'un polynôme P non-nul qui divise le polynôme caractéristique, P peut être le polynôme caractéristique ou le polynôme minimal, mais ce n'est pas necessaire.

Une fois trouvé les valeurs propres il faut, pour toute valeur propre λ, déterminer une base pour son espace caractéristique F_λ t.q. u_{F_λ} est triangulaire. Ceci peut se faire ainsi. On pose

$$h_\lambda = u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$$

alors h_λ est nilpotent. (L'indice de nilpotence de h_λ est q_λ, mais ce n'est pas necessaire de connaitre q_λ en avant.) Résoudre

$$h_\lambda(b_1) = 0_E, \qquad b_1 \neq 0_E$$

et puis par récurrence

$$h_\lambda(b_{j+1}) \in \mathrm{Vect}\{b_1, \cdots, b_j\}, \qquad b_{j+1} \notin \mathrm{Vect}\{b_1, \cdots, b_j\} \tag{6.1}$$

jusqu'il n'y a plus de solution. En effet, par construction $\{b_1, \cdots, b_j\}$ est une famille libres et $h_\lambda^j(b_j) = 0_E$. Donc ainsi on construit une base pour $\ker h_\lambda^{q_\lambda} = F_\lambda$. La matrice de h_λ dans la base $\mathcal{B}_\lambda = \{b_1, \cdots, b_k\}$ est alors

$$[h_\lambda]_{\mathcal{B}_\lambda} = (C_1 \cdots C_k)\,, \qquad C_i = \begin{pmatrix} c_{1i} \\ \vdots \\ c_{i-1\,i} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

C_1 est une colonne de 0 et les c_{ji} sont donnés par la solution de l'équation

$$h_\lambda(b_i) = c_{1i}b_1 + \cdots + c_{i-1\,i}b_{i-1}.$$

u_{F_λ} est alors triangulaire dans la base $\mathcal{B}_\lambda$, notament

$$[u_{F_\lambda}]_{\mathcal{B}_\lambda} = \begin{pmatrix} \lambda & & c_{ij} \\ \vdots & \ddots & \\ 0 & \cdots & \lambda \end{pmatrix}$$

c.à.d. sur la diagonale on trouve la valeur propre et en haut à droite les coefficients de la matrice $[h_\lambda]_{\mathcal{B}_\lambda}$.

Une fois les matrices $[u_{F_\lambda}]_{\mathcal{B}_\lambda}$ trouvé pour toutes valeurs propres, la matrice de u dans la base $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_k}$ est la matrice diagonale en blocs où les blocs correspondent aux matrices $[u_{F_{\lambda_i}}]_{\mathcal{B}_{\lambda_i}}$.

Exemple 6.4.1. Voici un exemple avec un seul bloc. Soit

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

1. Calcul des valeurs propres : Le polynôme caractéristique est

   $$P_A(\lambda) = -\lambda^3$$

   donc la seule valeur propre de A est $\lambda = 0$. Elle a multiplicité algébrique 3. L'endormorphisme h_λ correspond ici alors à A.
2. On résoud $Ab_1 = 0$ avec $b_1 \neq 0$ donc

   $$\begin{pmatrix} 0 & 0 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

   Ceci est équivalent à $z = 0$ et $x - y = 0$. Donc $b_1 = (1, 1, 0)$ est une solution. (D'ailleurs la multiplicité géométrique de 0 est égale à 1.)
3. On résoud $Ab_2 \in \text{Vect}\{b_1\}$ avec $b_2 \notin \text{Vect}\{b_1\}$. Autrement dit on résoud $Ab_2 = cb_1$ où $c \in \mathbb{C}$. On peut ajuster c comme on veut, mais tel que b_2 n'est pas un multiple de b_1. Ceci donne

   $$\begin{pmatrix} 0 & 0 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \\ c \\ 0 \end{pmatrix}$$

   Ceci est équivalent à $z = c$ et $x - y + z = 0$. Donc $b_2 = (0, 1, 1)$ avec $c = 1$ est une solution qui n'appartient pas à $\text{Vect}\{b_1\}$.

4. On résoud $Ab_3 \in \mathrm{Vect}\{b_1, b_2\}$ avec $b_3 \notin \mathrm{Vect}\{b_1, b_2\}$. Autrement dit on résoud $Ab_3 = c'b_1 + db_2$ où $c', d \in \mathbb{C}$. On peut ajuster c', d comme on veut, mais tel que b_3 n'est pas combinaison linéaire de b_1 et b_2. Ceci donne
$$\begin{pmatrix} 0 & 0 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \\ c+d \\ d \end{pmatrix}$$
Ceci est équivalent à $z = c$ et $x - y + z = \frac{d}{2}$. Donc $b_3 = (0, 0, 1)$ avec $c = 1$ et $d = 2$ est une solution qui n'appartient pas à $\mathrm{Vect}\{b_1, b_2\}$.

En résumé on a trouvé une base $\mathcal{B} = \{b_1, b_2, b_3\}$. Donc la matrice de passage est $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ et $T = P^{-1}AP$. On n'a pas besoin de calculer ce produit de matrices[1] pour trouver T, car on sait que T a colonnes $C_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $C_2 = \begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix}$ et $C_3 = \begin{pmatrix} c' \\ d \\ 0 \end{pmatrix}$. Donc
$$T = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

6.5 Puissances des matrices trigonalisables

Lemme 17. *Soit u et v deux endomorphisme de E qui commutent. Alors*
$$(u+v)^n = \sum_{k=0}^{n} c_n^k u^{n-k} \circ v^k, \quad c_n^k = \frac{n!}{(n-k)!k!}$$

Démonstration. C'est la même formule que pour les nombres. La démonstration (par exemple, par récursion) se fait de la même manière ; on utilise qu'on peut commuter tous les u vers la gauche. $\square$

Soit A une matrice carrée trigonalisable. Par Cor. 16 qu'il existe une matrice de passage P et une matrice T diagonale en blocs, les blocs étant triangulaires, t.q.
$$T = P^{-1}AP.$$

Plus precisement
$$T = \begin{pmatrix} T_{\lambda_1} & 0 & \cdots & 0 \\ 0 & T_{\lambda_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & T_{\lambda_k} \end{pmatrix} \tag{6.2}$$
où les λ_i sont les valeurs propres de A et T_λ est une matrice triangulaire de taille $\dim F_\lambda \times \dim F_\lambda$ qui à valeur λ sur la diagonale
$$T_\lambda = \lambda \mathbf{1}_{q_\lambda} + N_\lambda, \quad N_\lambda = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & 0 & 0 \end{pmatrix} \tag{6.3}$$

1. On aura besoin de P^{-1} plutard, quand on veut calculer les puissances de A.

De plus, $N_\lambda^{q_\lambda} = 0$.

Nous calculons, avec Lemme 17,

$$T_\lambda^n = \sum_{j=0}^{n} c_n^j \lambda^{n-j} \mathbf{1}_{q_\lambda} N_\lambda^j = \sum_{j=0}^{\min\{n, q_\lambda - 1\}} \frac{n!}{(n-j)!j!} \lambda^{n-j} N_\lambda^j \tag{6.4}$$

ce qui est à rentrer dans l'expression pour la puissance de A :

$$A^n = PT^nP^{-1} = P \begin{pmatrix} T_{\lambda_1}^n & 0 & \cdots & 0 \\ 0 & T_{\lambda_2}^n & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & T_{\lambda_k}^n \end{pmatrix} P^{-1}$$

Exemple 6.5.1. On veut déterminer les puissances de

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 45 & -39 & 11 \end{pmatrix}$$

On commence par la trigonalisation de A. Le polynôme caractéristique est

$$P_A(\lambda) = (5-\lambda)(\lambda-3)^2$$

Donc les valuers propres sont $\lambda_1 = 5$ avec multiplicité algébrique $m_5 = 1$, et $\lambda_2 = 3$ avec multiplicité algébrique $m_3 = 2$.

Déterminons les espaces propres.

$$E_5 = \ker(A - 5\,\mathbf{1}_3) = \ker \begin{pmatrix} -5 & 1 & 0 \\ 0 & -5 & 1 \\ 45 & -39 & 6 \end{pmatrix} = \mathrm{Vect}\{b_1\}, \quad b_1 = \begin{pmatrix} 1 \\ 5 \\ 25 \end{pmatrix}$$

$$E_3 = \ker(A - 3\,\mathbf{1}_3) = \ker \begin{pmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 45 & -39 & 8 \end{pmatrix} = \mathrm{Vect}\{b_2\}, \quad b_2 = \begin{pmatrix} 1 \\ 3 \\ 9 \end{pmatrix}.$$

En particulier la multiplicité géométrique de λ_2 et $g_3 = 1$ montrant que A n'est pas diagonalisable.

Comme $F_5 = E_5$ l'espace caractéristique F_{λ_2} est de dimension 2. On suit l'algorithme de triagonalisation décrit au dessus, ce que veut dire trouver un troisième vecteur b_3 t.q. $\ker(A - 3\,\mathbf{1}_3)(b_3) \in \mathrm{Vect}\{b_2\}$ mais avec $b_3 \notin \mathrm{Vect}\{b_2\}$. On doit alors résoudre

$$\begin{pmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 45 & -39 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \\ 3c \\ 9c \end{pmatrix}$$

ce qui est équivalent à

$$-3x + y = c \quad \text{et} \quad -3y + z = 3c$$

(la derniére équation est une combinaison linéaire des deux autres). Toute solution avec $c \neq 0$ est éligible et on choisit $x = 0$, $y = 1$, $z = 6$, $c = 1$. Donc $b_3 = (0, 1, 6)$.

La matrice de passage est

$$P = \begin{pmatrix} b_1 b_2 b_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 5 & 3 & 1 \\ 25 & 9 & 6 \end{pmatrix} \tag{6.5}$$

et les vecteurs de la nouvelle base (b_1, b_2, b_3) satisfont

$$Ab_1 = 5b_1, \quad Ab_2 = 3b_2, \quad Ab_3 = 3b_3 + b_2.$$

D'où

$$T = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \tag{6.6}$$

Comme $T_5 = (5)$ on a $T_3^n = (5^n)$. De plus

$$T_3^n = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}^n = \begin{pmatrix} 3^n & n3^{n-1} \\ 0 & 3^n \end{pmatrix}.$$

D'où

$$T^n = \begin{pmatrix} 5^n & 0 & 0 \\ 0 & 3^n & n3^{n-1} \\ 0 & 0 & 3^n \end{pmatrix} \tag{6.7}$$

Pour obtenir A^n il faut inverser P et calculer le produit PT^nP^{-1}. On trouve

$$P^{-1} = \frac{1}{4} \begin{pmatrix} 9 & -6 & 1 \\ -5 & 6 & -1 \\ -30 & 16 & -2 \end{pmatrix} \tag{6.8}$$

et

$$A^n = \frac{1}{4} \begin{pmatrix} 9\times 5^n - 5\times 3^n - 10n\times 3^{n+1} & 2(-3\times 5^n + 3^{n+1} + 8n\times 3^n) & 5^n - 3^n - 2n\times 3^n \\ 45(5^n - 3^n - 2n\times 3^n) & 2(-3\times 5^{n+1} + 17\times 3^n + 8n\times 3^{n+1}) & 5^{n+1} - 5\times 3^n - 2n\times 3^{n+1} \\ 45(5^{n+1} - 5\times 3^n - 2n\times 3^{n+1}) & 6(-5^{n+2} + 25\times 3^n + 8n\times 3^{n+1}) & 5^{n+2} - 7\times 3^{n+1} - 2n\times 3^{n+2} \end{pmatrix}$$

Exemple 6.5.2. On considère la suite $(u_n)_{n\geq 0}$ définie par $u_0 = 1$, $u_1 = 1$, $u_2 = 1$ et la récurrence

$$\forall n \geq 0, \quad u_{n+3} = 45u_n - 39u_{n+1} + 11u_{n+2}.$$

Quelle est la solution ? Il faut alors calculer u_n pour tout n.

Comme dans Chapitre 4 on commenc par une écriture de la relation de récurrence sous la forme d'une équation matricielle. La relation de récurrence est équivalente à

$$\begin{pmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 45 & -39 & 11 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$$

Avec $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 45 & -39 & 11 \end{pmatrix}$ on obtient donc l'équation matricielle

$$X_{n+1} = AX_n.$$

On s'intéresse à u_n qui est le premier coefficient de X_n,

$$u_n = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$$

Les conditions initiales $u_0 = 1$, $u_1 = 1$, $u_2 = 1$ correspondent à

$$X_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

d'où

$$u_n = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} A^n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

On pourrait utiliser le resultat pour A^n trouvé en haut. Mais ce n'est pas necessaire de faire le calcul (un peu compliqué) du produit PT^nP^{-1}, c'est plus rapide de calculer, avec (6.5)

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

puis, avec (6.8) [2]

$$P^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}$$

puis, avec (6.7)

$$u_n = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} T^n \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} = 5^n - 4n \times 3^n.$$

2. On n'a même pas besoin de calculer l'inverse de P entièrement, il suffit de résoudre l'équation $PY_0 = X_0$ pour les conditions initiales X_0 données, alors $Y_0 = P^{-1}X_0$.

Chapitre 7

Projecteurs spectraux

Dans ce chapitre on va retrouver le résultat du Cor. 16, mais d'une manière plus abstraite. Ceci nous donnerait aussi une approche alternative à trigonaliser des endomorphisme.

7.1 Notion de projecteur spectral

7.1.1 Projecteur

Définition 7.1.1. Un projecteur de E est un endormorphisme $\Pi : E \to E$ qui satisfait $\Pi \circ \Pi = \Pi$.

Si Π est un projecteur de E alors $E = \text{im } \Pi \oplus \ker \Pi$. En effet,

$$x = \Pi(x) + (\text{id} - \Pi)(x)$$

est une décomposition de x en une somme $y + z$ avec $y \in \text{im } \Pi$ et $z \in \ker \Pi$, et cette décomposition est unique : $x \in \ker \Pi \cap \text{im } \Pi$ implique qu'il existe $y \in E$ t.q. $x = \Pi(y)$ et $\Pi(\Pi(y)) = 0$; mais comme $\Pi(\Pi(y)) = \Pi(y)$ on a $x = 0$.

D'une manière réciproque, si $E = F \oplus G$ alors il existe un unique projecteur Π t.q. $F = \text{im } \Pi$ et $G = \ker \Pi$. En effet, chaque vecteur x de E se décompose donc d'une manière unique en une somme

$$x = y + z$$

où $y \in F$ et $z \in G$. Π est alors donné par $\Pi(x) = y$.

On a deja vu que $X^2 - X$ est un polynôme annulateur de chaque projecteur. Un projecteur d'un espace vectorielle de dimension finie est donc diagonalisable.

Proposition 27. *Soient u, v deux endomorphismes diagonalisables d'un espace vectoriel de dimension finie. On suppose que u et v commutent, $u \circ v = v \circ u$. Il existe une base $\mathcal{B}$ qui consistent en des vecteurs propres pour les deux, u et v. Dans cette base $[u]_{\mathcal{B}}$ et $[v]_{\mathcal{B}}$ sont diagonals. En particulier, $u - v$ est diagonalisable.*

Démonstration. u commute aussi avec $v - \lambda\text{id}$. Soit λ une valeur propre de v. D'après Prop. 13 u preserve l'espace propre $E_\lambda(v)$. D'après Cor. 18 $u_{E_\lambda(v)}$ est diagonalisable. Il existe donc une base $\mathcal{B}_\lambda$ de $E_\lambda(v)$ t.q. $[u]_{\mathcal{B}_\lambda}$ est diagonal. Biensûre, $[v]_{\mathcal{B}_\lambda}$ est diagonal. Soit $\mathcal{B}$ l'union des bases $\mathcal{B}_\lambda$ sur les valeurs propres de v. Alors $[u]_{\mathcal{B}}$ et $[v]_{\mathcal{B}}$ sont diagonals. De plus $[u - v]_{\mathcal{B}}$ est diagonal. □

7.1.2 Projecteur spectral de u

Nous considérons un endomorphisme u d'un espace vectoriel E de dimension finie. On suppose que $\lambda \in \mathbb{K}$ est une valeur propre de u. Ceci est le cas si et seulement si λ est une racine du polynôme minimal m_u de u. Ce dernier doit donc être de la forme

$$m_u(X) = (X - \lambda\mathrm{id})^{q_\lambda} Q(X)$$

où Q est un polynôme t.q. $Q(\lambda) \neq 0$ ou, autrement dit, Q et $(X - \lambda\mathrm{id})^{q_\lambda}$ sont premiers entre eux. On rappelle que $F_\lambda = \ker(X - \lambda\mathrm{id})^{q_\lambda}$ est l'espace caractéristique de λ. D'après le lemme des noyaux on a

$$E = F_\lambda \oplus \ker Q(u).$$

Chaque vecteur x de E se décompose d'une manière unique en une somme

$$x = y + z \tag{7.1}$$

avec $y \in F_\lambda$ et $z \in \ker Q(u)$.

Définition 7.1.2. Le projecteur spectral de u associé à λ est l'endomorphisme $\Pi_\lambda : E \to E$ donné par $\Pi_\lambda(x) = y$ où y est déterminé par (7.1).

Lemme 18. *Les propriétées fondamentales du projecteur spectral sont*

1. $\Pi_\lambda \circ \Pi_\lambda = \Pi_\lambda$,
2. $\mathrm{im}\Pi_\lambda = F_\lambda$,
3. $\ker \Pi_\lambda = \ker Q(u)$.
4. $u \circ \Pi_\lambda = \Pi_\lambda \circ u$.

Démonstration. Soit $x = y + z$ dans la décomposition (7.1). Alors $\Pi_\lambda(x) = x$ si et seulement si $z = 0$. De plus $\Pi_\lambda(x) = 0$ si et seulement si $y = 0$.

Comme F_λ est stable pour u on a, pour tout $y \in F$, $\Pi_\lambda(u(y)) = u(y) = u(\Pi_\lambda(y))$. De plus, $\ker Q(u)$ est stable pour u donc, pour tout $z \in \ker Q(u)$, $\Pi_\lambda(u(z)) = 0$ et $u(\Pi_\lambda(z)) = u(0) = 0$. □

Lemme 19. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie et λ, μ deux valeurs propres distinctes. Alors*

$$\Pi_\mu \circ \Pi_\lambda = \Pi_\lambda \circ \Pi_\mu = 0_{\mathrm{End}(E)}.$$

Démonstration. L'hypothèse implique que $m_u(X) = (X - \lambda\mathrm{id})^{q_\lambda}(X - \mu\mathrm{id})^{q_\mu} Q(X)$ et $Q(X)$ et $(X - \lambda\mathrm{id})^{q_\lambda}$ et $(X - \mu\mathrm{id})^{q_\mu}$ sont deux à deux premiers entre eux. Le lemme des noyaux nous donne alors que

$$E = F_\lambda \oplus F_\mu \oplus \ker Q(u).$$

Chaque vecteur x de E se décompose donc d'une manière unique dans une somme

$$x = y + z + w$$

où $y \in F_\lambda$, $z \in F_\mu$ et $w \in \ker Q(u)$. Nous avons $\Pi_\lambda(x) = y$ et $\Pi_\mu(y) = 0$. De plus $\Pi_\mu(x) = z$ et $\Pi_\lambda(z) = 0$. D'où le resultat. □

Théorème 19. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie. Toute projecteur spectral de u est un polynôme de u.*

Démonstration. Soit λ une valuer propre de u. Alors $m_u(X) = (X - \lambda)^{q_\lambda} Q(X)$ ou Q est un polynôme qui est premier avec $(X - \lambda)^{q_\lambda}$. On applique le Théorème de Bézout à $P_1(X) = (X - \lambda)^{q_\lambda}$ et $P_2 = Q$. Il existe alors deux polynômes Q_1, Q_2 t.q.

$$Q_2(X)Q(X) = 1 - Q_1(X)(X - \lambda)^{q_\lambda}. \tag{7.2}$$

Montrons que $\Pi_\lambda = Q_2(u) \circ Q(u)$. Par le lemme des noyaux $E = F_\lambda \oplus \ker Q(u)$. Soit $x = y + z$ avec $y \in F_\lambda$ et $z \in \ker Q(u)$. Alors

$$Q_2(u) \circ Q(u)(x) = Q_2(u) \circ Q(u)(y) = y - Q_1(u)((u - \lambda)^{q_\lambda}(y)) = y - Q_1(u)(0) = y.$$

□

7.2 Décomposition spectrale de Dunford

Théorème 20 (Décomposition spectrale de Dunford)**.** *Soit u un endomorphisme d'un espace vectoriel E de dimension finie. On suppose que le polynôme minimal de u est scindé (d'une manière équivalente, u est trigonalisable). Alors il existe un unique couple (δ, η) d'endomorphismes de E t.q. δ est diagonalisable, η est nilpotent, $\delta \circ \eta = \eta \circ \delta$ et*

$$u = \delta + \eta.$$

L'indice de nilpotence de η est $\max_i q_{\lambda_i}$ où q_{λ_i} est la multiplicité de λ_i en tant que racine du polynôme minimal de u.

Démonstration. Soient $\lambda_1, \cdots, \lambda_k$ les valeurs propres distinctes de u. Par hypothèse m_u est scindé et ses racines sont les λ_i. Donc

$$m_u(X) = \prod_{i=1}^{k} (X - \lambda_i)^{q_{\lambda_i}}.$$

Posons

$$\delta = \sum_{i=1}^{k} \lambda_i \Pi_{\lambda_i} \quad \text{et} \quad \eta = u - \delta.$$

Soit $P(X) = \prod_{i=1}^{k}(X - \lambda_i)$ et $x \in F_{\lambda_j}$. On pose $Q_j(X) = \prod_{i \neq j}(X - \lambda_i)$. Alors

$$P(\delta)(x) = Q_j(\delta) \circ (\delta - \lambda_j \mathrm{id})(x) = Q_j(\delta)(0) = 0.$$

D'où P est un polynôme annulateur de δ qui est scindé est à racine simple. Il en suit que le polynôme minimal de δ est scindé est à racine simple. Donc δ est diagonalisable.

D'après Lemme 19 u commute avec chaque Π_{λ_i} donc aussi avec δ. Il en suit que $\delta \circ \eta - \eta \circ \delta = (\delta + \eta) \circ \delta - \delta \circ (\delta + \eta) = u \circ \delta - \delta \circ u = 0$.

Montrons que η est nilpotent. Comme la restriction de Π_λ à F_λ est l'identité on a $\eta_{F_\lambda} = u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$. On a vu que $u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$ est nilpotent d'indice q_λ. D'où η est nilpotent d'indice $p = \max_i q_{\lambda_i}$.

Montrons l'unicité. Si (δ', η') est un autre couple avec les proprietés au dessus. En particulier, δ' et η' commutent entre eux et donc aussi avec $\delta' + \eta' = u$. Comme δ est un polynôme de u il commute avec δ'. Il en suit que aussi η et η' commutent. On peut donc commuter dans developement de l'expression $(\eta - \eta')^{2n}$ tout les η vers la gauche pour obténir

$$(\eta - \eta')^{2n} = \sum_{i=1}^{2n} (-1)^i c_{2n}^i \eta^{2n-i} \circ \eta'^i.$$

Si $i \leq n$ alors $\eta^{2n-i} = 0$ comme η est nilpotent d'indice $\leq n$, et si $i \geq n$ alors $\eta'^i = 0$ comme η' est nilpotent d'indice $\leq n$. D'où $(\eta - \eta')^{2n} = 0$ et $\eta - \eta'$ est nilpotent. De plus, d'après Prop. 27 $\delta - \delta' = \eta' - \eta$ est diagonalisable. Mais un endomorphisme nilpotent diagonalisable est 0. Donc $\eta = \eta'$ et $\delta = \delta'$. □

Ce resultat est à comparer avec Cor. 16. Si nous choisissons une base $\mathcal{B}$ qui est l'union de bases $\mathcal{B}_{\lambda_i}$ pour les espaces caractéristiques alors $[\delta]_{\mathcal{B}}$ est diagonal. En effet

$$[\delta_{F_\lambda}]_{\mathcal{B}_\lambda} = \lambda 1_{q_\lambda}$$

Cor. 16 nous dit qu'il existe une base $\mathcal{B}_\lambda$ dans laquelle $[\eta_{F_\lambda}]_{\mathcal{B}_\lambda}$ est triangulaire avec 0 sur la diagonale, donc en particulier nilpotent. Donc Cor. 16 est plus spécifique de la décomposition spectrale de Dunford. Autrement dit la décomposition spectrale de Dunford est Cor. 16 "sans choix de base".

7.3 Comment déterminer les projecteurs spectraux d'un endomorphisme ?

On a vu que

$$\Pi_\lambda = Q_2(u) \circ Q(u) \tag{7.3}$$

ou Q est déterminé par $m_u(X) = (X - \lambda)^{q_\lambda} Q(X)$, donc

$$Q(X) = \prod_{\lambda' \neq \lambda} (X - \lambda')^{q_{\lambda'}}$$

(produit sur toutes les valeurs propres distinctes de λ) et Q_2 est solution de l'équation (7.2). Avec un peu de chance, cet équation peut se résoudre directement avec l'ansatz $Q_2(X) = \sum_{i=0}^{q_\lambda - 1} a_i X^i$ et Q_1 un polynôme de degré un de moins que le degré de Q.

Le plus souvent on veut déterminer tous les projecteurs spectraux d'un endomorphisme trigonalisable. Soient $\lambda_1, \cdots, \lambda_k$ les valeurs propres de u. Pour trouver le polynôme Q_2 de (7.3) pour chaque valeurs propres on décompose $\frac{1}{m_u}$ en éléments simples, c.à.d. on détermine des polynômes $A_i(X)$ de degré au plus $q_{\lambda_i} - 1$ t.q.

$$\frac{1}{m_u(X)} = \frac{A_1(X)}{(X - \lambda_1)^{q_{\lambda_1}}} + \cdots + \frac{A_k(X)}{(X - \lambda_k)^{q_{\lambda_k}}}. \tag{7.4}$$

Si on combine (7.2) et (7.4) on obtient, pour la valeur propre $\lambda = \lambda_i$

$$Q_1(X)(X - \lambda_i)^{q_{\lambda_i}} + Q_2(X) \prod_{j \neq i} (X - \lambda_j)^{q_{\lambda_j}} = 1 = \sum_{l=1}^{k} A_l(X) \prod_{j \neq l} (X - \lambda_j)^{q_{\lambda_j}}.$$

Du coup [1]

$$Q_2(X) = A_i(X).$$

Ceci donne lieu à la formule suivante pour le projecteur spectrale associé à la valeur propre λ_i

$$\Pi_{\lambda_i} = A_i(u) \prod_{j \neq i} (u - \lambda_j)^{q_{\lambda_j}}.$$

Exemple 7.3.1. Soit

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 1 & -1 & -6 \\ -1 & 2 & 5 \end{pmatrix}.$$

On trouve que le polynôme caractéristique est $P_A(\lambda) = (1 - \lambda)(2 - \lambda)^2$. D'ou $\lambda_1 = 1$ et $\lambda_2 = 2$ sont les valeurs propres. On trouve que les multiplicités géométriques des deux valeurs propres sont 1. Donc A n'est pas diagonalisable et le polynôme minimal coincide avec le polynôme caractéristique, $m_A = P_A$. Pour trouver les projecteurs spectraux on doit alors résoudre

$$\frac{1}{(X-1)(X-2)^2} = \frac{A_1(X)}{(X-1)} + \frac{A_2(X)}{(X-2)^2}.$$

avec A_1 de degré 0 et A_2 de degré au plus 1. La solution est

$$A_1(X) = 1, \quad A_2(X) = 3 - X.$$

1. et $Q_1(X) = \sum_{l \neq i}^{k} A_l(X) \prod_{j \neq l} (X - \lambda_j)^{q_{\lambda_j}}$, mais on n'a pas besoin de ca

D'où

$$\Pi_1 = (A - 2\mathbf{1}_3)^2 = \begin{pmatrix} -1 & 0 & -3 \\ 1 & -3 & -6 \\ -1 & 2 & 3 \end{pmatrix}^2 = \begin{pmatrix} 4 & -6 & -6 \\ 2 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\Pi_2 = (3\mathbf{1}_3 - A) \circ (A - \mathbf{1}_3) = \begin{pmatrix} -3 & 6 & 6 \\ -2 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

La décomposition de Dunford est donc donné par $A = \delta + \eta$ avec

$$\delta = \Pi_1 + 2\Pi_2 = \begin{pmatrix} -2 & 6 & 6 \\ -2 & 5 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

et

$$\eta = A - \delta = \begin{pmatrix} 3 & -6 & -9 \\ 3 & -6 & -9 \\ -1 & 2 & 3 \end{pmatrix}.$$

Chapitre 8

Décomposition en sous espaces irréductibles et forme de Jordan

L'algorithme de trigonalisation d'un endomorphisme u, qu'on a vu dans Chapitre 6 (qui correspond à la décomposition spectrale de Dunford), se base sur la décomposition de l'espace vectoriel en espaces caractéristiques. Il nous a permit de trouver une base dans laquelle l'endomorphisme se décompose en blocs qui sont triangulaires, chaque bloc correspondant à la reduction de u à un de ses espaces caractéristiques. Il y a donc une correspondance un à un entre les blocs et les valeurs propres de u. Les espaces caractéristiques sont stables pour u, mais pas forcément irreductibles. Dans ce chapitre on étudie une décomposition plus fine, en sous-espaces irreductibles. Plus précisement, on va décomposer les espaces caractéristiques en sous espaces irreductibles. Les blocs obtenues par l'algorithme de trigonalisation du Chapitre 6 se décomposent alors en des blocs plus petites.

8.1 Sous-espaces irréductibles

Proposition 28. *Soit u un endomorphisme nilpotent sur un espace vectoriel E, de dimension n. Sont équivalent :*

1. *E est cyclique pour u, c.à.d. is existe $x \in E$ t.q. $E = \langle x\rangle_u$.*
2. *Il existe $x \in E$ t.q. $u^{n-1}(x) \neq 0_E$.*
3. *L'indice de nilpotence de u est n.*
4. *$\ker u$ est de dimension 1.*

Démonstration.

$1 \Rightarrow 2$ Soit $x \in E$ t.q. $E = \langle x\rangle_u$. Par construction $\langle x\rangle_u = \mathrm{Vect}\{x, u(x), \cdots, u^{k-1}(x)\}$ avec $u^{k-1}(x) \neq 0_E$ et $u^k(x)$ est combinaison linéaire des $x, u(x), \cdots, u^{k-1}(x)$. Comme $k = \dim\langle x\rangle_u$ l'hypothèse 1 implique que $k = n$.

$2 \Rightarrow 1$ Soit $x \in E$ t.q. $u^{n-1}(x) \neq 0_E$. Alors $\langle x\rangle_u = \mathrm{Vect}\{x, \cdots u^{n-1}(x)\}$ car $u^n = 0$. Soit $k = \dim\langle x\rangle_u$. Alors $u_{\langle x\rangle_u}$ est nilpotent avec indice de nilpotence $\leq k$. Comme $u^{n-1}(x) \neq 0_E$ on a $n \leq k$. Donc $n = k$. Donc $\{x, \cdots u^{n-1}(x)\}$ est une famille libre, qui engendre alors E.

$2 \Leftrightarrow 3$ Vu que l'indice de nilpotence de u est au plus n cette équivalence est clair.

$1 \Rightarrow 4$ Soit $x \in E$ t.q. $E = \langle x\rangle_u$. $(x, u(x), \cdots, u^{n-1}(x))$ est alors une base pour E. Soit $y \in E$. Il existe $a_0, \cdots, a_{n-1} \in \mathbb{K}$ t.q. $y = a_0x + a_1u(x) + \cdots + a_{n-1}u^{n-1}(x)$. On a

$$u(y) = a_0u(x) + a_1u^2(x) + \cdots + a_{n-2}u^{n-1}(x).$$

Comme les vecteurs $u(x), \cdots, u^{n-1}(x)$ sont libres on a $u(y) = 0_E$ si et seulement si $a_0, \cdots, a_{n-2} = 0$. Donc $\ker u$ est de dimension 1.

$4 \Rightarrow 3$ Par le théorème du rang, $\mathrm{rg}(u^{k+1}) \geq \mathrm{rg}(u^k) - \dim\ker u$. Comme $\dim\ker u = 1$ on a $\mathrm{rg}(u) = n-1$ et donc l'inégalité implique $\mathrm{rg}(u^k) \geq n-k$. D'où l'indice de nilpotence de u est n.

□

Corollaire 19. *Soit u un endomorphisme nilpotent sur E. Si E est cyclique, alors il existe une base $\mathcal{B}$ t.q.*

$$[u]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & 0 & \vdots \\ & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \cdots & & \cdots & 0 \end{pmatrix}$$

Démonstration. D'après Prop. 28 il existe $b_1 \in E$ t.q. $u^{\dim E - 1}(b_1) \neq 0_E$. Alors $\mathcal{B}' = (b_1, \cdots, b_n) = (b_1, u(b_1), \cdots, u^{n-1}(b_1))$ est une base pour E. Comme $u(b_j) = b_{j+1}$ la matrice de u dans cette base est triangulaire inférieure

$$[u]_{\mathcal{B}'} = \begin{pmatrix} 0 & \vdots & & \cdots & 0 \\ 1 & 0 & & & \vdots \\ 0 & \ddots & \ddots & & \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

D'ou la matric de u dans la base $\mathcal{B} = (u^{n-1}(b_1), \cdots, b_1)$ a la forme démandée. □

Corollaire 20. *Soit u un endomorphisme nilpotent sur E de dimension finie. Si E est cyclique alors il est irréductible.*

Démonstration. Supposons, par absurde, que $E = F_1 \oplus F_2$ ou F_1, F_2 sont non-triviaux (différent de $\{0\}$) et stable par u. u_{F_1} est nilpotent avec indice de nilpotence $\leq \dim F_1 \leq \dim E - 1$. De même u_{F_2} est nilpotent avec indice de nilpotence $\leq \dim F_2 \leq \dim E - 1$. Alors l'indice de nilpotence de u est au plus $\dim E - 1$. E ne contient donc pas d'élément x t.q. $u^{\dim -1}(x) \neq 0_E$. Donc E n'est pas cyclique. □

8.2 Décomposition en sous espaces irréductibles

Lemme 20. *Soit $F \subset E$ un sous espace cyclique d'un endomorphisme nilpotent u. Si $y \in F \backslash \mathrm{im}\, u$ alors $F = \langle y \rangle_u$.*

Démonstration. Par l'hypothèse il existe $x \in F$ t.q. $F = \langle x \rangle_u = \mathrm{Vect}\{x, u(x), \cdots u^{k-1}(x)\}$ où $k = \dim F$. Soit $y \in F \backslash \mathrm{im}\, u$. Il existent alors λ_i t.q. $y = \lambda_0 x + \lambda_1 u(x) + \cdots + \lambda_{k-1} u^{k-1}(x)$. Comme $y \notin \mathrm{im}\, u$ ceci implique $\lambda_0 \neq 0$. D'où $u^{k-1}(y) = \lambda_0 u^{k-1}(x) \neq 0$. Il en suit que $\{y, u(y), \cdots u^{k-1}(y)\}$ est une famille libre, donc génératrice pour F. □

Théorème 21. *Soit u un endomorphisme nilpotent sur un espace vectoriel E de dimension finie. Soit $k = \dim\ker u$. Alors E est la somme directe de k sous espaces cycliques (donc irreductibles).*

Démonstration. La preuve est par récurrence sur la dimension de E. Si $\dim E = 1$ il n'y a rien à montrer. Soit $\dim E = n > 1$. Comme u est nilpotent, $\mathrm{rg}(u) < n$. Il existe donc un sous-espace H de dimension $n-1$ t.q. $\mathrm{im}\ u \subset H$. On a $u(H) \subset \mathrm{im}\ u \subset H$, donc H est stable pour u.

Par hypothèse de récurrence appliqué à H et u_H il existe des sous-espaces cycliques $H_1, \cdots, H_r$ pour u_H t.q.

$$H = H_1 \oplus \cdots \oplus H_r.$$

On note que les sous-espaces H_i sont aussi cyclique pour u.

Choisit $e \in E$, $e \notin H$ (ceci est possible car $H \neq E$). Comme $u(e) \in \text{im } u$ il existent $y_i \in H_i$ t.q.

$$u(e) = y_1 + \cdots + y_r. \tag{8.1}$$

Si $y_i = u(x_i)$ pour $x_i \in H_i$, alors on remplace e par $e - x_i$, ce qui appartient à E, mais pas à H. En faisant cette remplacement pour tout les i necessaires, nous pouvons alors supposer que, pour tout $i = 1, \cdots, r$, soit $y_i = 0$ soit $y_i \notin u(H_i)$.

On considère d'abord le cas $u(e) = 0$. Dans ce cas $\langle e \rangle_u = \mathbb{K}e$ ce qui implique $\langle e \rangle_u \cap H = \{0\}$ et donc $E = \mathbb{K}e \oplus H = \mathbb{K}e \oplus H_1 \oplus \cdots \oplus H_r$, une décomposition en sous espaces cycliques.

Maintenant on considère le cas que $u(e) \neq 0$. Il existe alors i t.q. le $y_i \neq 0$ dans la décomposition de (8.1). Quitte à ré-indexer les espaces H_i, on peut supposer que $m = \dim H_1$ est la plus grande valeur de $\dim H_i$ t.q. $y_i \neq 0$. Par construction on a $y_1 \notin \text{im } u$ et donc, par Lemme 20, $H_1 = \langle y_1 \rangle_u$. Alors $u^{m-1}(y_1) \neq 0$ et $u^m(y_1) = 0$ ce qui entraine que $u^m(e) = u^{m-1}(y_1) + \cdots u^{m-1}(y_r) \neq 0$ et $u^m(u(e)) = 0$. Ces deux dernières équations implique que $\dim\langle e \rangle_u = m + 1$. D'où

$$\dim\langle e \rangle_u + \dim(H_2 \oplus \cdots \oplus H_r) = m + 1 + (n - m - 1) = n$$

Du coup, la somme $\langle e \rangle_u + H_2 \oplus \cdots \oplus H_r$ est directe et égal à E si

$$\langle e \rangle_u \cap H_2 \oplus \cdots \oplus H_r = \{0\}.$$

Pour vérifier cela nous prenons un élément z dans cette intersection. Il est donc de la forme $z = \lambda_0 e + \lambda_1 u(e) + \cdots + \lambda_m u^m(e)$ et appartient à $H_2 \oplus \cdots \oplus H_r$. Alors

$$\lambda_0 e = z - u(\lambda_1 e + \cdots + \lambda_m u^{m-1}(e)) \in H_2 \oplus \cdots \oplus H_r + \text{im } u \subset H.$$

Or, on a supposé que $e \notin H$. Donc $\lambda_0 = 0$. D'où

$$z = \lambda_1 u(e) + \cdots + \lambda_m u^m(e) = \lambda_1 y_1 + \cdots + \lambda_m u^{m-1}(y_1) + w$$

avec $w \in H_2 \oplus \cdots \oplus H_r$. En effet, par (8.1) $w = \sum_{i=2}^{n} \lambda_1 y_i + \cdots + \lambda_m u^{m-1}(y_i)$. On a $z - w \in H_1$. Mais aussi $z - w \in H_2 \oplus \cdots \oplus H_r$. D'où

$$0 = z - w = \lambda_1 u(e) + \cdots + \lambda_m u^m(e) = z - \lambda_0 e = z.$$

Comme la réduction d'un endomorphisme nilpotent sur un space espace cyclique a un noyau de dimension 1, u a un noyau de dimension égal au nombre de composantes irreductibles dans le décomposition. □

8.3 Forme de Jordan d'un endomorphisme trigonalisable

Une matrice de taille $n \times n$ de la forme

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

s'appelle un bloc de Jordan.

Théorème 22 (Forme de Jordan)**.** *Soit u un endomorphism trigonalisable de E, de dimension finie. Il existe une base $\mathcal{B}$ t.q. $[u]_{\mathcal{B}}$ est diagonal en blocs et chaque bloc à la forme de Jordan.*

Démonstration. On décompose E en les espaces caractéristiques F_{λ_i} de u. L'endomorphisme

$$h_\lambda = u_{F_\lambda} - \lambda \mathrm{id}_{F_\lambda}$$

est nilpotent sur F_λ. Par Théorème 21 F_λ se décompose en $r_\lambda = \dim\ker h_\lambda$ sous-espaces cycliques $F_{\lambda,j}$. Par Cor. 19 il existe une base $\mathcal{B}_{\lambda,j}$ pour $F_{\lambda,j}$ t.q. $[h_{\lambda F_{\lambda,j}}]_{\mathcal{B}_{\lambda,j}} = J_{0,n'}$ avec $n' = \dim F_{\lambda,j}$. Il en suit que

$$[u_{F_{\lambda,j}}]_{\mathcal{B}_{\lambda,j}} = J_{\lambda,\dim F_{\lambda,j}}.$$

On prenant $\mathcal{B}$ comme l'union de ces bases on trouve que $[u]_{\mathcal{B}}$ est diagonal en blocs ou chaque bloc corresponds à une matrice $[u_{F_{\lambda,j}}]_{\mathcal{B}_{\lambda,j}}$. □

Calcul des puissances II

La forme de Jordan simplifie le calcul les puissances. En effet

$$J_{\lambda,m} = \lambda \mathbf{1}_m + N, \quad N = J_{0,m}$$

est la décomposition spectral de $J_{\lambda,n}$, et les puissances N^k sont faciles à calculer : le coefficient ${N^k}_{ij} = 1$ si $j - i = k$ et 0 ailleurs. D'ou, équation 6.4 implique

$$J_{\lambda,m}^n = \begin{pmatrix} \lambda^n & c_n^1\lambda^{n-1} & c_n^2\lambda^{n-2} & \cdots & c_n^{m-1}\lambda^{n-m+1} \\ \vdots & \lambda^n & c_n^1\lambda^{n-1} & c_n^2\lambda^{n-2} & \vdots \\ & & \ddots & \ddots & \\ & & & & c_n^2\lambda^{n-2} \\ \vdots & & & \lambda^n & c_n^1\lambda^{n-1} \\ 0 & \cdots & & \cdots & \lambda^n \end{pmatrix}$$

avec $c_n^j = \frac{n!}{(n-j)!j!}$ et $c_n^j = 0$ si $j > n$.

Calcul de l'exponentiel d'une matrice trigonalisable

On rappelle (analyse) que la fonction exponentielle admet une developpement illimité

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

De plus, la série converge pour to $x \in \mathbb{C}$ et donc peut servir aussi pour définir $\exp(z)$ pour $z \in \mathbb{C}$. De plus, on peut définir la fonction exponentielle pour une matrice (carrée) de la même manière

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Nous admettons que la série matricielle converge, c.à.d. que tous les coefficients convergent. En particulier, pour tout $t \in \mathbb{R}$,

$$\exp(tA) = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}.$$

La forme de Jordan pour A permet de calculer explicitement $\exp(tA)$. Il suffit de connaître ce resultat pour un bloc de Jordan.

Lemme 21. *On a*

$$\exp tJ_{\lambda,m} = \begin{pmatrix} e^{t\lambda} & te^{t\lambda} & \frac{t^2}{2}e^{t\lambda} & \cdots & \frac{t^{m-1}}{(m-1)!}e^{t\lambda} \\ \vdots & e^{t\lambda} & te^{t\lambda} & \frac{t^2}{2}e^{t\lambda} & \vdots \\ & & \ddots & \ddots & \\ & & & & \frac{t^2}{2}e^{t\lambda} \\ \vdots & & & e^{t\lambda} & te^{t\lambda} \\ 0 & \cdots & & \cdots & e^{t\lambda} \end{pmatrix}$$

Démonstration. Soit $0 \leq k \leq n$. Alors

$$\frac{t^n}{n!}c_n^k\lambda^{n-k} = \frac{t^k}{k!}\frac{t^{n-k}}{(n-k)!}\lambda^{n-k}.$$

Comme $c_n^k = 0$ si $k > n$ ceci implique

$$\sum_{n=0}^{\infty}\frac{t^n}{n!}c_n^k\lambda^{n-k} = \sum_{n=k}^{\infty}\frac{t^k}{k!}\frac{t^{n-k}}{(n-k)!}\lambda^{n-k} = \frac{t^k}{k!}e^{t\lambda}$$

D'où le resultat. □