La classification des endomorphismes. Le théorème de Jordan

Andrei Teleman

Aix-Marseille Université

21 octobre 2021

Table des matières

1 Polynômes et équations algébriques **1**
1.1 L'anneau des polynômes . . . 1
1.2 La division euclidienne dans l'anneau des polynômes . . . 4
1.3 Le théorème de Gauss-d'Alembert . . . 6

2 Diagonalisation et trigonalisation **8**
2.1 Sous-espaces en sommes directe, décompositions en sommes directe . . . 8
2.2 Valeurs propres et espaces propres . . . 9
2.3 Critères de diagonalisabilité . . . 12
2.4 Trigonalisation . . . 14

3 Polynôme d'un endomorphisme. Idéal annulateur. Polynôme minimal. Le lemme des noyaux **16**
3.1 Polynôme d'un endomorphisme. Idéal annulateur. Polynôme minimal . . . 16
3.2 Le lemme des noyaux. Sous-espaces caractéristiques . . . 19
3.2.1 Le lemme des noyaux . . . 19
3.2.2 Espaces caractéristiques . . . 22

4 Endomorphismes nilpotents. La réduction de Jordan des endomorphismes nilpotents **25**
4.1 Endomorphismes nilpotents . . . 25
4.2 La réduction de Jordan des endomorphismes nilpotents . . . 27

5 Le théorème de Jordan et la décomposition de Dunford **30**
5.1 Le théorème de Jordan . . . 30
5.2 La décomposition de Dunford . . . 32
5.3 L'exponentielle d'un endomorphisme . . . 34

1 Polynômes et équations algébriques

1.1 L'anneau des polynômes

Soit $\mathbb{K}$ l'un corps $\mathbb{R}$ ou $\mathbb{C}$. On commence par la

Définition 1.1 *L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est défini par*

$$\mathbb{K}[X] := \{(a_k)_{k\geqslant 0} |\ \forall k \in \mathbb{N},\ a_k \in \mathbb{K},\ \exists N \in \mathbb{N}\ t.q.\ \forall k > N,\ a_k = 0\}\ .$$

Donc la donnée d'un polynôme à coefficients dans $\mathbb{K}$ est équivalente à la donnée d'une suite $(a_k)_{k\geqslant 0}$ de $\mathbb{K}$ dont tous les termes sont nuls à partir d'un certain indice. Cet ensemble a une structure naturelle de $\mathbb{K}$-espace vectoriel, l'addition dans cet espace vectoriel étant donnée par

$$((a_k)_{k\geqslant 0}) + ((b_k)_{k\geqslant 0}) := (a_k + b_k)_{k\geqslant 0}\ ,$$

et la multiplication par les scalaires dans cet espace vectoriel étant donnée par la formule

$$a\left((a_k)_{k\geqslant 0}\right) := (aa_k)_{k\geqslant 0} \ .$$

Pour tout $k \in \mathbb{N}$ notons par $e_k \in \mathbb{K}[X]$ la suite définie par

$$(e_k)_i = \delta_{ik} \ ,$$

donc

$$e_0 = (1, 0, 0 \dots) \ \ e_1 = (0, 1, 0, 0, \dots) \ , \ e_k = (0, \dots, 0, 1, 0, 0 \dots) \text{ avec 1 sur la } k\text{ème place} \ .$$

C'est facile de vérifier que la famille $(e_k)_{k\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$, donc $\mathbb{K}[X]$ est un $\mathbb{K}$-espace vectoriel de dimension infinie. Un polynôme de la forme $ae_0 = (a, 0, 0 \dots)$ sera appelé *polynôme constant*.

Définition 1.2 *La multiplication dans $\mathbb{K}[X]$ est la loi de composition interne donnée par la formule*

$$((a_k)_{k\geqslant 0})((b_l)_{l\geqslant 0}) = (c_n)_{n\geqslant 0} \ \textit{où} \ c_n := \sum_{k+l=n} a_k b_l = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{l=0}^{n} a_{n-l} b_l \ .$$

C'est facile de vérifier que cette loi de composition interne a les propriétés suivantes :

1. est associative,
2. admet un élément neutre, à savoir e_0,
3. est commutative,
4. est distributive par rapport à l'addition dans $\mathbb{K}[X]$.

On va désigner l'élément $e_0 \in \mathbb{K}[X]$ par 1 (en sous-entendant l'élément neutre de la multiplication définie dans $\mathbb{K}[X]$) et le polynôme constant $ae_0 = (a, 0, 0 \dots)$ sera noté simplement par a_0, donc chaque scalaire $a \in \mathbb{K}$ sera identifié avec le polynôme constant $ae_0 = (a, 0, 0 \dots)$ qui lui correspond.

En utilisant la définition de la multiplication c'est facile de vérifier que pour tout $k \in \mathbb{N}^*$ on a

$$(e_1)^k = e_k \ .$$

On va poser $X := e_1$. Avec la convention $X^0 := 1$ on obtient l'égalité suivante

$$(a_k)_{k\geqslant 0} = \sum_{k\geqslant 0} a_k X^k \ ,$$

(somme qui contient seulement un nombre fini de termes non-nuls). Cette égalité importante fait la liaison entre la définition moderne de la notion de polynôme (voir la définition 1.1), et la manière élémentaire d'introduire cette notion, à savoir comme une expression algébrique de la forme

$$a_0 + a_1 X + \cdots + a_N X^N \ .$$

La multiplication des polynômes donnée par la définition 1.2 correspond à la multiplication des expressions algébriques introduite au lycée.

On peut montrer facilement que $\mathbb{K}[X]$, muni de l'addition et de la multiplication des polynômes, est *un anneau commutatif avec unité*.

Définition 1.3 *Le degré d'un polynôme $P(X) = \sum_{k\geqslant 0} a_k X^k$ est défini par*

$$\deg(P(X)) := \begin{cases} \max\{k \in \mathbb{N} |\ a_k \neq 0\} & \text{si} \quad P(X) \neq 0 \\ -\infty & \text{si} \quad P(X) = 0 \end{cases} \ .$$

Donc le degré d'un polynôme *non-nul* est (l'exposant de) la puissance maximale de X qui intervient effectivement dans l'expression de $P(X)$. Avec cette définition on obtient facilement la formule générale

$$\deg(P(X)Q(X)) = \deg(P(X)) + \deg(Q(X)) \tag{1}$$

En utilisant cette formule on obtient

Remarque 1.4 *1. L'anneau $\mathbb{K}[X]$ est intègre, i.e. pour tous deux polynômes $P(X)$, $Q(X) \in \mathbb{K}[X]$ nous avons l'implication*

$$P(X)Q(X) = 0 \Rightarrow \big(P(X) = 0 \text{ ou } Q(X) = 0\big) \ .$$

2. Les seuls polynômes inversibles dans l'anneau $\mathbb{K}[X]$ sont les polynômes constants non-nuls.

Dans cette remarque la condition "inversible" signifie "inversible par rapport à la multiplication des polynômes", donc, par définition, un polynôme $P(X)$ est inversible si et seulement si il existe $Q(X) \in \mathbb{K}[X]$ tel que $P(X)Q(X) = 1$. Si on multiplie un polynôme $P(X)$ par un élément inversible (donc par un polynôme constant non-nul) on va obtenir un polynôme qui a les mêmes propriétés de divisibilité (par exemple les mêmes diviseurs et les mêmes multiples) que $P(X)$. Cette remarque montre que, pour l'étude des problèmes de divisibilité dans $\mathbb{K}[X]$ il suffit de se concentrer sur les *polynômes unitaires* au sens de la définition suivante :

Définition 1.5 *Un polynôme $P(X) \in \mathbb{K}[X]\backslash\{0\}$ est dit unitaire si son coefficient du terme de plus haut degré (son coefficient dominant) est égal à 1.*

Par définition, un polynôme unitaire s'écrit donc sous la forme

$$X^d + \sum_{i=0}^{d-1} a_i X^i \text{ où } d = \deg(P(X)) \ .$$

Remarquons que tout polynôme $P(X) \in \mathbb{K}[X]\backslash\{0\}$ s'écrit d'une manière unique comme un produit

$$P(X) = a\tilde{P}(X)$$

où $\tilde{P}(X)$ est un polynôme unitaire et $a \in \mathbb{K}^*$ est le coefficient du terme de plus haut degré de $P(X)$.

Définition 1.6 *Soient $P(X)$, $S(X) \in \mathbb{K}(X)$. On dit que $P(X)$ divise $S(X)$ (ou que $S(X)$ est divisible par $P(X)$) s'il existe $Q(X) \in \mathbb{K}(X)$ tel que $S(X) = Q(X)P(X)$.*

Définition 1.7 *Un polynôme $P(X)$ avec $\deg(P(X)) > 0$ est dit irréductible s'il n'admet aucune factorisation de la forme $P(X) = Q_1(X)Q_2(X)$ avec $\deg(Q_i(X)) > 0$ pour $i \in \{1, 2\}$.*

Donc un polynôme $P(X)$ de degré strictement positif est irréductible si et seulement si l'un des deux facteurs de toute factorisation de $P(X)$ est un polynôme constant (qui sera nécessairement non-nul, donc inversible). En utilisant la définition et la formule (1) on obtient facilement

Remarque 1.8 *Tout polynôme de degré 1 est irréductible.*

Dans la section suivante on va donner la classification complète de tous les polynômes irréductibles dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$. En particulier on va voir que, tandis que dans $\mathbb{C}[X]$ la réciproque de la remarque 1.8 est vraie (donc un polynôme $P(X) \in \mathbb{C}[X]$ est irréductible si et seulement si $\deg(P(X)) = 1$) dans $\mathbb{R}[X]$ la réciproque de cette remarque est fausse. En effet, il suffit de remarquer que le polynôme $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$.

Définition 1.9 *Soit $P(X) = \sum_{k\geqslant 0} a_k X^k \in \mathbb{K}(X)$. La fonction polynômiale associée à $P(X)$ est la fonction $p : \mathbb{K} \to \mathbb{K}$ définie par*

$$p(x) = P(x) = \sum_{k\geqslant 0} a_k x^k \ \forall x \in \mathbb{K} \ .$$

L'équation algébrique associée à $P(X)$ est l'équation

$$P(x) = 0 \ .$$

Une solution $x_0 \in \mathbb{K}$ de cette équation s'appelle racine de $P(X)$.

Notons que l'application

$$\mathbb{K}[X] \to \mathcal{F}(\mathbb{K}, \mathbb{K})$$

qui associe à tout polynôme sa fonction polynomiale est un morphisme d'anneaux, en particulier pour deux polynômes $P(X)$, $Q(X) \in \mathbb{K}[X]$ on a les identités

$$(P + Q)(x) = P(x) + Q(x) \ , \ (PQ)(x) = P(x)Q(x) \ \forall x \in \mathbb{K} \ .$$

1.2 La division euclidienne dans l'anneau des polynômes

On va admettre sans démonstration le théorème important suivant :

Théorème 1.10 *(le théorème de division euclidienne pour les polynômes) Soit $S(X)$, $P(X) \in \mathbb{K}[X]$ avec $P(X) \neq 0$. Alors il existe une unique paire $(Q(X), R(X)) \in \mathbb{K}[X] \times \mathbb{K}[X]$ avec $\deg(R(X)) < \deg(P(X))$ telle que*

$$S(X) = Q(X)P(X) + R(X)$$

Le polynôme $Q(X)$ s'appelle le quotient et le polynôme $R(X)$ s'appelle le reste de la division euclidienne de $S(X)$ par $P(X)$. Ce théorème a quelques corollaires très importants :

Corollaire 1.11 *Soit $a \in \mathbb{K}$. Un polynôme $P(X) \in \mathbb{K}(X)$ est divisible par $X - a$ si et seulement si $P(a) = 0$, i.e. si et seulement si a est une racine de $P(X)$.*

Démonstration: En effet si $P(X)$ est divisible par $X - a$, alors on peut écrire $P(X) = (X - a)Q(X)$ pour un polynôme $Q(X) \in \mathbb{K}[X]$. En remplaçant X par a on obtient $P(a) = 0 \cdot Q(a) = 0$. Réciproquement, supposons que $P(a) = 0$. En appliquant le théorème de division euclidienne aux polynômes $P(X)$, $X - a$ on obtient

$$P(X) = Q(X)(X - a) + R$$

où $\deg(R) \leqslant 0$, donc R est un polynôme constant. En remplaçant X par a on obtient $R = 0$, donc $P(X) = Q(X)(X - a)$, ce qui montre que $P(X)$ est divisible par $X - a$. ■

Définition 1.12 *Soit $P(X) \in \mathbb{K}(X)$ avec $\deg(P(X)) > 0$ et $a \in \mathbb{K}$ une racine de $P(X)$. La multiplicité (ou l'ordre de multiplicité) de la racine a est définie par*

$$\mathrm{mult}_P(a) := \max\{k \in \mathbb{N}^* |\ (X - a)^k | P(X)\}\ .$$

Donc $\mathrm{mult}_P(a)$ est l'exposant de la puissance maximale de $X - a$ qui divise $P(X)$.

Définition 1.13 *Soient $P_1(X), \ldots, P_k(X) \in \mathbb{K}[X]$ non tous nuls.*

1. *Un polynôme $P(X) \in \mathbb{K}(X)$ est dit diviseur commun des polynômes $P_i(X)$ si $P(X)|P_i(X)$ pour $1 \leqslant i \leqslant k$.*
2. *Un polynôme $P(X) \in \mathbb{K}(X)$ s'appelle le plus grand commun diviseur (le pgcd) des polynômes $P_i(X)$ si les conditions suivantes sont vérifiées :*
   - *(a) $P(X)$ est un diviseur commun des polynômes $P_i(X)$,*
   - *(b) $P(X)$ est un diviseur commun de degré maximal de ces polynômes,*
   - *(c) $P(X)$ est un polynôme unitaire.*

La condition 2c dans cette définition est imposée par convention pour assurer l'unicité du plus grand commun diviseur. Si on n'impose pas cette condition, le pgcd sera seulement bien défini à multiplication par une constante non-nulle près.

Corollaire 1.14 *Soient $P_1(X), \ldots, P_k(X) \in \mathbb{K}[X]$ non tous nuls. Alors*

1. *Il existe un unique pgcd des polynômes $P_i(X)$, polynôme qui sera noté $\mathrm{pgcd}(P_1(X), \ldots, P_k(X))$,*
2. *Il existe des polynômes $U_1(X), \ldots, U_k(X) \in \mathbb{K}[X]$ tels que*

$$\mathrm{pgcd}(P_1(X), \ldots, P_k(X)) = \sum_{i=1}^{k} U_i(X)P_i(X)\ . \tag{2}$$

Démonstration: On va utiliser le théorème de division euclidienne. Commençons par le cas $k = 2$.

Pour $k = 2$ on obtient facilement une démonstration effective (qui donne un algorithme explicite de calcul du pgcd) par récurrence par rapport à $\deg(P_2(X))$. Plus précisément on va démontrer par récurrence par rapport à d que pour $d \in \{-\infty\} \cup \mathbb{N} = \{-\infty, 0, 1, 2, \ldots\}$ l'affirmation suivante est vraie :

($\mathcal{P}_d$) *Toute paire de polynômes non tous nuls* $(P_1(X), P_2(X))$ *avec* $\deg(P_2(X)) = d$ *admet un pgcd unique qui est donné par une expression de la forme*

$$\mathrm{pgcd}(P_1(X), P_2(X)) = U_1(X)P_1(X) + U_1(X)P_2(X) \ .$$

Pour $d = -\infty$ (initialisation) on a $P_2(X) = 0$, donc (en tenant compte de l'hypothèse), on a nécessairement $P_1(X) \neq 0$. Remarquons que, dans ce cas, l'ensemble des diviseurs communs de $P_1(X)$ et $P_2(X)$ coïncide avec l'ensemble des diviseurs de $P_1(X)$, donc dans ce cas on a $\mathrm{pgcd}(P_1(X), P_2(X)) = \tilde{P}_1(X)$, où $\tilde{P}_1(X)$ désigne le polynôme unitaire associé à $P_1(X)$. Si on désigne par a le coefficient du terme de plus haut degré de $P_1(X)$ on obtient dans ce cas un unique pgcd donné par $\mathrm{pgcd}(P_1(X), P_2(X)) = a^{-1}P_1(X)$, et notre affirmation pour $d = -\infty$ est démontrée en choisissant $U_1(X) = a^{-1}$, $U_2(X) = 0$.

Supposons maintenant que $d > -\infty$ et que la proposition $\mathcal{P}_{d'}$ est vraie pour tout $d' < d$ (l'hypothèse de récurrence). Puisque $d > -\infty$ on a $P_2(X) \neq 0$, donc on peut appliquer le théorème de division euclidienne à la paire $(P_1(X), P_2(X))$. On obtient

$$P_1(X) = Q(X)P_2(X) + R(X) \ , \tag{3}$$

avec $\deg(R(X)) < \deg(P_2(X))$. Par l'hypothèse de récurrence on sait que la paire $(P_2(X), R(X))$ admet un unique pgcd donné par une expression de la forme

$$\mathrm{pgcd}(P_2(X), R(X)) = V_1(X)P_2(X) + V_2(X)R(X) \ . \tag{4}$$

Mais en utilisant (3) on constate que l'ensemble des diviseurs communs des polynômes $P_1(X)$, $P_2(X)$ coïncide avec l'ensemble des diviseurs communs des polynômes $P_2(X)$, $R(X)$. Donc la paire $(P_1(X), P_2(X))$ admet aussi un unique pgcd et

$$\mathrm{pgcd}(P_1(X), P_2(X)) = \mathrm{pgcd}(P_2(X), R(X)) \ .$$

En utilisant (4) puis (3) on obtient

$$\begin{aligned}\mathrm{pgcd}(P_1(X), P_2(X)) &= V_1(X)P_2(X) + V_2(X)R(X) = V_1(X)P_2(X) + V_2(X)(P_1(X) - Q(X)P_2(X)) \\ &= V_2(X)P_1(X) + (V_1(X) - V_2(X)Q(X))P_2(X) \ ,\end{aligned}$$

et notre affirmation est démontrée pour d en choisissant $U_1(X) = V_2(X)$, $U_2(X) = V_1(X) - V_2(X)Q(X)$.

Pour $k > 2$ le théorème est démontré par récurrence par rapport à k en utilisant la formule récursive (exercice)

$$\mathrm{pgcd}(P_1(X), \ldots, P_{k-1}(X), P_k(X)) = \mathrm{pgcd}(\mathrm{pgcd}(P_1(X), \ldots, P_{k-1}(X)), P_k(X)) \ .$$

Par exemple, pour $k = 3$ on a (exercice)

$$\mathrm{pgcd}(P_1(X), P_2(X), P_3(X)) = \mathrm{pgcd}(\mathrm{pgcd}(P_1(X), P_2(X)), P_3(X)) \ ,$$

donc il suffit d'appliquer deux fois la théorème d'existence pour le pgcd de deux polynômes. ■

Remarque 1.15 *La démonstration du corollaire 1.14 fournit un algorithme explicite pour le calcul du pgcd des deux polynômes, algorithme qui s'appelle l'algorithme d'Euclid :*

(1) Est-ce que $P_2(X) = 0$ *? Si* $P_2(X) = 0$ *on pose* $\mathrm{pgcd}(P_1(X), P_2(X)) = \tilde{P}_1(X)$ *et on arrête l'algorithme. Sinon on passe à l'étape suivante.*

(2) Si $P_2(X) \neq 0$ *on fait la division euclidienne de* $P_1(X)$ *par* $P_2(X)$*, on remplace la paire initiale* $(P_1(X), P_2(X))$ *par la paire* $(P_2(X), R(X))$ *où* $R(X)$ *est le reste de cette division euclidienne, et on revient à la question (1), posée pour la nouvelle paire.*

Donc le pgcd cherché coïncide avec `le polynôme unitaire associé au` *dernier reste non-nul obtenu dans cette suite finie de divisions euclidiennes.*
Pour le calcul du pgcd dans la cas général $k \geqslant 2$ on utilise la formule générale

$$\mathrm{pgcd}(P_1(X), \ldots, P_{k-1}(X), P_k(X)) = \mathrm{pgcd}(\mathrm{pgcd}(P_1(X), \ldots, P_{k-1}(X)), P_k(X)) \ ,$$

qui permet de réduire le problème au cas $k = 2$.

Définition 1.16 *On dit que k polynômes $P_1(X), \ldots, P_k(X) \in \mathbb{K}[X]$, non tous nuls, sont dits premiers entre eux si $\mathrm{pgcd}(P_1(X), \ldots, P_k(X)) = 1$, i.e. si ces polynômes n'admettent aucun diviseur commun de degré strictement positif.*

Corollaire 1.17 *(Théorème de Bézout) Soient $P_1(X), \ldots, P_k(X) \in \mathbb{K}[X]$ non tous nuls. Alors les conditions suivantes sont équivalentes :*

1. *$P_1(X), \ldots, P_k(X)$ sont premiers entre eux,*
2. *il existe $U_1(X), \ldots, U_k(X) \in \mathbb{K}[X]$ tels que $\sum_{i=1}^{k} U_i(X)P_i(X) = 1$.*

Le corollaire suivant sera énoncé sans démonstration. La démonstration est proposée comme exercice. Pour comprendre l'idée de démonstration commencer par le cas particulier d'un polynôme de degré 2, puis essayer de généraliser votre argument.

Corollaire 1.18 *Tout polynôme $P(X) \in \mathbb{K}(X)$ avec $\deg(P(X) > 0$ admet une factorisation de la forme*

$$P(X) = a(Q_1(X))^{m_1} \ldots (Q_k(X))^{m_k} = a \prod_{i=1}^{k} (Q_i(X))^{m_i} \ ,$$

où a est le coefficient du terme de plus haut degré de $P(X)$, $m_i \in \mathbb{N}^*$, et $Q_i(X)$ sont des polynômes irréductibles unitaires distincts deux à deux. Cette factorisation est unique, à permutations des facteurs près.

Si tous les polynômes irréductibles unitaires $Q_i(X)$ dans cette factorisation sont des polynômes du premier degré, on va dire que $P(X)$ est *scindé dans* $\mathbb{K}$. Remarquer qu'un polynôme *unitaire* du premier degré s'écrit sous la forme $X - a$ avec $a \in \mathbb{K}$. Notre définition devient :

Définition 1.19 *Un polynôme $P(X) \in \mathbb{K}(X)$ avec $\deg(P(X) > 0$ est dit scindé dans $\mathbb{K}$ s'il admet une factorisation de la forme*

$$P(X) = a(X - a_1)^{m_1} \ldots (X - a_k)^{m_k} = a \prod_{i=1}^{k} (X - a_i)^{m_i}$$

où $m_i \in \mathbb{N}^*$, et $a_i \in \mathbb{K}$ sont distincts deux à deux.

Remarquer que si $P(X)$ admet une telle factorisation, alors

1. a est le coefficient du terme de plus haut degré de $P(X)$,
2. $\deg(P(X)) = \sum_{i=1}^{k} m_i$,
3. l'ensemble des racines de $P(X)$ est $\{a_1, \ldots, a_k\}$,
4. $\mathrm{mult}_P(a_i) = m_i$ pour $1 \leqslant i \leqslant k$.

1.3 Le théorème de Gauss-d'Alembert

Le théorème de Gauss d'Alembert, où le théorème fondamental de l'algèbre, a beaucoup de conséquences importantes dans plusieurs domaines des mathématiques modernes. Nous allons utiliser ce résultat plus tard dans l'étude des endomorphismes d'un espace vectoriel de dimension finie. La démonstration de ce théorème dépasse le niveau et le but de ce cours, donc sera omise.

Théorème 1.20 *Tout polynôme $P(X) \in \mathbb{C}[X]$ avec $\deg(P(X)) > 0$ admet une racine complexe.*

En utilisant la terminologie de l'algèbre moderne, ce théorème affirme que $\mathbb{C}$ est un corps algébriquement clos. En général, un corps K est dit algébriquement clos si tout polynôme $P(X) \in K[X]$ avec $\deg(P(X)) > 0$ admet une racine dans K. Remarquer que $\mathbb{R}$ n'est pas algébriquement clos parce que, par exemple, le polynôme $X^2 + 1 \in \mathbb{R}[X]$ d'admet aucune racine réelle.

D'après le corollaire 1.11 l'existence d'une racine $a \in \mathbb{C}$ d'un polynôme $P(X) \in \mathbb{C}[X]$ implique la divisibilité de $P(X)$ par $X - a$. Si le quotient de $Q(X)$ de $P(X)$ par $X - a$ est encore un polynôme de degré strictement positif (i.e. si $\deg(P(X)) \geqslant 2$), on peut appliquer le théorème 1.20 à $Q(X)$ et on déduit que $Q(X)$ est aussi divisible par une polynôme de la forme $X - b$. Par récurrence on obtient donc

Corollaire 1.21 *Tout polynôme $P(X) \in \mathbb{C}[X]$ avec $\deg(P(X)) > 0$ est scindé dans $\mathbb{C}$ donc se factorise sous la forme*

$$P(X) = a(X - a_1)^{m_1} \dots (X - a_k)^{m_k} = a \prod_{i=1}^{k} (X - a_i)^{m_i} \tag{5}$$

où a est le coefficient du terme de plus haut degré de $P(X)$, $m_i \in \mathbb{N}^*$ et $a_i \in \mathbb{K}$ sont distincts deux à deux.

Comme nous avons vu dans la section précédente, dans la factorisation (5) a est le coefficient du terme de plus haut degré de $P(X)$, on a $\deg(P(X)) = \sum_{i=1}^{k} m_i$, l'ensemble des racines de $P(X)$ est $\{a_1, \dots, a_k\}$, et $\mathrm{mult}_P(a_i) = m_i$ pour $1 \leqslant i \leqslant k$.

Le dernier corollaire donne une classification complète des polynômes irréductibles dans $\mathbb{K}[X]$ pour $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$:

Corollaire 1.22 *1. L'ensemble des polynômes irréductibles unitaires de $\mathbb{C}[X]$ est*

$$\{X - a |\ a \in \mathbb{C}\}.$$

2. L'ensemble des polynômes irréductibles unitaires de $\mathbb{R}[X]$ est

$$\{X - a \mid a \in \mathbb{R}\} \cup \{X^2 + bX + c \mid a,\ b \in \mathbb{R},\ b^2 - 4c < 0\}.$$

La première partie du corollaire 1.22 est une conséquence directe de la remarque 1.8 et du corollaire 1.21. Pour la deuxième partie remarquons d'abord que tout polynôme du 2me degré à coefficients réels *à discriminant strictement négatif* est irréductible dans $\mathbb{R}[X]$. En effet, puisque un tel polynôme n'admet pas de racines réelles, il n'admet aucune factorisation (dans $\mathbb{R}[X]$!) en produit de polynômes du 1er degré. Pour démontrer que, réciproquement, tout polynôme irréductible dans $\mathbb{R}[X]$ est soit de la forme $X - a$, soit de la forme $X^2 + bX + c$ avec $b^2 - 4c < 0$, on utilise la proposition suivante concernant l'ensemble des racines *complexes* d'un polynôme à coefficients réels :

Proposition 1.23 *Soit $P(X) \in \mathbb{R}[X]$ avec $\deg(P(X)) > 0$ est soit $\mathcal{R}_P \subset \mathbb{C}$ l'ensemble des racines complexes de $P(X)$. Alors*

1. *$\mathcal{R}_P$ est stable par conjugaison, i.e on a l'implication $z \in \mathcal{R}_P \Rightarrow \bar{z} \in \mathcal{R}_P$.*
2. *Si $z \in \mathcal{R}_P$ on a*
$$\mathrm{mult}_P(z) = \mathrm{mult}_P(\bar{z})\ .$$
3. *En posant*
$$\mathcal{R}_P = \{a_1, \dots, a_l, \alpha_1, \bar{\alpha}_1, \dots, \alpha_m, \bar{\alpha}_m\}$$
avec $a_i \in \mathbb{R}$ et $\Im(\alpha_i) > 0$, et en désignant par a le coefficient du terme de plus haut degré de $P(X)$ on obtient la factorisation
$$P(X) = a \prod_{i=1}^{l} (X - a_i)^{m_i} \prod_{j=1}^{m} (X - \alpha_j)^{\mu_j} (X - \bar{\alpha}_j)^{\mu_j} = a \prod_{i=1}^{l} (X - a_i)^{m_i} \prod_{j=1}^{m} (X^2 - 2\Re(\alpha_j)X + |\alpha_j|^2)^{\mu_j}\ .$$

Remarquer que (puisque $\Im(\alpha_i) > 0$) le discriminant du polynôme $X^2 - 2\Re(\alpha_j)X + |\alpha_j|^2 \in \mathbb{R}[X]$ est bien strictement négatif. La proposition 1.23 montre en particulier qu'un polynôme $P(X) \in \mathbb{R}[X]$ de degré $d \geqslant 3$ ne peut pas être irréductible dans $\mathbb{R}[X]$, donc tout polynôme irréductible dans $\mathbb{R}[X]$ est de degré $\leqslant 2$. Le corollaire 1.22 est une conséquence directe de cette remarque.

2 Diagonalisation et trigonalisation

On commence par un récapitulatif sur les sommes directes.

2.1 Sous-espaces en sommes directe, décompositions en sommes directe

Soient K un corps, E un K-espace vectoriel de dimension finie, $k \in \mathbb{N}^*$ et $F_1, \ldots, F_k \subset E$ des sous-espaces vectoriels de E.

Définition 2.1 *On dit que les sous-espaces $F_1, \ldots, F_k$ sont en somme directe si une des conditions équivalentes suivantes est vérifiée :*

1. *Pour toute famille $(x_1, \ldots, x_k)$ de vecteurs de E tels que $x_i \in F_i$ pour $1 \leqslant i \leqslant k$, l'implication suivante est vraie*
$$\left\{ \sum_{i=1}^{k} x_i = 0_E \right\} \Rightarrow \{x_1 = \cdots = x_k = 0_E\} \ .$$
2. *Pour tout $j \in \{1, \ldots, k\}$ on a*
$$F_j \cap \left(\sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} F_i \right) = \{0_E\} \ .$$
3. *Tout vecteur $x \in \sum_{i=1}^{k} F_i$ admet une unique décomposition de la forme*
$$x = \sum_{i=1}^{k} x_i \ \text{ où } x_i \in F_i \ \forall i \in \{1, \ldots, k\} \ .$$
4. *Si B_i est une base de F_i pour $1 \leqslant i \leqslant k$ alors la famille concaténée $B_1 \ldots B_k$ est libre dans E.*

Définition 2.2 *On dit que les sous-espaces $F_1, \ldots, F_k$ donnent une décomposition en somme directe de E, et on écrit $E = \bigoplus_{i=1}^{k} F_i$, si une des conditions équivalentes suivantes est vérifiée :*

1. *$\sum_{i=1}^{k} F_i = E$ et les sous-espaces $F_1, \ldots, F_k$ sont en somme directe.*
2. *$\sum_{i=1}^{k} F_i = E$ et pour tout $j \in \{1, \ldots, k\}$ on a*
$$F_j \cap \left(\sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} F_i \right) = \{0_E\} \ .$$
3. *Tout vecteur $x \in E$ admet une unique décomposition de la forme*
$$x = \sum_{i=1}^{k} x_i \ \text{ où } x_i \in F_i \ \forall i \in \{1, \ldots, k\} \ .$$
4. *Si B_i est une base de F_i pour $1 \leqslant i \leqslant k$ alors la famille concaténée $B_1 \ldots B_k$ est une base de E.*

Si les sous-espaces $F_1, \ldots, F_k$ donnent une décomposition en somme directe de E alors la projection sur le $j$ème facteur de la somme directe $E = \bigoplus_{i=1}^{k} F_i$ est l'application linéaire surjective $p_j : E \to F_j$ définie par
$$p_j(x) = x_j$$
où $x = \sum_{i=1}^{k} x_i$ est l'unique décomposition de x avec $x_i \in F_i$ pour $1 \leqslant i \leqslant k$. La composition de p_j avec l'application évidente $F_j \hookrightarrow F$ est un endomorphisme de E qui sera noté par le même symbole p_j. Rappelons l'identité importante $p_j^2 = p_j$.

2.2 Valeurs propres et espaces propres

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n.

Définition 2.3 *Un endomorphisme de E est une application linéaire $\phi : E \to E$. On pose*

$$\mathrm{End}(E) := \{\phi : E \to E |\ \phi \text{ est linéaire}\}$$

(l'espace des endomorphismes de E).

Rappelons que, si E est un K-espace vectoriel de dimension finie n, ϕ un endomorphisme de E, et $B = (f_1, \ldots, f_n)$ une base de E, alors on désigne par $M_B(\phi)$ la matrice de ϕ dans la base B. Si on pose $M_B(\phi) = (a_{ij})_{1 \leqslant i,j \leqslant n}$, alors les éléments a_{ij} de cette matrice sont donnés par les formules

$$\phi(f_j) = \sum_{i=1}^{n} a_{ij} f_i \ \forall j \{1, \ldots, n\}, \tag{6}$$

donc la j-me colonne de $M_B(\phi)$ est formée des coordonnées du vecteur $\phi(f_j)$ dans la base B. En utilisant les notations introduites dans le cours d'algèbre linéaire on peut écrire $A = M_{B,B}(\phi) = [\phi]_{B,B}$. Notons que, dans le cas des endomorphismes (le cas où l'espace de départ coïncide avec l'espace d'arrivé) utiliser deux bases différentes pour le calcul de la matrice associée serait une complication inutile. Les formules (6) s'écrivent d'une manière compacte :

$$\phi(B) = BM_B(\phi) \ .$$

Si on fait un changement de base $B \mapsto B' = (f'_1, \ldots, f'_n)$ on peut utiliser la matrice de passage P définie par les formules $f'_j = \sum_{i=1}^{n} p_{ij} f_i$ pour $j \leqslant j \leqslant n$. On peut écrire l'ensemble de ces formules sous la forme compacte $(f'_1, \ldots, f'_n) = (f_1, \ldots, f_n)P$, soit

$$B' = BP \ ,$$

où à droite on utilise formellement la règle de multiplication matricielle. La matrice $M_{B'}(\phi)$ de ϕ dans la "nouvelle" base B' est définie par la formule $\phi(B') = B'M_{B'}(\phi)$. En utilisant la linéarité de ϕ on obtient

$$\phi(B') = B'M_{B'}(\phi) \Rightarrow \phi(BP) = BPM_{B'}(\phi) \Rightarrow \phi(B)P = BPM_{B'}(\phi) \ .$$

Mais, puisque $\phi(B) = BM_B(\phi)$ on a aussi $\phi(B)P = BM_B(\phi)P$. On obtient donc

$$BM_B(\phi)P = BPM_{B'}(\phi) \ .$$

En utilisant l'unicité de la décomposition par rapport à une base, on remarque que la dernière égalité implique $M_B(\phi)P = M_{B'}(\phi)P$, c'est à dire

$$M_{B'}(\phi) = P^{-1}M_B(\phi)P \tag{7}$$

Cette formule justifie la définition importante suivante :

Définition 2.4 *Deux matrices carrées A, $A' \in M_{n,n}(\mathbb{K})$ sont dites semblables s'il existe une matrice $P \in M_{n,n}(\mathbb{K})$ inversible telle que $A' = P^{-1}AP$.*

Notons que la relation de similitude est un relation d'équivalence (c'est à dire c'est une relation reflexive, symétrique et transitive) sur l'espace $M_{n,n}(\mathbb{K})$. La formule (7) montre que

Remarque 2.5 *Les matrices $M_B(\phi)$, $M_{B'}(\phi)$ associées à un endomorphisme $\phi \in \mathrm{End}(E)$ dans deux bases B, B' de E sont semblables.*

Rappelons qu'une matrice carrée $A \in M_{n,n}(\mathbb{K})$ est dite *diagonale* si tous ses éléments non-diagonaux sont nuls, c'est dire si $a_{ij} = 0$ pour $i \neq j$.

Le problème de diagonalisation pour les endomorphismes : Soit $\phi \in \mathrm{End}(E)$. Est-ce qu'il existe une base B de E telle que la matrice associée à ϕ dans cette base (c'est à dire la matrice $M_B(\phi)$) soit diagonale ? Si c'est le cas déterminer une telle base.

Terminologie : Si $M_B(\phi)$ est diagonale, on dit que ϕ est diagonal dans la base B, ou que la base B diagonalise l'endomorphisme ϕ.

On peut commencer par écrire ϕ dans une base arbitraire (par exemple dans la base canonique si $E = \mathbb{K}^n$) et puis diagonaliser ϕ en effectuant un changement de base. En utilisant la formule (7) on voit que le problème de diagonalisation est équivalent au :

Problème de diagonalisation pour les matrices : Soit $A \in M_{n,n}(\mathbb{K})$ une matrice carrée réelle. Est-ce qu'il existe une matrice A' semblable à A qui est diagonale ? Si c'est le cas, déterminer une telle matrice et la matrice de passage P intervenant dans l'égalité $A' = P^{-1}AP$.

Définition 2.6 *Soit $\phi \in \mathrm{End}(E)$.*

1. *Un scalaire $\lambda \in \mathbb{K}$ s'appelle valeur propre de ϕ s'il existe $v \in E\backslash\{0_E\}$ tel que $\phi(v) = \lambda v$.*
2. *Le spectre de ϕ est défini par*

$$\mathrm{Spec}(\phi) := \{\lambda \in \mathbb{K} |\ \lambda \text{ est valeur propre de } \phi\} \ .$$

3. *Si $\lambda \in \mathrm{Spec}(\phi)$ on définit le sous-espace propre de ϕ associé à λ par*

$$E_\lambda(\phi) := \{v \in E |\ \phi(v) = \lambda v\} = \ker(\phi - \lambda \mathrm{id}_E) \ .$$

Les éléments non-nuls de $E_\lambda(\phi)$ s'appellent vecteurs propres de valeur propre λ.

Si ϕ est fixé on va utiliser la notation simplifié E_λ au lieu de $E_\lambda(\phi)$.

Remarque 2.7 1. *$E_\lambda(\phi)$ est un sous-espace vectoriel de E pour toute valeur propre $\lambda \in \mathrm{Spec}(\phi)$.*

2. *Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de ϕ si et seulement si $\ker(\phi - \lambda \mathrm{id}_E) \neq \{0_E\}$, donc un espace propre $E_\lambda(\phi)$ n'est jamais trivial.*

Donc tout sous-espace propre est de dimension $\geqslant 1$. Notons que la matrice $M_B(\phi)$ est diagonale si et seulement si les formules (6) se réduisent à

$$\phi(f_j) = a_{jj} f_j \ ,$$

c'est à dire si et seulement si f_j est un vecteur propre de valeur propre a_{jj}. On obtient donc la remarque suivante qui justifie l'importance des vecteurs propres pour le problème de diagonalisation :

Remarque 2.8 *Soit $\phi \in \mathrm{End}(E)$. Une base $B = (f_1, \ldots, f_n)$ de E diagonalise ϕ si et seulement si elle est formée de vecteurs propres de ϕ. Si c'est le cas, le j-ème élément diagonal de $M_B(\phi)$ est la valeur propre qui correspond à f_j.*

Le problème de diagonalisation pour ϕ se réduit donc au problème :

Est-ce qu'il existe une base de E formée de vecteurs propres de ϕ ? Si c'est la cas, déterminer une telle base.

Définition 2.9 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, soit $\phi \in \mathrm{End}(E)$ et soit A la matrice de ϕ dans une base B de E. Le polynôme caractéristique de ϕ est défini par $P_\phi(X) := \det(A - XI_n)$.*

Le degré de ce polynôme est $n = \dim(E)$. On peut préciser facilement trois coefficients de ce polynôme :

$$P_\phi(X) = (-1)^n X^n + (-1)^{n-1}\mathrm{Tr}(A)X^{n-1} + \cdots + \det(A) \ .$$

Remarque 2.10 *Le polynôme caractéristique de ϕ dépend seulement de ϕ (pas de la base choisie). En particulier, on peut définir d'une manière cohérente les applications*

$$\mathrm{Tr} : \mathrm{End}(E) \to \mathbb{K} \ , \ \det : \mathrm{End}(E) \to \mathbb{K}$$

en posant $\mathrm{Tr}(\phi) := \mathrm{Tr}(A)$, $\det(\phi) := \det(A)$ où $A := M_B(\phi)$.

Démonstration: Si on passe d'une base B à une nouvelle base B', la matrice A' associée à ϕ dans la nouvelle base sera donnée par $A' = P^{-1}AP$, où P est la matrice de passage de B à B'. On a

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

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

■

Le calcul du spectre et des sous-espaces propres d'un endomorphisme s'appuie sur la proposition suivante :

Proposition 2.11 *Soit Soit E un $\mathbb{K}$ espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$.*

1. *Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de ϕ si et seulement si il est une racine du polynôme caractéristique $P_\phi(X)$.*
2. *Soit A la matrice associée à ϕ dans une base B de E et soit $\lambda \in \mathrm{Spec}(\phi)$. L'espace propre $E_\lambda(\phi)$ s'identifie dans la base B à l'espace des solutions du système linéaire homogène*

$$(A - \lambda I_n)x = 0\ .$$

Démonstration: 1. La matrice associée à $\phi - \lambda \mathrm{id}_E$ dans la base B est $A - \lambda I_n$, donc le noyau $\ker(\phi - \lambda \mathrm{id}_E)$ est donné dans cette base par le système linéaire homogène $(A - \lambda I_n)x = 0$. D'après la Remarque 2.7 $\lambda \in \mathrm{Spec}(\phi)$ si et seulement si ce noyau est non-trivial, donc si et seulement si ce système admet une solution non-triviale, donc si et seulement si les colonnes de cette matrices sont linéairement dépendantes, i.e. le rang de la matrice $A - \lambda I_n$ est strictement inférieur à n. Mais

$$\mathrm{rang}(A - \lambda I_n) < n \Leftrightarrow \det(A - \lambda I_n) = 0 \Leftrightarrow P_\phi(\lambda) = 0\ .$$

2. Par définition l'espace propre $E_\lambda(\phi)$ associé à une valeur propre λ coïncide avec $\ker(\phi - \lambda \mathrm{id}_E)$, donc s'identifie dans la base B, à l'espace des solutions du système linéaire homogène $(A - \lambda I_n)x = 0$. ■

Définition 2.12 *Soient E un $\mathbb{K}$ espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$.*

1. *la multiplicité algébrique $\mathrm{m}_a(\lambda)$ de λ est la multiplicité de λ en tant que racine du polynôme $P_\phi(X) \in \mathbb{K}[X]$, i.e. (l'exposant de) la puissance maximale de $(X - \lambda)$ qui divise $P_\phi(X)$.*
2. *la multiplicité géométrique $\mathrm{m}_g(\lambda)$ est la dimension de l'espace propre E_λ.*

Proposition 2.13 *On a l'inégalité $1 \leqslant \mathrm{m}_g(\lambda) \leqslant \mathrm{m}_a(\lambda)$ pour tout $\lambda \in \mathrm{Spec}(\phi)$.*

Démonstration: La multiplicité algébrique $\mathrm{m}_a(\lambda)$ est (l'exposant de) la puissance maximale de $(X - \lambda)$ qui divise $P_\phi(X)$, donc pour démontrer l'inégalité $\mathrm{m}_g(\lambda) \leqslant \mathrm{m}_a(\lambda)$, il suffit de montrer que $(X - \lambda)^{\mathrm{m}_g(\lambda)}$ divise $P_\phi(X)$. Pour ceci choisissons une base B_0 de E_λ et complétons B_0 (qui est une famille libre de E) en une base $B = B_0B_1$ de E. En posant $d := \mathrm{m}_g(\lambda) = \dim(E_\lambda)$ on a :

$$M_B(\phi) = \begin{pmatrix} \lambda I_d & B \\ 0_{n-d,d} & C \end{pmatrix},\ M_B(\phi) - XI_n = \begin{pmatrix} \lambda I_d - XI_d & B \\ 0_{n-d,d} & C - XI_{n-d} \end{pmatrix}\ ,\ \text{où } B \in M_{d,n-d},\ C \in M_{n-d,d,n-d}\ .$$

En utilisant la règle de calcul du déterminant d'une matrice *triangulaire par blocs* on obtient :

$$P_\phi(X) = \begin{vmatrix} \lambda I_d - XI_d & B \\ 0_{n-d,d} & C - XI_{n-d} \end{vmatrix} = |\lambda I_d - XI_d||C - XI_{n-d}| = (-1)^d(X - \lambda)^d \det(C - XI_{n-d})\ ,$$

ce qui montre que $(X - \lambda)^{\mathrm{m}_g(\lambda)}$ est un diviseur de $P_\phi(X)$. ■

Corollaire 2.14 *Si $\mathrm{m}_a(\lambda) = 1$ (donc si λ est une racine simple de $P_\phi(X)$), alors $\mathrm{m}_g(\lambda) = \mathrm{m}_a(\lambda) = 1$, donc le sous-espace propre E_λ est une droite vectorielle.*

Proposition 2.15 *Soient $\lambda_1, \ldots, \lambda_k \in \mathrm{Spec}(\phi)$ valeurs propres distinctes deux à deux. Alors les sous-espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_k}$ sont en somme directe.*

Démonstration: On va démontrer cette affirmation par récurrence par rapport à k. Pour $k = 1$ l'affirmation est évidente. Soit $k > 1$ et supposons que (l'hypothèse de récurrence) toute famille de $(k-1)$ sous-espaces propres de ϕ est en somme directe.

D'après la définition 2.1 pour montrer que $E_{\lambda_1}, \ldots, E_{\lambda_k}$ sont en somme directe il suffit de montrer que, pour tout $j \in \{1, \ldots, k\}$, on a

$$E_{\lambda_j} \cap \left(\sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} E_{\lambda_i} \right) = \{0_E\} \ .$$

Soit $x \in E_{\lambda_j} \cap \left(\sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} E_{\lambda_i} \right)$. On a donc $x \in E_{\lambda_j}$ et une décomposition de la forme

$$x = \sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} x_i \ . \tag{8}$$

En appliquant ϕ on obtient

$$\lambda_j x = \sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} \lambda_i x_i \ . \tag{9}$$

Si on multiplie (8) par $-\lambda_j$ et on ajoute (9) on obtient

$$\sum_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} (\lambda_i - \lambda_j) x_i = 0_E \ .$$

Puisque, par l'hypothèse de récurrence, les sous-espaces $(E_{\lambda_i})_{\substack{1 \leqslant i \leqslant k \\ i \neq j}}$ sont en somme directe, cette relation implique (voir la définition 2.1)

$$(\lambda_i - \lambda_j) x_i = 0_E \ \forall i \in \{1, \ldots k\} \backslash \{j\} \ ,$$

donc (puisque $\lambda_i \neq \lambda_j$) on obtient

$$x_i = 0_E \ \forall i \in \{1, \ldots k\} \backslash \{j\} \ .$$

En revenant à (8) on obtient $x = 0_E$. ■

2.3 Critères de diagonalisabilité

Le théorème suivant nous donne quatre critères de diagonalisabilité :

Théorème 2.16 *Soit E un $\mathbb{K}$ espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$. Alors les propriétés suivantes sont équivalentes :*

1. *ϕ est diagonalisable,*
2. *E admet une base formée de vecteurs propres de ϕ,*
3. *Les sous-espaces propres E_λ donnent une décomposition en somme directe de E, i.e.*

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} E_\lambda \ .$$

4. *$\sum_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{m}_g(\lambda) = n$,*
5. *$P_\phi(X)$ est scindé dans $\mathbb{K}$ et pout toute valeur propre $\lambda \in \mathrm{Spec}(\phi)$ on a $\mathrm{m}_g(\lambda) = \mathrm{m}_a(\lambda)$.*

Démonstration: L'équivalence (1)$\Leftrightarrow$(2) est évidente. En effet, soit $B = (f_1, \dots, f_n)$ une base de E et $A := M_B(\phi)$ la matrice de ϕ dans la base B. On a donc

$$\phi(f_j) = \sum_{i=1} a_{ij} f_j \ .$$

Donc $M_B(\phi)$ est diagonale si et seulement si ces formules se réduisent à

$$\phi(f_j) = a_{jj} f_j \ ,$$

donc si et seulement si pour tout $j \in \{1, \dots, n\}$ f_j est vecteur propre de valeur propre a_{jj}.

2.$\Rightarrow$3. Soit B une base de E formée de vecteurs propres de ϕ. En regroupant ces vecteurs propres si nécessaire on peut supposer que $B = B_1 B_2 \dots B_l$ où B_i est formée de (disons) μ_i vecteurs propres associés à la valeur propre $\lambda_i \in \mathrm{Spec}(\phi)$ et $(\lambda_1, \dots, \lambda_l)$ sont distinctes deux à deux. La matrice de ϕ dans la base B s'écrit

$$M_B(\phi) = \begin{pmatrix} \lambda_1 I_{\mu_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{\mu_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_l I_{\mu_l} \end{pmatrix} ,$$

donc

$$P_\phi(X) = \det(M_B(\phi) - X I_n) = \Pi_{i=1}^{l} (\lambda_i - X)^{\mu_i} \ . \tag{10}$$

D'autre part, puisque B_i est une famille libre à μ_i éléments dans E_{λ_i} on obtient pour $1 \leqslant i \leqslant l$ les inégalités $1 \leqslant \mu_i \leqslant \mathrm{m}_g(\lambda_i)$, donc en tenant compte de la proposition 2.13 on a :

$$1 \leqslant \mu_i \leqslant \mathrm{m}_g(\lambda_i) \leqslant \mathrm{m}_a(\lambda_i) \ . \tag{11}$$

En comparant les degrés à gauche et à droite dans (10) et en additionnant les inégalités (11) on obtient

$$\deg(P_\phi(X)) = \sum_{i=1}^{l} \mu_i \leqslant \sum_{i=1}^{l} \mathrm{m}_g(\lambda_i) \leqslant \sum_{i=1}^{l} \mathrm{m}_a(\lambda_i) \leqslant \deg(P_\phi(X)) \ . \tag{12}$$

La dernière égalité dans (12) est obtenue en tenant compte que $(X - \lambda_i)^{\mathrm{m}_a(\lambda_i)}$ divise $P_\phi(X)$ pour $1 \leqslant i \leqslant l$ (par la définition de $\mathrm{m}_a(\lambda_i)$) et que $\lambda_1, \dots, \lambda_l$ étant distinctes deux à deux, on aura aussi

$$\prod_{i=1}^{l} (X - \lambda_i)^{\mathrm{m}_a(\lambda_i)} | P_\phi(X) \ .$$

Remarquer que la dernière inégalité dans (12) est une égalité si et seulement si $P_\phi(X)$ est scindé dans $\mathbb{K}$ et $\mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_l\}$.

Mais toutes les inégalités dans (12) sont des égalités, ce qui implique que toues les inégalités (11) sont aussi des égalités. Nous avons donc démontré que $\mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_l\}$ et que B_i est une base de E_{λ_i}. Nous savons que, en concanténant les bases B_i de E_{λ_i} pour $1 \leqslant i \leqslant l$, on obtient une base de E. D'après la définition 2.2 ceci implique

$$E = \bigoplus_{i=1}^{l} E_{\lambda_i} = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} E_\lambda \ .$$

3.$\Rightarrow$4. Puisque la dimension est additive par rapport aux sommes directes, 3. implique

$$n = \dim(E) = \sum_{\lambda \in \mathrm{Spec}(\phi)} \dim(E_\lambda) = \sum_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{m}_g(\lambda) \ .$$

4.$\Rightarrow$5. Puisque $\mathrm{m}_g(\lambda) \leqslant \mathrm{m}_a(\lambda)$ pour tout $\lambda \in \mathrm{Spec}(\phi)$ et $\Pi_{\lambda \in \mathrm{Spec}(\phi)} (X - \lambda)^{\mathrm{m}_a(\lambda)} | P_\phi(X)$ on a

$$\sum_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{m}_g(\lambda) \leqslant \sum_{\lambda \in \mathrm{Spec}(\phi)} \mathrm{m}_a(\lambda) \leqslant n.$$

D'après l'hypothèse (la propriété 4.) ces inégalités sont des égalités, ce qui implique $\mathrm{m}_g(\lambda) = \mathrm{m}_a(\lambda)$ pour tout $\lambda \in \mathrm{Spec}(\phi)$ et $\deg\left(\Pi_{\lambda\in\mathrm{Spec}(\phi)}(X-\lambda)^{\mathrm{m}_a(\lambda)}\right) = \deg(P_\phi(X))$, donc $P_\phi(X)$ est scindé dans $\mathbb{K}$.

5.⇒2. Posons $\mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_k\}$ et soit B_i une base de E_{λ_i}. L'hypothèse (la propriété 5) implique

$$n = \deg(P_\phi(X)) = \sum_{i=1} \mathrm{m}_a(\lambda_i) = \sum_{i=1} \mathrm{m}_g(\lambda_i) \ . \tag{13}$$

Puisque les sous-espaces E_{λ_i} sont en somme directe, la famille concaténée $B := B_1 B_2 \dots B_k$ sera libre. Puisque le cardinal de cette famille est $n = \dim(E)$, d'après (13) on déduit que B est une base de E. Cette base est formée de vecteurs propres évidemment. ■

Question : Comment on diagonalise un endomorphisme ?
Réponse : Il suffit de trouver une base B_λ dans E_λ pour tout $\lambda \in \mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_k\}$. Si l'endomorphisme ϕ est diagonalisable, alors la famille concaténée $B = B_{\lambda_1} \dots B_{\lambda_k}$ sera une base de E (formée évidemment de vecteurs propres). Dans les exercices de diagonalisation on suppose d'habitude que $E = \mathbb{K}^n$ et ϕ est donné dans la base canonique par une matrice A. Dans cette situation, la matrice de passage P de la base canonique de $\mathbb{K}^n$ à la base B qui diagonalise ϕ est formée des vecteurs de B en tant que colonnes. On aura alors

$$P^{-1}AP = \begin{pmatrix} \lambda_1 I_{d_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k I_{d_k} \end{pmatrix}$$

où $d_i = \mathrm{m}_a(\lambda_i) = \mathrm{m}_g(\lambda_i)$.

2.4 Trigonalisation

Définition 2.17 *Une matrice carrée $A \in M_{n,n}(\mathbb{K})$ est dite triangulaire supérieure (respectivement inférieure) si $a_{ij} = 0$ pour toute paire d'indices (i,j) telle que $i > j$ (respectivement pour toute paire d'indices (i,j) telle que $i < j$).*

Remarque 2.18 *Le déterminant d'une matrice triangulaire supérieure (ou inférieure) coïncide avec le produit des éléments diagonaux.*

On peut démontrer facilement cette remarque soit par récurrence (en utilisant le développement du déterminant par rapport à la première ligne), soit en utilisant la définition du déterminant et en remarquant que la seule permutation $\sigma \in \mathrm{S}_n$ ayant la propriété $\sigma_i \leqslant i$ pour tout $i \in \{1, \dots, n\}$ est la permutation identique.

Définition 2.19 *Un endomorphisme $\phi \in \mathrm{End}(E)$ est dit trigonalisable s'il existe une base B de E telle que la matrice $A := M_B(\phi)$ soit triangulaire supérieure.*

Interprétation géométrique : Soient $\phi \in \mathrm{End}(E)$ et $B = (f_1, \dots, f_n)$ une base de E. Pour tout $i \in \{1, \dots, n\}$ posons :

$$E_i := \mathrm{Vect}(f_1, \dots, f_i) \ ,$$

donc E_i est le sous-espace engendré par les premiers i vecteurs de la base B. Alors

Remarque 2.20 *La matrice $M_B(\phi)$ de ϕ dans la base B est triangulaire supérieure si et seulement si on a $\phi(E_i) \subset \phi(E_i)$ pour tout $i \in \{1, \dots, n\}$, donc si et seulement si les sous-espaces E_i sont ϕ-invariants.*

En utilisant cette remarque et le théorème de la base incomplète, on peut montrer facilement que ϕ est trigonalisable si et seulement s'il existe une famille croissante $E_1 \subset E_2 \subset \cdots \subset E_n = E$ de sous-espaces vectoriels tels que pour $1 \leqslant i \leqslant n$ on a $\dim(E_i) = i$ et $\phi(E_i) \subset E_i$.

Si, dans la définition 2.19 on remplace "triangulaire supérieure" par "triangulaire inférieure" on obtient une définition équivalente. Pourquoi ?

Théorème 2.21 *Soient E un $\mathbb{K}$ espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$. Les propriétés suivantes sont équivalentes :*

1. *ϕ est trigonalisable,*
2. *$P_\phi(X)$ est scindé dans $\mathbb{K}$.*

Démonstration: (1)⇒(2) Supposons que ϕ est trigonalisable, et calculons sont polynôme caractéristique en utilisant une base B dans laquelle sa matrice $A := M_B(\phi)$ est triangulaire supérieure . Alors

$$P_\phi(X) = \det(A - X\mathrm{I}_n) = \Pi_{i=1}^n (a_{ii} - X) \ ,$$

donc $P_\phi(X)$ est scindé dans $\mathbb{K}$.

(2)⇒(1) : Récurrence par rapport à $\dim(E)$. Si $\dim(E) = 1$ l'affirmation est évidente. Supposons que $\dim(E) \geqslant 2$ et l'affirmation est vraie pour tout espace vectoriel F de dimension $n-1$. Puisque $P_\phi(X)$ est scindé dans $\mathbb{K}$ on a $\mathrm{Spec}(\phi) \neq \emptyset$, donc on peut choisir une valeur propre $\lambda \in \mathrm{Spec}(\phi)$. Soit $v_1 \in E_\lambda \backslash \{0_E\}$ un vecteur propre pour cette valeur propre, et soit $F \subset R$ un supplémentaire de la droite vectorielle $\mathbb{K}v_1$. On peut obtenir pratiquement un tel supplémentaire en appliquant le théorème de la base incomplète à la famille libre (v_1). Ce théorème fournit une base $B = (v_1, v_2, \ldots, v_n)$ de E et on pose $F := \mathrm{Vect}(v_2, \ldots v_n)$. Désignons par p et p' les projections de E sur les deux termes de la somme directe $E = \mathbb{K}v_1 \oplus F$ ainsi obtenue. Puisque le premier terme est une droite vectorielle on peut écrire $p' = fv_1$, où $f : E \to \mathbb{K}$ est une forme linéaire sur E. Le premier vecteur de B est un vecteur propre de ϕ de valeur propre λ, donc la matrice A de ϕ dans la base B aura la forme

$$A = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \lambda & a \\ 0 & A' \end{pmatrix} \quad \text{avec } a = (a_{12} \cdots a_{1n}) \in M_{1,n-1} \ , A' \in M_{n-1,n-1} \ .$$

On a donc $a_{11} = \lambda$ et $a_{i1} = 0$ pour $i > 1$.

Remarquer que la première ligne $(\lambda \ a_{12} \cdots a_{1n})$ de A est formée des coefficients de la forme linéaire f dans la base B et la matrice

$$A' := \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} \in M_{n-1,n-1}(\mathbb{K})$$

est la matrice de l'endomorphisme $\phi' := p' \circ \phi\big|_F \in \mathrm{End}(F)$ dans la base $(v_2, \ldots, v_n)$. On a donc

$$P_\phi(X) = (\lambda - X)\det(A' - X\mathrm{I}_{n-1}) = (\lambda - X)P_{\phi'}(X) \ .$$

Puisque $P_\phi(X)$ est scindé dans $\mathbb{K}$ par l'hypothèse, on en déduit que $P_{\phi'}(X)$ est aussi scindé sur $\mathbb{K}$ (parce que tout diviseur irréductible de degré $\geqslant 2$ de $P_\phi(X)$ serait aussi un diviseur irréductible de $P_{\phi'}(X)$). Donc, par l'hypothèse de récurrence, ϕ' est trigonalisable. Soit B' une base de F dans laquelle la matrice de ϕ' est triangulaire supérieure. La matrice $\tilde{A}$ de ϕ dans la base $\tilde{B} := (v_1, B')$ de E aura la forme

$$\tilde{A} = \begin{pmatrix} \lambda & \tilde{a} \\ 0 & M_{B'}(\phi') \end{pmatrix} \quad \text{avec } \tilde{a} \in M_{1,n-1} \ , \ M_{B'}(\phi') \in M_{n-1,n-1},$$

donc est triangulaire supérieure, parce que $M_{B'}(\phi')$ est triangulaire supérieure. ■

Rappelons que, d'après le théorème de Gauss-d'Alembert, tout polynôme à coefficients complexes est scindé dans $\mathbb{C}$. Pour $\mathbb{K} = \mathbb{C}$ on a donc :

Corollaire 2.22 *Si E est un espace vectoriel complexe de dimension finie n, alors tout endomorphisme $\phi \in \mathrm{End}(E)$ est trigonalisable.*

3 Polynôme d'un endomorphisme. Idéal annulateur. Polynôme minimal. Le lemme des noyaux

3.1 Polynôme d'un endomorphisme. Idéal annulateur. Polynôme minimal

On va adopter la convention standard pour les cours de licence : un anneau est un anneau avec unité.

Rappel sur les anneaux : Soit $(A, +, \cdot)$ un anneau commutatif. On désigne par 0 et 1 les éléments neutres pour les deux opérations.

Définition 3.1 *Un idéal de A est un sous-groupe du groupe $(A, +)$ tel que $aI \subset I$ pour tout $a \in A$.*

Pour tout $a \in A$ le sous-ensemble $(a) := Aa := \{xa|\ x \in A\}$ est un idéal de A, qui s'appelle l'idéal principal engendré par a. Un idéal I de cette forme s'appelle idéal principal, et un élément $a \in I$ pour lequel $I = Aa$ s'appelle générateur de I.

Définition 3.2 *Un anneau commutatif avec unité $(A, +, \cdot)$ est dit anneau principal si tout idéal de A est un idéal principal.*

Exemple : $(\mathbb{Z}, +, \cdot)$ est un anneau principal, parce que tout idéal I de $\mathbb{Z}$ s'écrit sous la forme $I = n\mathbb{Z}$, où $n \in \mathbb{N}$. Pour obtenir un générateur positif d'un idéal *non-nul* $I \subset \mathbb{Z}$ il suffit de choisir un élément strictement positif minimal de I.

Théorème 3.3 *Soit K un corps. L'anneau $K[X]$ des polynômes à coefficients dans K est principal.*

Démonstration: Soit $I \subset K[X]$ un idéal. On peut supposer $I \neq \{0\}$. Soit $P(X) \in I\backslash\{0\}$ tel que

$$\deg(P(X)) = \min\{\deg(S(X))|\ S(X) \in I\backslash\{0\}\}.$$

Nous affirmons que $(P(X)) = I$, donc $P(X)$ est un générateur de I. L'inclusion $(P(X)) \subset I$ est évidente. Pourquoi ?

Pour l'inclusion $I \subset (P(X))$: soit $S(X) \in I$. En appliquant le théorème de division euclidienne à $(S(X), P(X))$ on obtient

$$S(X) = Q(X)P(X) + R(X) \text{ avec } \deg(R(X)) < \deg(P(X).$$

Nous allons démontrer que $R(X) = 0$, donc $S(X) \in (P(X))$ et l'inclusion $I \subset (P(X))$ sera démontrée. Nous avons $R(X) = S(X) - Q(X)P(X) \in I$, donc si par l'absurde $R(X) \neq 0$, nous obtenons un élément de $I\backslash\{0\}$ de degré strictement inférieur à $\deg(P(X))$, ce qui contredit la définition de $P(X)$. ■

Remarque 3.4 *1. Si $(A, +, \cdot)$ un anneau commutatif, alors l'anneau $A[X]$ des polynômes à coefficients dans A est principal si et seulement si A est un corps. Donc $\mathbb{Z}[X]$ n'est pas un anneau principal, mais l'anneau $\mathbb{K}[X]$ (qui intervient dans la théorie des endomorphismes) est bien principal.*

2. En utilisant la même méthode de démonstration on obtient un théorème plus général : Tout anneau euclidien est principal.

Une question naturelle se pose :

Question : Comment déterminer un générateur d'un idéal *non-nul* $I \subset \mathbb{K}[X]$?

La réponse est très simple :

Réponse : D'après notre démonstration, il suffit de choisir un élément *non-nul* $Q(X) \in I$ de degré minimal.

Remarque 3.5 *Si $P_1(X)$, $P_2(X)$ sont deux générateurs d'un idéal non-nul $I \subset \mathbb{K}[X]$, alors il existe une constante $a \in \mathbb{K}^*$ tel que $P_2(X) = aP_1(X)$. Ceci montre que I possède un unique générateur qui est un polynôme unitaire, c'est à dire un polynôme dont le coefficient du terme de plus haut degré est 1.*

Définition 3.6 *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$ et*

$$P(X) = a_0 + a_1 X + \cdots + a_m X^m = \sum_{i=0}^{m} a_i X^i \in \mathbb{K}[X].$$

Posons

$$P(\phi) = a_0 \mathrm{id}_E + a_1 \phi + \cdots + a_m \phi^m = \sum_{i=0}^{m} \phi^i \in \mathrm{End}(E) \ ,$$

où $\phi^0 := \mathrm{id}_E$, $\phi^1 := \phi$ et $\phi^i := \underbrace{\phi \circ \phi \circ \cdots \circ \phi}_{i \text{ fois}}$ pour $i \geqslant 2$.

Rappelons que $\mathrm{End}(E)$, muni de l'addition usuelle et la composition des endomorphismes, est un anneau qui est non-commutatif si $n > 1$.

Remarque 3.7 *Soit $\phi \in \mathrm{End}(E)$. L'application $\Gamma_\phi : \mathbb{K}[X] \to \mathrm{End}(E)$ donnée par $P(X) \mapsto P(\phi)$ est un morphisme d'anneaux. En particulier pour tous deux polynômes $P_1(X)$, $P_2(X) \in \mathbb{K}[X]$ on a l'identité*

$$(P_1 P_2)(\phi) = P_1(\phi) \circ P_2(\phi).$$

Puisque $\mathbb{K}[X]$ est commutatif il en résulté $P_1(\phi) \circ P_2(\phi) = P_2(\phi) \circ P_1(\phi)$, donc deux endomorphismes obtenus comme polynômes du même endomorphisme ϕ commutent.

La démonstration est proposée comme exercice. Rappelons que la composition des endomorphismes (comme la multiplication matricielle) n'est pas commutative donc, en général, deux endomorphismes ne commutent pas, ce qui montre l'importance de la deuxième partie de cette remarque.

Corollaire 3.8 *Soit $\phi \in \mathrm{End}(E)$ et $Q(X) \in \mathbb{K}[X]$. Alors les sous-espaces $\ker(Q(\phi))$ et $\mathrm{im}(Q(\phi))$ de E sont ϕ-invariants, i.e. on a $\phi(\ker(Q(\phi))) \subset \ker(Q(\phi))$, $\phi(\mathrm{im}(Q(\phi))) \subset \mathrm{im}(Q(\phi))$.*

Démonstration: En effet, d'après la remarque 3.7 pour tout $v \in E$ on a

$$Q(\phi)(\phi(v)) = (Q(\phi) \circ \phi)(v) = (\phi \circ Q(\phi))(v) = \phi(Q(\phi)(v)) \ ,$$

ce qui démontre les implications désirées : Premièrement, si $v \in \ker(Q(\phi))$ alors le dernier terme dans l'égalité s'annule, donc le premier terme s'annule, donc $\phi(v) \in \ker(Q(\phi))$. D'autre part, si $w \in \mathrm{im}(Q(\phi))$ alors w s'écrit sous la forme $w = Q(\phi)(v)$ et l'égalité ci-dessus montre que $\phi(w) = Q(\phi)(\phi(v)) \in \mathrm{im}(Q(\phi))$. ■

Remarque 3.9 *Si $\lambda \in \mathrm{Spec}(\phi)$ alors $Q(\lambda) \in \mathrm{Spec}(Q(\phi))$. En particulier, si $\lambda \in \mathrm{Spec}(\phi)$ et $Q(\phi) = 0_{\mathrm{End}(E)}$, alors λ est nécessairement une racine du polynôme $Q(X)$.*

Démonstration: En effet, soit $v \in E \backslash \{0_E\}$ tel que $\phi(v) = \lambda v$. Alors par récurrence on obtient facilement $\phi^i(v) = \lambda^i v$ pour tout $i \in \mathbb{N}$, ce qui implique $Q(\phi)(v) = Q(\lambda)v$. ■

Théorème 3.10 *(Théorème de Cayley-Hamilton) Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$. Alors $P_\phi(\phi) = 0_{\mathrm{End}(E)}$.*

Démonstration: Le cas réel se réduit au cas complexe. En effet, dans le cas réel on peut supposer que $E = \mathbb{R}^n$ et alors ϕ est donné par une matrice $A \in M_{n,n}(\mathbb{R})$. Mais, regardée comme matrice complexe, A définit aussi un endormorphisme de $\mathbb{C}^n$ qui a le même polynôme caractéristique que A (donc que ϕ). Si le théorème est vrai dans le cas complexe alors on obtient $P_\phi(A) = 0_{n,n}$, ce qui implique évidemment $P_\phi(\phi) = 0_{\mathrm{End}(E)}$.

On va supposer donc que $\mathbb{K} = \mathbb{C}$. Soit $B = (f_1, \ldots, f_n)$ une base E qui trigonalise ϕ (voir le corollaire 2.22), donc telle que la matrice $A := M_B(\phi)$ soit triangulaire supérieure. Posons $\lambda_i = a_{ii}$ et notons que $P_\phi(X) = (-1)^n \Pi_{i=1}^n (X - \lambda_i)$. Posons $Q_j(X) := (X - \lambda_j)$. Remarquer que $Q_j(\phi) = \phi - \lambda_j \mathrm{id}_E$ et que la matrice de cet endomorphisme dans la base B est encore triangulaire supérieure. A remarquer aussi que le j-me élément diagonal de la matrice de $Q_j(\phi)$ dans la base B est 0, donc $Q_j(\phi)(f_j)$ s'écrit comme combinaison linéaire de $(f_1, \ldots, f_{j-1})$ pour $j \geqslant 2$ et $Q_1(\phi)(f_1) = 0_E$. En utilisant ces remarques on peut montrer facilement les propriétés suivantes (exercice) :

1. $Q_j(\phi)(f_i) \in \mathrm{Vect}(f_1, \ldots, f_i)$ pour $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$.
2. $Q_j(\phi)(f_j) \in \mathrm{Vect}(f_1, \ldots, f_{j-1})$ pour $1 \leqslant j \leqslant n$, où pour $j = 1$, on pose $\mathrm{Vect}(f_1, \ldots, f_{j-1}) := \{0_E\}$.
3. $Q_j(\phi)(f_i) \in \mathrm{Vect}(f_1, \ldots, f_{j-1})$ pour $1 \leqslant i \leqslant j \leqslant n$.

Ces relations ont des interprétations géométriques : si on pose $E_0 := \{0_E\}$, $E_j := \mathrm{Vect}(f_1, \ldots, f_j)$, alors les relations (1) sont équivalentes aux inclusions $Q_j(\phi)(E_i) \subset E_i$ (i.e. chaque sous-espace E_i est $Q_j(\phi)$-invariant) et les relations (3) sont équivalentes aux inclusions $Q_j(\phi)(E_j) \subset E_{j-1}$. Mais alors

$$Q_1(\phi)(E_1) = \{0_E\},\ (Q_1Q_2)(\phi)(E_2) = Q_1(\phi)(Q_2(\phi)(E_2)) \subset Q_1(\phi)(E_1) = \{0_E\}$$

et par récurrence on obtient $(\prod_{i=1}^{j} Q_i)(\phi)(E_j) \subset \{0_E\}$, soit

$$(\prod_{i=1}^{j} Q_i)(\phi)\big|_{E_j} = 0 \ \text{ pour } 1 \leqslant j \leqslant n. \tag{14}$$

Mais $E_n = E$ et $\prod_{i=1}^{n} Q_i(X) = (-1)^n P_\phi(X)$, donc en choisissant $j = n$ dans (14) on obtient $P_\phi(\phi) = 0$. ■

Définition 3.11 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$. On polynôme $Q(X) \in \mathbb{K}[X]$ est dit polynôme annulateur de ϕ si $Q(\phi) = 0_{\mathrm{End}(E)}$. L'ensemble des polynômes annulateurs de ϕ est un idéal non-nul de $\mathbb{K}[X]$, qui s'appelle l'idéal annulateur de ϕ et sera noté par $\mathrm{Ann}(\phi)$ (ou $\mathcal{I}(\phi)$). L'unique générateur unitaire de $\mathrm{Ann}(\phi)$ est un polynôme nun-nul qui s'appelle le polynôme minimal de ϕ et sera désigné par $m_\phi(X)$.*

Le fait que $\mathrm{Ann}(\phi)$ est un idéal de $\mathbb{K}[X]$ est une conséquence directe de la remarque 3.7. En effet cette remarque montre que $\mathrm{Ann}(\phi)$ est le noyau d'un morphisme d'anneaux, et c'est facile de voir que le noyau de tout morphisme d'anneaux est un idéal. Le fait que cet idéal est *non-nul* résulte directement du théorème de Cayley-Hamilton. Retenir que, par définition, $m_\phi(X)$ divise tout polynôme annulateur de ϕ.

Exemple : Si $\varphi = 0$ alors $\mathrm{Ann}(\varphi)$ est l'idéal $(X) = \{P(X) |\ P(0) = 0\}$ engendré par X. Il s'agit donc de l'idéal des polynômes de la forme $\sum_{i=1}^{d} a_i X^i$ (des polynômes sans terme de degré 0).

La remarque 3.9 montre que

Proposition 3.12 *Soit $Q(X) \in \mathrm{Ann}(\phi)$. Alors pour toute $\lambda \in \mathrm{Spec}(\phi)$ on a $Q(\lambda) = 0$.*

Donc *toute valeur propre de ϕ est racine de tout polynôme annulateur $Q(X)$*. Il en résulté que pour toute $\lambda \in \mathrm{Spec}(\phi)$ on a $(X - \lambda)|Q(X)$, donc si on pose $\mathrm{Spec}(\phi) = \{\lambda_1, \ldots, \lambda_k\}$ (avec $\lambda_i \neq \lambda_j$ pour $i \neq j$) alors $\prod_{i=1}^{k}(X - \lambda_i)|Q(X)$. C'est important de remarquer que

Remarque 3.13 *Si $\mathbb{K} = \mathbb{R}$, alors toute racine* `complexe` *de $P_\phi(X) \in \mathbb{R}[X]$ est aussi racine* `complexe` *de tout polynôme annulateur $Q(X) \in \mathbb{R}(X)$ de ϕ.*

Cette remarque résulte facilement en supposant $E = \mathbb{R}^n$ et en appliquant la proposition précédente à l'endomorphisme de $\mathbb{C}^n$ définit par la matrice de ϕ dans la base canonique.

Exemple : Si E un $\mathbb{R}$-espace vectoriel de dimension finie n, et $\phi \in \mathrm{End}(E)$ tel que $(X^2 + 1)$ divise $P_\phi(X)$ alors nécessairement $(X^2 + 1)$ divise aussi tout polynôme annulateur $Q(X) \in \mathbb{R}(X)$ (en particulier le polynôme minimal $m_\phi(X)$). En effet, l'hypothèse montre que i et $-i$ sont racines complexes de $P_\phi(X)$, donc elles seront aussi racines complexes de $Q(X)$ d'après la remarque ci-dessus. Donc $(X + i)(X - i) = X^2 + 1$ divise $Q(X)$.

Pour le polynôme minimal nous avons une information plus précise :

Proposition 3.14 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$. Alors les polynômes $P_\phi(X)$ et $m_\phi(X)$ ont les mêmes racines complexes.*

Démonstration: Nous savons déjà que, puisque $m_\phi(X) \in \mathrm{Ann}(\phi)$, toute racine complexe de $P_\phi(X)$ est aussi racine complexe de $m_\phi(X)$. D'autre part, d'après le théorème de Cayley-Hamilton $P_\phi(X) \in \mathrm{Ann}(\phi) = (m_\phi(X))$, donc $m_\phi(X)$ divise $P_\phi(X)$. Ceci montre que, réciproquement, toute racine complexe de $m_\phi(X)$ est aussi racine complexe de $P_\phi(X)$. ■

Corollaire 3.15 *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n, et $\phi \in \mathrm{End}(E)$. Alors P_ϕ est scindé dans $\mathbb{R}$ si et seulement si m_ϕ est scindé dans $\mathbb{R}$.*

Démonstration: En effet, un polynôme a coefficients dans $\mathbb{R}$ est scindé dans $\mathbb{R}$ si et seulement si toutes ses racines dans $\mathbb{C}$ sont réelles. Il suffit d'appliquer la proposition 3.14. ■

Corollaire 3.16 *Supposons que $P_\phi(X)$ est scindé dans $\mathbb{K}$, donc il s'écrit sous la forme*

$$P_\phi(X) = (-1)^n \Pi_{\lambda \in \mathrm{Spec}(\phi)} (X - \lambda)^{\mathrm{m}_a(\lambda)} \ .$$

Alors

$$m_\phi(X) = \Pi_{\lambda \in \mathrm{Spec}(\phi)} (X - \lambda)^{\beta(\lambda)} \ ,$$

où $1 \leqslant \beta(\lambda) \leqslant \mathrm{m}_a(\lambda)$ pour toute $\lambda \in \mathrm{Spec}(\phi)$.

3.2 Le lemme des noyaux. Sous-espaces caractéristiques

3.2.1 Le lemme des noyaux

Pour expliquer l'importance du lemme des noyaux qu'on va énoncer et démontrer dans ce chapitre, on va commencer par un lemme très simple :

Lemme 3.17 *Soient K un corps, E un K-espace vectoriel de dimension finie, $E = \bigoplus_{i=1}^k F_i$ une décomposition en somme directe de sous-espaces, et ϕ un endomorphisme de E tel que chaque sous-espace F_i soit ϕ-invariant (i.e. $\phi(F_i) \subset F_i$). Désignons par $\phi_i \in \mathrm{End}(F_i)$ l'endomorphisme de F_i induit par ϕ. Alors*

1. *En choisissant une base B_i de F_i pour $1 \leqslant i \leqslant k$, la matrice de ϕ dans la base concaténée $B = B_1 \dots B_k$ est*
$$M_B(\phi) = \begin{pmatrix} M_{B_1}(\phi_1) & 0 & \cdots & 0 \\ 0 & M_{B_2}(\phi_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_{B_k}(\phi_k) \end{pmatrix}.$$
2. *On a*
$$P_\phi(X) = \prod_{i=1}^k P_{\phi_i}(X).$$

Démonstration: La première affirmation est évidente. Pour la deuxième affirmation posons $d_i := \dim(F_i)$. En utilisant la règle de calcul pour le déterminant d'un matrice diagonale par blocs, on obtient

$$P_\phi(X) = \begin{vmatrix} M_{B_1}(\phi_1) - XI_{d_1} & 0 & \cdots & 0 \\ 0 & M_{B_2}(\phi_2) - XI_{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_{B_k}(\phi_k) - XI_{d_k} \end{vmatrix} = \prod_{i=1}^k |M_{B_i}(\phi_i) - XI_{d_i}| =$$

$$= \prod_{i=1}^k P_{\phi_i}(X) \ .$$

■

Si les hypothèses du lemme 3.17 sont vérifiées, on va écrire $\phi = \oplus_{i=1}^k \phi_i$ où $\phi = \phi_1 \oplus \cdots \oplus \phi_k$. Par exemple

Exemple : Si ϕ est diagonalisable, alors ϕ se décompose en somme directe d'homothéties linéaires, plus précisément on a

$$\phi = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} \lambda \mathrm{id}_{E_\lambda} \ ,$$

donc dans ce cas, le lemme 3.17 fournit :

1. une base $B = B_1 \dots B_k$ (où B_i est une base de E_{λ_i}) dans laquelle la matrice de ϕ s'écrit sous la forme diagonale connue

$$\begin{pmatrix} \lambda_1 I_{d_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k I_{d_k} \end{pmatrix} \text{ où } d_i = \mathrm{m}_g(\lambda_i) \ .$$

2. L'égalité $P_\phi(X) = \prod_{j=1}^k (\lambda_j - X)^{d_j}$.

Le lemme 3.17 nous dit que, en général, pour comprendre la structure d'un endomorphisme $\phi \in \mathrm{End}(E)$, c'est très utile d'avoir une décomposition de E en somme directe de sous-espaces ϕ-invariants. Le lemme des noyaux ci-dessous nous donne une méthode générale pour obtenir une telle décomposition :

Théorème 3.18 *(le lemme des noyaux). Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$. Soient $Q_1(X), \dots, Q_k(X) \in \mathbb{K}[X]$ tels que*

H1. Les polynômes $Q_1(X), \dots, Q_k(X)$ sont premiers entre eux `deux à deux`*, c'est à dire* $\mathrm{pgcd}(Q_i, Q_j) = 1$ *pour $i \neq j$.*

H2. $\Pi_{j=1}^k Q_j(X) \in \mathrm{Ann}(\phi)$.

Alors

C1. Les sous-espaces $\ker(Q_j(\phi)) \subset E$ *sont ϕ-invariants.*

C2. Les sous-espaces $\ker(Q_j(\phi))$ *donnent une décomposition de E en somme directe*

$$E = \ker(Q_1(\phi)) \oplus \cdots \oplus \ker(Q_k(\phi)). \tag{15}$$

C3. Pour tout $i \in \{1, \dots, k\}$ la projection $p_i : E \to \ker(Q_i(\phi))$ associée à cette décomposition en somme directe, regardée comme endomorphisme de E, est un polynôme de ϕ, donc s'écrit sous la forme $p_i = P_i(\phi)$ pour un polynôme $P_i(X) \in \mathbb{K}[X]$.

Démonstration: C1. Le fait que $\ker(Q_i(\phi))$ sont ϕ-invariants résulte directement de la remarque 3.8. Pour cette affirmation on n'a pas besoin des propriétés des polynômes Q_i données par l'hypothèse du théorème.

Pour démontrer la deuxième et la troisième partie de la conclusion, supposons d'abord $k = 2$.

C2. Puisque $Q_1(X)$, $Q_2(X)$ sont premiers entre eux, on a $\mathrm{pgcd}(Q_1(X), Q_2(X)) = 1$, donc, d'après le corollaire 1.17 (le théorème de Bézout) il existe deux polynômes $U_1(X)$, $U_2(X) \in \mathbb{K}[X]$ tels que

$$1 = U_1(X)Q_2(X) + U_2(X)Q_1(X).$$

En appliquant ces polynômes à ϕ et en tenant compte de le remarque 3.7 on obtient

$$\mathrm{id}_E = U_1(\phi) \circ Q_2(\phi) + U_2(\phi) \circ Q_1(\phi) = Q_2(\phi) \circ U_1(\phi) + Q_1(\phi) \circ U_2(\phi), \tag{16}$$

donc pour tout $v \in E$ on a les identités

$$v = U_1(\phi)(Q_2(\phi(v)) + U_2(\phi)(Q_1(\phi)(v)) \ , \tag{17}$$

$$v = Q_2(\phi)(U_1(\phi)(v)) + Q_1(\phi)(U_2(\phi)(v)) \ . \tag{18}$$

Nous devons démontrer

(i) $\ker(Q_1(\phi)) \cap \ker(Q_2(\phi)) = \{0_E\}$,

(ii) $\ker(Q_1(\phi)) + \ker(Q_2(\phi)) = E$.

Pour (i) soit $v \in \ker(Q_1(\phi)) \cap \ker(Q_2(\phi))$, c'est à dire $Q_2(\phi)(v) = Q_1(\phi)(v) = 0_E$. En utilisant l'identité (17) on obtient $v = 0_E$. Donc $\ker(Q_1(\phi)) \cap \ker(Q_2(\phi)) = \{0_E\}$.

Pour (ii) soit $v \in E$. Nous devons démontrer l'existence de $v_1 \in \ker(Q_1(\phi))$, et $v_2 \in \ker(Q_2(\phi))$ tels que $v = v_1 + v_2$.

En utilisant l'identité (18) il suffit de poser $v_1 := Q_2(\phi)(U_1(\phi)(v))$, $v_2 := Q_1(\phi)(U_2(\phi)(v))$ et de remarquer que $Q_2(\phi)(U_1(\phi)(v)) \in \ker(Q_1(\phi))$ et $Q_1(\phi)(U_2(\phi)(v)) \in \ker(Q_2(\phi))$. En effet,

$$Q_1(\phi)\left(Q_2(\phi)(U_1(\phi)(v))\right) = (Q_1(\phi) \circ Q_2(\phi))(U_1(\phi)(v)) = (Q_1Q_2)(\phi)(U_1(\phi)(v)) = 0_E$$

parce que, par l'hypothèse, $Q_1(X)Q_2(X) \in \mathrm{Ann}(\phi)$. Un calcul similaire montre que $Q_1(\phi)(U_2(\phi)(v)) \in \ker(Q_2(\phi))$.

C3. Pour démontrer la 3me partie de la conclusion (encore dans le cas $k = 2$) remarquer que, d'après la démonstration de (ii) les projections sur $\ker(Q_1(\phi))$, $\ker(Q_2(\phi))$ sont données par

$$p_1 = (Q_2U_1)(\phi),\ p_2 = (Q_1U_2)(\phi,$$

donc il suffit de poser $P_1(X) = Q_2(X)U_1(X)$, $P_2(X) = Q_1(X)U_2(X)$.

Pour $k > 2$ on peut faire soit une démonstration par récurrence, soit généraliser pour k arbitraire la démonstration connue pour $k = 2$. On va suivre la 2-ème méthode et proposer la démonstration par récurrence comme exercice.

Pour $1 \leqslant j \leqslant k$ posons

$$R_j(X) := \prod_{i \neq j} Q_i(X)\ ,$$

c'est à dire

$$R_1(X) = Q_2(X) \dots Q_k(X)\ ,$$
$$R_j(X) = Q_1(X) \dots Q_{j-1}(X)Q_{j+1}(X) \dots Q_k(X) \text{ pour } 1 < j < k$$
$$R_k(X) = Q_1(X) \dots Q_{k-1}(X)\ .$$

Remarquer que dans le cas $k = 2$ traité antérieurement on a $R_1(X) = Q_2(X)$ et $R_2(X) = Q_1(X)$.

C2. (k arbitraire). La démonstration commence avec la remarque

R *Puisque les polynômes $Q_1(X), \dots, Q_k(X)$ sont premiers entre eux deux à deux, les polynômes associés $R_1(X), \dots, R_k(X)$ sont premiers entre eux, c'est à dire* $\mathrm{pgcd}(R_1(X), \dots, R_k(X)) = 1$.

Nous proposons la démonstration de cette remarque comme exercice. Avant de réfléchir à cette démonstration, énoncer et démontrer d'abord une implication similaire pour un système de k nombre naturels $q_1, \dots, q_k$ premier entre eux deux à deux. Commencer par $k = 3$.

Puisque $R_1(X), \dots, R_k(X)$ sont premiers entre eux, d'après le théorème de Bézout (voir le corollaire 1.17), il existe des polynômes $U_j(X)$ $(1 \leqslant j \leqslant k)$ tels que

$$\sum_{i=1} U_i(X)R_i(X) = 1.$$

Comme dans le cas $k = 2$ nous obtenons les identités

$$v = \sum_{i=1}^{k} U_i(\phi)(R_i(\phi)(v))\ \forall v \in E, \tag{19}$$

$$v = \sum_{j=1}^{k} R_j(\phi)(U_j(\phi)(v))\ \forall v \in E. \tag{20}$$

En tenant compte de la définition d'une décomposition en somme directe (voir la définition 2.2), pour démontrer C2 dans le cas général nous devons montrer que

(i) Pour tout $j \in \{1, \ldots, k\}$ nous avons

$$\ker(Q_j(\phi)) \cap \left\{ \sum_{i \neq j} \ker(Q_i(\phi)) \right\} = \{0_E\} \ .$$

(ii)

$$\sum_{j=1}^{k} \ker(Q_j(\phi)) = E \ .$$

Pour démontrer (i) on utilise l'identité de (19) en remarquant que, pour j fixé, tout élément de $\ker(Q_j(\phi))$ appartient à $\ker(R_i(\phi))$ pour $i \neq j$ et tout élément de $\sum_{i \neq j} \ker(Q_i(\phi))$ appartient à $\ker(R_j(\phi))$. Pourquoi ?

Pour démontrer (ii) on utilise l'identité (20) et on obtient pour tout $v \in E$ la décomposition

$$v = \sum_{j=1}^{k} R_j(\phi)(U_j(\phi)(v)).$$

Mais $R_j(\phi)(U_j(\phi)(v)) \in \ker(Q_j(\phi))$ parce que

$$(Q_j(\phi))\,(R_j(\phi)(U_j(\phi)(v))) = \{(Q_j R_j)(\phi)\}\,(U_j(\phi)(v)) = 0_E.$$

Pour la dernière égalité on a utilisé l'égalité $Q_j(X)R_j(X) = \prod_{i=1}^{k} Q_i(X)$ et l'hypothèse $\prod_{i=1}^{k} Q_i(X) \in \mathrm{Ann}(\phi)$.

C3. (k arbitraire) D'après la démonstration de (ii), la projection $p_j : E \to \ker(Q_j(\phi))$ est donnée par

$$p_j = (R_j U_j)(\phi),$$

donc il suffit de poser $P_j(X) = R_j(X)U_j(X)$. ■

3.2.2 Espaces caractéristiques

Définition 3.19 *Soit $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$. On définit le sous-espace caractéristique N_λ de λ par*

$$N_\lambda = \ker((\phi - \lambda \mathrm{id}_E)^{\beta(\lambda)})$$

où, comme d'habitude, $\beta(\lambda)$ désigne la multiplicité de λ en tant que racine du polynôme minimal $m_\phi(X)$.

Donc le sous-espace caractéristique N_λ de λ est le noyau de l'endomorphisme $(X - \lambda)^{\beta(\lambda)}(\phi)$. Notons que

Remarque 3.20 *N_λ est ϕ-invariant et $E_\lambda \subset N_\lambda$.*

Expliquer pourquoi ?

Corollaire 3.21 *1. Soient $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$. Décomposons $m_\phi(X) = (X - \lambda)^{\beta(\lambda)} Q(X)$, avec $Q(X) \in \mathbb{K}[X]$ et $Q(\lambda) \neq 0$. Alors on a une décomposition en somme directe*

$$E = N_\lambda \oplus \ker(Q(\phi)).$$

2. (a) $N_\lambda = \ker((\phi - \lambda \mathrm{id}_E)^k)$ pour tout $k \geqslant \beta(\lambda)$, en particulier pour $k = \mathrm{m}_a(\lambda)$.

(b) $\ker((\phi - \lambda \mathrm{id}_E)^k) \subsetneq N_\lambda$ pour tout $k < \beta(\lambda)$.

3. *Supposons que $P_\phi(X)$ est scindé dans $\mathbb{K}$ donc (voir corollaire le 3.15),*

$$P_\phi(X) = \Pi_{\lambda\in\mathrm{Spec}(\phi)}(X-\lambda)^{\mathrm{m}_a(\lambda)},\ m_\phi(X) = \Pi_{\lambda\in\mathrm{Spec}(\phi)}(X-\lambda)^{\beta(\lambda)}.$$

Alors E se décompose en somme directe de sous-espaces caractéristiques :

$$E = \bigoplus_{\lambda\in\mathrm{Spec}(\phi)} N_\lambda\ .$$

Démonstration: Pour (1) on applique le lemme des noyaux aux polynômes $(X-\lambda)^{\beta(\lambda)}$, $Q(X)$.

Pour (2)(a), soit $k \geqslant \beta(\lambda)$. Nous avons évidemment $N_\lambda \subset \ker((\phi-\lambda\mathrm{id}_E)^k)$. On applique le lemme des noyaux aux polynômes $(X-\lambda\mathrm{id}_E)^k$, $Q(X)$. Il en résulte que $\ker((\phi-\lambda\mathrm{id}_E)^k)$ est aussi un supplémentaire de $\ker(Q(\phi))$ dans E, donc $\dim(\ker(\phi-\lambda\mathrm{id}_E)^k)) = \dim(N_\lambda)$ donc l'inclusion $N_\lambda \subset \ker((\phi-\lambda\mathrm{id}_E)^k)$ est une égalité.

Pour (2)(b), soit $k < \beta(\lambda)$. Nous avons évidemment $\ker((\phi-\lambda\mathrm{id}_E)^k) \subset N_\lambda$. Si on a égalité, alors $(\phi-\lambda\mathrm{id}_E)^k$ s'annulera sur N_λ. En utilisant la décomposition en somme directe $E = N_\lambda \oplus \ker(Q(\phi))$ fournie par (1)), $(\phi-\lambda\mathrm{id}_E)^k Q(\phi)$ s'annulera sur E, donc $(X-\lambda)^k Q(X) \in \mathrm{Ann}(\phi)$. Mais

$$\deg((X-\lambda)^k Q(X)) < \deg((X-\lambda)^{\beta(\lambda)} Q(X)) = \deg(m_\phi(X)),$$

donc ceci contredit la définition de $m_\phi(X)$.

Pour (3) on applique le même lemme aux polynômes $Q_\lambda := (X-\lambda)^{\beta(\lambda)}$, $\lambda \in \mathrm{Spec}(\phi)$. ■

Corollaire 3.22 *Soit $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$. Alors $\dim(N_\lambda) = \mathrm{m}_a(\lambda)$.*

Démonstration: On va utiliser la décomposition en somme directe $E = N_\lambda \oplus \ker(Q(\phi))$ donnée par le corollaire 3.21. Soit ϕ_1 l'endomorphisme induit par ϕ sur N_λ et soit ϕ_2 l'endomorphisme induit par ϕ sur $\ker(Q(\phi))$. On a évidemment $(X-\lambda)^{\beta(\lambda)} \in \mathrm{Ann}(\phi_1)$ (pourquoi ?). D'après la remarque 3.13, toute racine complexe de $P_{\phi_1}(X)$ doit être racine complexe de $(X-\lambda)^{\beta(\lambda)}$ donc doit coïncider avec λ. Donc $P_{\phi_1}(X)$ s'écrit sous la forme $P_{\phi_1}(X) = (-1)^k(X-\lambda)^k$ où $k = \dim(N_\lambda)$. Mais d'après le lemme 3.17 on a $P_\phi(X) = P_{\phi_1}(X)P_{\phi_2}(X)$ donc $P_\phi(X) = (-1)^k(X-\lambda)^k P_{\phi_2}(X)$.

D'autre part $Q(X) \in \mathrm{Ann}(\phi_2)$ (pourquoi ?), donc d'après la remarque 3.13 toute racine complexe de $P_{\phi_2}(X)$ doit être racine complexe de $Q(X)$, donc doit être différente de λ, parce que $Q(\lambda) \neq 0$. Ceci montre que $k = \max\{j \in \mathbb{N}^* |\ (X-\lambda)^j | P_\phi(X)\} = \mathrm{m}_a(\lambda)$. ■

En utilisant la 2me décomposition du corollaire 3.21, nous obtenons un nouveau critère de diagonalisabilité :

Corollaire 3.23 *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$. Sont équivalentes*

1. *ϕ est diagonalisable.*
2. *$P_\phi(X)$ est scindé dans $\mathbb{K}$ et toutes les racines du polynôme minimal $m_\phi(X)$ sont simples (i.e. $\beta(\lambda) = 1$ pour tout $\lambda \in \mathrm{Spec}(\phi)$ et donc $m_\phi(X) = \prod_{\lambda\in\mathrm{Spec}(\phi)}(X-\lambda)$).*

Démonstration: 1. $\Rightarrow$ 2. Si ϕ est diagonalisable nous avons une décomposition en somme directe

$$E = \bigoplus_{\lambda\in\mathrm{Spec}(\phi)} E_\lambda\ .$$

Mais la restriction de $\phi-\lambda\mathrm{id}_X$ à E_λ s'annule pour tout $\lambda \in \mathrm{Spec}(\phi)$. En posant $\mathrm{Spec}(\phi) = \{\lambda_1,\dots,\lambda_k\}$ ceci montre que la composition

$$(\phi-\lambda_1\mathrm{id}_X)\circ\cdots\circ(\phi-\lambda_k\mathrm{id}_X)$$

est nulle sur tout l'espace E, donc $\Pi_{\lambda_\in\mathrm{Spec}(\phi)}(X-\lambda)$ est un polynôme annulateur de ϕ. Il en résulte que $m_\phi(X)$ est un diviseur de ce produit, donc $\beta(\lambda) \leqslant 1$ pour toute valeur propre λ. Puisque $\beta(\lambda) \geqslant 1$ (voir le corollaire 3.16) on obtient $\beta(\lambda) = 1$.

2. $\Rightarrow$ 1. Si $\beta(\lambda) = 1$ alors $N_\lambda = E_\lambda$, donc si 2. est vraie, alors la 3me affirmation du corollaire 3.21 nous donne la décomposition $E = \bigoplus_{\lambda\in\mathrm{Spec}(\phi)} E_\lambda$, ce qui montre que ϕ est diagonalisable. ■

Soient $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$.

Question importante : Comment calculer explicitement l'exposant $\beta(\lambda)$ avec lequel $(X-\lambda)$ intervient dans la factorisation de $m_\phi(X)$ et comment calculer l'espace caractéristique N_λ ?

La réponse est simple :

On va déterminer explicitement *les noyaux successifs*

$$K_j(\lambda) := \ker((\phi - \lambda \mathrm{id}_E)^j).$$

On a évidement une chaîne d'inclusions

$$E_\lambda = K_1(\lambda) \subset K_2(\lambda) \subset \dots.$$

Puisqu'il s'agit d'une chaine croissante de sous-espaces vectoriels d'un espace de dimension finie, il existe un indice *minimal* j_0 pour lequel $K_{j_0}(\lambda) = K_{j_0+1}(\lambda)$. C'est facile de voir que, à partir de j_0 la chaîne d'inclusions se stabilise, donc pour tout $j \geqslant j_0$ on aura aussi $K_j(\lambda) = K_{j+1}(\lambda)$. Expliquer pourquoi.

Puisque $K_{\beta(\lambda)} = K_{\beta(\lambda)+1} = N_\lambda$ (voir le corollaire 3.21 (2)(a)), la définition de j_0 (la minimalité) nous donne $j_0 \leqslant \beta(\lambda)$. Mais la suite de noyaux se stabilise à partir de j_0, donc ceci implique implique $K_{j_0} = K_{\beta(\lambda)} = N_\lambda$.

D'après le corollaire 3.21 (2)(b), l'inégalité $j_0 \leqslant \beta(\lambda)$ ne peut pas être stricte . En conclusion $j_0 = \beta(\lambda)$.

En tenant compte aussi que $\dim(N_\lambda) = \mathrm{m}_a(\lambda)$ on obtient le résultat suivant, qui nous donne un algorithme explicite pour le calcul de l'exposant $\beta(\lambda)$ et de l'espace caractéristique N_λ :

Remarque 3.24 *L'exposant $\beta(\lambda)$ coïncide avec le premier indice j_0 pour lequel l'une des deux conditions équivalentes suivantes est vérifiée :*

1. $\dim(K_{j_0}) = \mathrm{m}_a(\lambda)$.
2. $K_{j_0}(\lambda) = K_{j_0+1}(\lambda)$.

Pour ce nombre $j_0 \in \mathbb{N}^*$ on a $N_\lambda = K_{j_0}(\lambda)$.

Remarque 3.25 *Si $P_\phi(X)$ est scindé dans $\mathbb{K}$ alors $m_\phi(X) = \prod_{\lambda \in \mathrm{Spec}(\phi)} (X-\lambda)^{\beta(\lambda)}$, donc l'algorithme précédent, appliqué à chaque valeur propre $\lambda \in \mathrm{Spec}(\phi)$, va fournir le polynôme minimal $m_\phi(X)$.*

Théorème 3.26 *(réduction selon les espaces caractéristiques) Soient $\phi \in \mathrm{End}(E)$ avec $P_\phi(X)$ scindé dans $\mathbb{K}$ et $\mathrm{Spec}(\phi) = \{\lambda_1, \dots, \lambda_k\}$ son spectre. Pour $1 \leqslant i \leqslant k$ il existe un base B^i de N_{λ_i} telle que la matrice de ϕ dans la base concaténée $B = B^1 \dots B^k$ soit une matrice diagonale par blocs de la forme*

$$M_B(\phi) = \begin{pmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & T_k \end{pmatrix}$$

où chaque bloc T_{λ_i} est une matrice supérieure triangulaire supérieure ayant λ_i sur la diagonale.

Autrement dit, la matrice de ϕ dans la base B aura la forme

$$\begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & 0 & \cdots & 0 \\ 0 & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \boxed{\begin{matrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{matrix}} \end{pmatrix}.$$

Démonstration: D'après le corollaire 3.21 et le lemme 3.17 il suffit de trouver, pour $1 \leqslant i \leqslant k$, une base B^i de N_{λ_i} telle que la matrice de l'endomorphisme $\phi_i \in \mathrm{End}(N_{\lambda_i})$ induit par ϕ soit supérieure triangulaire avec λ_i sur la diagonale.

Fixons $1 \leqslant i \leqslant k$. Nous avons une suite croissante de noyaux successifs

$$\{0_E\} =: K_0(\lambda_i) \subsetneq E_{\lambda_i} = K_1(\lambda_i) \subsetneq \cdots \subsetneq K_{\beta(\lambda_i)}(\lambda_i) = N_{\lambda_i}.$$

En appliquant successivement le théorème de la base incomplète on obtient une base B^i de N_{λ_i} telle que, pour $1 \leqslant j \leqslant \beta(\lambda_i)$, la famille formée des premiers $\dim(K_j(\lambda_i))$ vecteurs de B^i soit une base de $K_j(\lambda_i)$. Une telle base s'écrit explicitement $B^i = B^i_1 \dots B^i_{\beta(\lambda_i)}$, où B^i_j est une base d'un supplémentaire de $K_{j-1}(\lambda_i)$ dans $K_j(\lambda_i)$. Puisque $(\phi - \lambda \mathrm{id}_E)K_j(\lambda_i) \subset K_{j-1}(\lambda_i)$, pour tout élément f_s de B^i_j on a

$$(\phi_i - \lambda_i \mathrm{id}_{N_{\lambda_i}})(f_s) \in \mathrm{Vect}(B^i_1 \dots B^i_{j-1}),$$

donc

$$\phi_i(f_s) = \lambda_i f_s + \text{ une combinaison linéaire de la famille } (B^i_1 \dots B^i_{j-1}).$$

Ceci montre que la $s$ème colonne de la matrice de ϕ_i dans la base B^i ne contient aucun terme sous la diagonale et son terme diagonal est λ_i. ■

4 Endomorphismes nilpotents. La réduction de Jordan des endomorphismes nilpotents

4.1 Endomorphismes nilpotents

Définition 4.1 *Un endomorphisme $u \in \mathrm{End}(E)$ est dit nilpotent s'il existe $k \in \mathbb{N}^*$ tel que $u^k = 0_{\mathrm{End}(E)}$. L'indice de nilpotence d'un endomorphisme nilpotent $u \in \mathrm{End}(E)$ est défini par*

$$\iota(u) := \min\{k \in \mathbb{N}^* |\ u^k = 0_{\mathrm{End}(E)}\}.$$

Comme dans le chapitre précédent, dans cette définition on a utilisé la notation $u^k := \underbrace{u \circ \cdots \circ u}_{k \text{ fois}}$.

Proposition 4.2 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et $u \in \mathrm{End}(E)$ un endomorphisme nilpotent. Alors*

1. $m_u(X) = X^{\iota(u)}$,
2. $\iota(u) \leqslant n$,
3. $P_u(X) = (-1)^n X^n$,
4. $\mathrm{Spec}(u) = \{0\}$.
5. *L'espace propre E_0 coïncide avec $\ker(u)$, en particulier $\ker(u)$ est un sous-espace non-nul de E.*
6. *L'endomorphisme u est diagonalisable si et seulement si $u = 0_{\mathrm{End}(E)}$.*

Démonstration: 1. En effet, par la définition même de $\iota(u)$ on constate que $X^{\iota(u)}$ est un polynôme annulateur de u, donc $m_u(X)$ est un diviseur *unitaire* de $X^{\iota(u)}$ dans $\mathbb{K}[X]$. Donc $m_u(X) = X^j$ où $1 \leqslant j \leqslant \iota(u)$. Pour $1 \leqslant j < \iota(u)$ on a $u^j \neq 0_{\mathrm{End}(E)}$ par la définition de $\iota(u)$. Donc $j = \iota(u)$ et $m_u(X) = X^{\iota(u)}$.

2. Puisque $P_u(X) \in \mathrm{Ann}(u) = (m_u(X))$, on a $m_u(X)|P_u(X)$, donc $n = \deg(P_u(X)) \geqslant \deg(m_u(X)) = \iota(u)$.

3., 4. D'après la proposition 3.14 les racines complexes de $P_u(X)$ coïncident avec les racines complexes de $m_u(X) = X^{\iota(u)}$. Mais $X^{\iota(u)}$ a une seule racine complexe dans $\mathbb{C}$, à savoir 0. La factorisation en facteurs irréductibles (dans $\mathbb{C}[X]$) de $P_u(X)$ sera donc $P_u(X) = (-1)^n(X-0)^n = (-1)^n X^n$. Ceci montre aussi $\mathrm{Spec}(u) = \{0\}$.

5. Résulte de la définition d'un sous-espace propre.

6. On a $\mathrm{m}_a(0) = n$ et $\mathrm{m}_g(0) = \dim(E_0) = \dim(\ker(u))$, donc u est diagonalisable si et seulement si $\dim(\ker(u)) = n$, ce qui est équivalent à $\ker(u) = E$, soit $u = 0_{\mathrm{End}(E)}$. ■

Remarque 4.3 *Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in \mathrm{End}(E)$. Les conditions suivantes sont équivalentes :*

1. *u est nilpotent.*
2. *$P_u(X) = (-1)^n X^n$.*
3. *$P_u(X)$ est scindé dans $\mathbb{K}$ et $\mathrm{Spec}(u) = \{0\}$.*
4. *Il existe une base B de E t.q. la matrice $A = M_B(u)$ de u dans B soit strictement supérieure triangulaire (i.e. $a_{ij} = 0$ pour $i \geqslant j$).*

Démonstration: Exercice. ■

Lemme 4.4 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $u \in \mathrm{End}(E)$ un endomorphisme nilpotent et $\iota := \iota(u)$ son indice de nilpotence. Choisissons $v \in E$ tel que $u^{\iota-1}(v) \neq 0_E$ [1]. Alors la famille*

$$(v, u(v), \ldots, u^{\iota-1}(v))$$

est libre.

Démonstration: Soit $(x_0, \ldots, x_{\iota-1}) \in \mathbb{K}^\iota$ tel que

$$\sum_{i=0}^{\iota-1} x_i u^i(v) = 0_E \ .$$

En appliquant $u^{\iota-1}$ aux deux membres de cette égalité et en tenant compte que $u^j(v) = 0$ pour $j \geqslant \iota$, on obtient $x_0 u^{\iota-1}(v) = 0$, donc (puisque $u^{\iota-1}(v) \neq 0$) on a $x_0 = 0$. En appliquant $u^{\iota-2}$ on obtient d'une manière similaire $x_1 = 0$ et, par recurrence, $x_i = 0$ pour tout $i \in \{0, \ldots, \iota - 1\}$. ■

Proposition 4.5 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $u \in \mathrm{End}(E)$ un endomorphisme nilpotent d'indice de nilpotence $\iota(u) = n$. Soit $v \in E$ tel que $u^{n-1}(v) \neq 0_E$. Alors*

1. *La famille $B := (u^{n-1}(v), u^{n-2}(v), \ldots, u(v), v)$ est une base de E,*
2. *La matrice de u dans cette base est*

$$M_B(u) = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n} .$$

Remarquer que la famille B a été obtenue *en renversant l'ordre* des éléments de la famille qui intervient dans le lemme 4.4.

Démonstration: 1. D'après ce lemme 4.4 B est une famille libre de E. Mais le cardinal de cette famille coïncide avec la dimension de l'espace, donc il s'agit d'une base.
2. Posons $A := M_B(u)$. Pour $1 \leqslant i \leqslant n$ posons $f_i := u^{n-i}(v)$. Alors

$$u(f_1) = u^n(v) = 0_E \ , \ u(f_j) = u(u^{n-j}(v)) = u^{n-(j-1)}(v) = f_{j-1} \text{ pour } 2 \leqslant j \leqslant n.$$

En comparant ces formules avec les formules connues $u(f_j) = \sum_{i=1}^n a_{ij} f_i$ pour la matrice associée à u, on voit que $a_{i,j} = 0$ pour $i \neq j - 1$ et $a_{j-1,j} = 1$ pour $2 \leqslant j \leqslant n$.

■

1. Un tel vecteur existe, parce que par la définition de l'indice de nilpotence, on a $u^{\iota-1} \neq 0_{\mathrm{End}(E)}$.

Définition 4.6 *La matrice carrée*

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n} \in M_{n,n}(\mathbb{K})$$

qui intervient dans l'énoncé de la proposition 4.5 s'appelle bloc de Jordan d'ordre n et valeur propre 0, et cette matrice sera notée J_0^n.

Donc la proposition 4.5 affirme que, dans une base convenable, la matrice d'un endomorphisme de nilpotence $u \in \mathrm{End}(E)$ avec $\iota(u) = \dim(E)$ est un bloc de Jordan d'ordre $\dim(E)$ et valeur propre 0.

La proposition 4.5 donne la forme canonique des endomorphismes nilpotents $u \in \mathrm{End}(E)$ d'indice de nilpotence maximal $\iota(u) = \dim(E)$. On va étudier maintenant la forme canonique des endomorphismes nilpotents en général.

4.2 La réduction de Jordan des endomorphismes nilpotents

Remarque 4.7 *Soit $u \in \mathrm{End}(E)$ un endomorphisme nilpotent. posons $\iota := \iota(u)$. Nous avons la suite l'inclusions* `strictes`

$$\{0_E\} \subsetneq \ker(u) \subsetneq \cdots \subsetneq \ker(u^{\iota-1}) \subsetneq \ker(u^\iota) = E \ ,$$

et $\ker(u^j) = E$ pour tout $j \geqslant \iota$.

Démonstration: Pour tout $j \in \mathbb{N}$ on a évidemment $\ker(u^j) \subset \ker(u^{j+1})$ et pour tout $j \geqslant \iota$ on a évidemment $\ker(u^j) = E$ (pourquoi ?). Il suffit donc de vérifier que les inclusions $\ker(u^j) \subset \ker(u^{j+1})$ sont strictes pour $0 \leqslant j \leqslant \iota-1$. En effet supposons par l'absurde qu'il existe $j \in \{0, \ldots, \iota-1\}$ tel que $\ker(u^j) = \ker(u^{j+1})$. Si $v \in \ker(u^{j+2})$ alors $u(v) \in \ker(u^{j+1}) = \ker(u^j)$, donc $u^j(u(v)) = 0_E$, i.e. $v \in \ker(u^{j+1})$. Nous avons démontré que l'inclusion suivante $\ker(u^{j+1}) \subset \ker(u^{j+2})$ est aussi une égalité. Par récurrence, on montre facilement que $\ker(u^{j'}) \subset \ker(u^{j+1})$ pour tout $j' \geqslant j$, donc

$$\ker(u^j) = \ker(u^{j+1}) = \cdots = \ker(u^\iota) = E \ .$$

Mais l'égalité $\ker(u^j) = E$ est équivalente à $u^j = 0_{\mathrm{End}(E)}$, ce qui implique $\iota = \iota(\phi) \leqslant j$, ce qui est évidemment une contradiction. ■

Proposition 4.8 *(la décomposition en somme directe des noyaux successifs) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $u \in \mathrm{End}(E)$ un endomorphisme nilpotent. Posons $\iota := \iota(u)$ et soit*

$$\{0_E\} \subsetneq K_1 = E_0 \subsetneq \cdots \subsetneq K_{\iota-1} \subsetneq K_\iota = E$$

la suite croissante des noyaux itérés successifs, donc $K_j := \ker(u^j)$ pour $1 \leqslant j \leqslant \iota$. Il existe des sous-espaces vectoriels $F_j \subset K_j$, où $F_\iota \neq \{0_E\}$, tels que les noyaux K_j se décomposent en somme directe comme suit

$$\begin{array}{ccccccccccccccccc}
E & = & K_\iota & = & K_{\iota-1} & \oplus & F_\iota & & & & & & & & & & \\
 & & K_{\iota-1} & = & K_{\iota-2} & \oplus & u(F_\iota) & \oplus & F_{\iota-1} & & & & & & & & \\
 & & K_{\iota-2} & = & K_{\iota-3} & \oplus & u^2(F_\iota) & \oplus & u(F_{\iota-1}) & \oplus & F_{\iota-2} & & & & & & \\
 & & \vdots & & & & \vdots & & \vdots & & \vdots & & & & & & \\
E_0 & = & K_1 & = & & & u^{\iota-1}(F_\iota) & \oplus & u^{\iota-2}(F_{\iota-1}) & \oplus & u^{\iota-3}(F_{\iota-2}) & \oplus & \ldots & \oplus & u(F_2) & \oplus & F_1.
\end{array}$$

Démonstration:

Nous devons construire des des sous-espaces vectoriels $F_j \subset K_j$ tels que pour tout $j \in \{1, \ldots, \iota\}$ on a la décomposition en somme directe :

$$K_j = K_{j-1} \oplus \left\{ \bigoplus_{i=j}^{\iota} u^{i-j}(F_i) \right\}, \tag{D_j}$$

où, pour $j = 1$, on a utilisé la notation $K_0 = \{0_E\}$. Remarquons d'abord que $u^s(K_r) \subset K_{r-s}$ pour $0 \leqslant s \leqslant r$ (pourquoi ?) donc dans le terme $u^{i-j}(F_i)$ dans le second membre de (D_j) est bien un sous-espace de K_j.

Pour démontrer la proposition on va obtenir les sous-espaces F_k par *récurrence descendante* en commençant par ι. Pour $k = \iota$ il suffit de choisir un sous-espace supplémentaire F_ι de $K_{\iota-1}$ dans $K_\iota = E$. Fixons $k \in \{2, \ldots, \iota\}$ et supposons que nous avons trouvé des sous-espaces $F_j \subset K_j$ pour $j \in \{k, \ldots, \iota\}$ tels que la décomposition en some directe (D_j) soit vérifiée pour $j \in \{k, \ldots, \iota\}$. On va montrer l'existence d'un sous-espace vectoriel $F_{k-1} \subset K_{k-1}$ tel que (D_{k-1}) soit vraie. À cette fin on va utiliser la remarque simple suivante :

R. *Supposons $k \geqslant 2$ et soient $F_j \subset K_j$ des sous-espaces vectoriels donnés pour $k \leqslant j \leqslant \iota$ tels que (D_k) est vraie. Alors les sous-espaces*

$$K_{k-2},\ u^{i-k+1}(F_i)\ k \leqslant i \leqslant \iota$$

de K_{k-1} sont des sous-espaces en somme directe.

En effet, supposons que nous avons une relation du type

$$x + \sum_{i=k}^{\iota} u^{i-k+1}(f_i) = 0_E \tag{21}$$

avec $x \in K_{k-2}$ et $f_i \in F_i$ pour $k \leqslant i \leqslant \iota$. En appliquant u^{k-2} aux deux membres et en tenant compte que $x \in K_{k-2}$ on obtient

$$\sum_{i=k}^{\iota} u^{i-1}(f_i) = 0_E\ ,$$

soit

$$u^{k-1}(\sum_{i=k}^{\iota} u^{i-k}(f_i)) = 0_E\ , \tag{22}$$

ce qui implique $\sum_{i=k}^{\iota} u^{i-k}(f_i) \in \ker(u^{k-1}) = K_{k-1}$. Mais, puisque (D_k) est vraie par l'hypothèse, les sous-espaces K_{k-1}, $u^{i-k}(F_{\iota-i})$ $(k \leqslant i \leqslant \iota)$ sont en somme directe, donc (22) implique $u^{i-k}(f_i) = 0_E$ pour $k \leqslant i \leqslant \iota$. Ceci implique $u^{i-k+1}(f_i) = 0_E$ pour $k \leqslant i \leqslant \iota$ et, en revenant à (21), on obtient $x = 0_E$. La remarque **R** est donc démontrée.

D'après la remarque **R** les sous-espaces K_{k-2}, $u^{i-k+1}(F_i)$ $(k \leqslant i \leqslant \iota)$ sont en somme directe dans K_{k-1}. Il suffit de choisir un sous-espace supplémentaire F_{k-1} de la somme directe

$$K_{k-2} \oplus \left\{ \bigoplus_{i=k}^{\iota} u^{i-k+1}(F_i) \right\}$$

dans K_{k-1}.

■

Important : D'après la proposition 4.8 le premier sous-espace supplémentaire F_ι est non-nul. En général, on peut avoir $F_j = \{0_E\}$ pour $j < \iota$.

Important : La démonstration de la proposition 4.8 a un caractère algorithmique (effectif), donc on peut l'utiliser pour déterminer les sous-espaces F_j effectivement. A chaque étape de l'algorithme on construit un sous-espace supplémentaire d'un sous-espace connu dans le noyau connu K_k. Dans les exercices concrets le "sous-espace connu" est muni naturellement d'une base, donc un sous-espace supplémentaire est déterminé en utilisant le théorème de la base incomplète.

Remarque 4.9 *En utilisant les hypothèses et les notations de la proposition 4.8, la restriction*

$$u^s\big|_{F_j} : F_j \to u^s(F_j)$$

est un isomorphisme pour $0 \leqslant s \leqslant j-1$.

En effet, la surjectivité est évidente. En ce qui concerne l'injectivité il suffit de la démontrer pour $s = j-1$. En effet si $u^{j-1}(f_j) = 0_E$ alors $f_j \in K_{j-1}$, mais K_{j-1} et F_j sont en somme directe donc $f_j = 0_E$.

Corollaire 4.10 *En utilisant les hypothèses et les notations de la proposition 4.8 on a une décomposition en somme directe*

$$E = \bigoplus_{\substack{1\leqslant j\leqslant \iota \\ 0\leqslant i\leqslant j-1}} u^i(F_j) \ .$$

Démonstration: En utilisant successivement les décompositions (D_ι), $(D_{\iota-1}), \ldots, (D_1)$ on obtient

$$E = K_\iota = K_{\iota-1} \oplus F_\iota = (K_{\iota-2} \oplus u(F_\iota) \oplus F_{\iota-1}) \oplus F_\iota = \cdots = \bigoplus_{\substack{1\leqslant j\leqslant \iota \\ 0\leqslant i\leqslant j-1}} u^i(F_j) \ .$$

■

Théorème 4.11 *(la réduction de Jordan des endomorphismes nilpotents) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $u \in \mathrm{End}(E)$ un endomorphisme nilpotent. Posons $\iota := \iota(u)$, soit*

$$\{0_E\} \subsetneq K_1 = E_0 \subsetneq \cdots \subsetneq K_{\iota-1} \subsetneq K_\iota = E$$

la suite croissante de noyaux des itérés successifs $K_j := \ker(u^j)$ et soit $F_j \subset K_j$ des sous-espaces vectoriels satisfaisant la conclusion de la proposition 4.8. Alors E admet une base B telle que

$$M_B(\phi) = \begin{pmatrix} J_0^{j_1} & 0 & \ldots & 0 \\ 0 & J_0^{j_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_0^{j_k} \end{pmatrix} ,$$

où

1. *L'ordre maximal $j_{\max} := \max_{1\leqslant i\leqslant k} j_i$ des blocs de Jordan $J_0^{j_i}$ qui interviennent dans $M_B(\phi)$ est ι.*
2. *Pour $1 \leqslant j \leqslant \iota$ on a précisément $\dim(F_j)$ blocs de Jordan d'ordre j qui interviennent dans $M_B(\phi)$. En particulier, si $F_j = \{0_E\}$, on n'a a aucun bloc de Jordan d'ordre j dans cette matrice.*
3. *Le nombre total k des blocs de Jordan qui interviennent dans $M_B(\phi)$ est $\mathrm{m}_g(0)$.*

Démonstration: Pour chaque sous-espace F_j posons

$$\tilde{F}_j := \bigoplus_{i=0}^{j-1} u^i(F_j) \ .$$

Remarquons que ce sous-espace est u-invariants et, d'après le corollaire 4.10, nous avons une décomposition en somme directe

$$E = \bigoplus_{j=1}^{\iota} \tilde{F}_j \ .$$

Posons $d_j := \dim(F_j)$. Supposons que $d_j > 0$ et soit $B^j = (f_1^j, \ldots, f_{d_j}^j)$ une base de F_j. Alors la famille

$$\tilde{B}_j :=$$

$$u^{j-1}(f_1^j), u^{j-2}(f_1^j), \ldots, u(f_1^j), f_1^j, u^{j-1}(f_2^j), u^{j-2}(f_2^j), \ldots, u(f_2^j), f_2^j, \ldots, u^{j-1}(f_{d_j}^j), u^{j-2}(f_{d_j}^j), \ldots, u(f_{d_j}^j), f_{d_j}^j$$

sera une base de $\tilde{F}_j$. En effet, cette famille a été obtenue an appliquant une permutation à la base concaténée naturelle

$$B_j \ u(B_j) \ \dots \ u^{j-1}(B_j)$$

de $F_j \oplus u(F_j) \oplus \cdots \oplus u^{j-1}(F_j) = \tilde{F}_j$. Les mêmes arguments que ceux utilisés dans la démonstration de deuxième affirmation de la proposition 4.5 montre que la matrice de l'endomorphisme $\phi_j \in \mathrm{End}(\tilde{F}_j)$ induit par ϕ est une matrice formée de d_j blocs de Jordan d'ordre j. En utilisant la base concaténée formée des familles $\tilde{B}_j$ (pour $d_j > 0$) on obtient une base de E, et (d'après le lemme 3.17) la matrice de ϕ dans cette base aura évidemment les propriétés 1., 2.

Pour démontrer 3. remarquer que chaque base $\tilde{B}_j$ a précisément d_j vecteurs qui sont éléments de K_1, à savoir les vecteurs $u^{j-1}(f_i^j)$ où $1 \leqslant i \leqslant d_j$. Mais en utilisant la dernière décomposition donnée par la proposition 4.8 on constate que la famille de vecteurs

$$(u^{j-1}(f_i^j))_{\substack{1 \leqslant j \leqslant \iota \\ 1 \leqslant i \leqslant d_j}}$$

est une base de l'espace propre $K_1 = E_0$. Donc

$$\mathrm{m}_g(0) = \dim(E_0) = \dim(K_1) = \sum_{j=1}^{\iota} d_j = \text{ le nombre total des blocs de Jordan de la matrice } M_B(\phi) \ .$$

■

Important : La démonstration du théorème 4.11 fournit un algorithme explicite pour la réduction de Jordan d'un endomorphisme nilpotent u, donc un algorithme qui donne une base B dans laquelle la matrice de u a une forme réduite et l'ordre des blocs de Jordan qui interviennent dans cette matrice. Pour appliquer cet algorithme dans des situations concrètes il faut commencer par trouver des bases B^j dans les sous-espaces supplémentaires F_j.

Nous achevons cette section par un résultat simple qui sera utilisé dans la démonstration du théorème de Jordan :

Proposition 4.12 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$ et $\lambda \in \mathrm{Spec}(\phi)$. Soit $\phi_\lambda \in \mathrm{End}(N_\lambda)$ l'endomorphisme induit par ϕ sur le sous-espace ϕ-invariant N_λ. Alors l'endomorphisme $u_\lambda := \phi_\lambda - \lambda \mathrm{id}_{N_\lambda}$ est nilpotent et $\iota(u_\lambda) = \beta(\lambda)$, où $\beta(\lambda)$ désigne la multiplicité de λ en tant que racine du polynôme minimal $m_\phi(X)$.*

Démonstration: Par définition $N_\lambda = \ker(\phi - \lambda \mathrm{id}_E)^{\beta(\lambda)}$, donc, d'une manière tautologique, la restriction de $(\phi - \lambda \mathrm{id}_E)^{\beta(\lambda)}$ à N_λ est nulle. Ceci implique $u_\lambda^{\beta(\lambda)} = 0_{\mathrm{End}(N_\lambda)}$, donc $\iota(u_\lambda) \leqslant \beta(\lambda)$. D'autre part, par la remarque **??**, pour $0 \leqslant j < \beta(\lambda)$ le noyau de u_λ^j est un sous-espace propre de N_λ, i.e. $u_\lambda^j \neq 0_{\mathrm{End}(N_\lambda)}$. D'après la définition de $\iota(u_\lambda)$ on a donc $\iota(u_\lambda) = \beta(\lambda)$. ■

5 Le théorème de Jordan et la décomposition de Dunford

5.1 Le théorème de Jordan

Définition 5.1 *La matrice carrée*

$$\underbrace{\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}}_{n} \in M_{n,n}(\mathbb{K})$$

s'appelle bloc de Jordan d'ordre n et valeur propre λ et sera noté J_λ^n.

Théorème 5.2 *(le théorème de Jordan) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, soit $\phi \in \mathrm{End}(E)$ un endomorphisme avec $P_\phi(X)$ scindé dans $\mathbb{K}$ et soit $\{\lambda_1, \dots, \lambda_k\}$ son spectre. Alors il existe une base B de E telle que*

$$M_B(\phi) = \begin{pmatrix} \tilde{J}_{\lambda_1} & 0 & \cdots & 0 \\ 0 & \tilde{J}_{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{J}_{\lambda_k} \end{pmatrix} ,$$

où chaque bloc $\tilde{J}_{\lambda_i}$ est une matrice d'ordre $\mathrm{m}_a(\lambda_i)$ formée de k_i blocs de Jordan (positionnés sur la diagonale) de valeur propre λ_i, donc une matrice de la forme

$$\tilde{J}_{\lambda_i} = \begin{pmatrix} J_{\lambda_i}^{n_1^i} & 0 & \cdots & 0 \\ 0 & J_{\lambda_i}^{n_2^i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_i}^{n_{k_i}^i} \end{pmatrix}$$

En plus

1. *L'ordre maximal des blocs de Jordan de valeur propre λ_i qui interviennent dans $\tilde{J}_{\lambda_i}$ coïncide avec la multiplicité $\beta(\lambda_i)$ de λ_i en tant que racine du polynôme minimal $m_\phi(X)$.*
2. *Le nombre k_i de blocs de Jordan de valeur propre λ_i qui interviennent dans $\tilde{J}_{\lambda_i}$ coïncide avec la multiplicité géométrique $\mathrm{m}_g(\lambda_i)$.*

Démonstration:

D'après le corollaire 3.21 nous avons une décomposition en somme directe

$$E = N_{\lambda_1} \oplus N_{\lambda_2} \oplus \cdots \oplus N_{\lambda_k}$$

de sous-espaces caractéristiques, qui sont ϕ-invariants. Soit $\phi_\lambda \in \mathrm{End}(N_\lambda)$ l'endomorphisme induit par ϕ sur l'espace ϕ-invariant N_λ, et soit

$$u_\lambda := \phi_\lambda - \lambda \mathrm{id}_{N_\lambda} \in \mathrm{End}(N_\lambda) .$$

D'après la proposition 4.12 l'endomorphisme u_λ de N_λ est nilpotent. Pour chaque $i \in \{1, \dots, k\}$ soit B_i une base de N_{λ_i} qui réduit u_{λ_i} au sens du théorème 4.11, donc telle que $M_{B_i}(u_{\lambda_i})$ soit une matrice formée de blocs de Jordan (positionné sur la diagonale) de valeur propre 0 ayant le propriétés formulées dans ce théorème, à savoir :

1. le nombre k_i de blocs de Jordan de cette matrice coïncide avec la multiplicité géométrique $\mathrm{m}_g^{u_{\lambda_i}}(0)$ de 0 associée à l'endomorphisme u_{λ_i},
2. l'ordre maximal des blocs de Jordan de cette matrice est l'indice de nilpotence $\iota(u_{\lambda_i})$ de l'endomorphisme nilpotent u_{λ_i}.

Puisque $\phi_{\lambda_i} = u_{\lambda_i} + \lambda_i \mathrm{id}_{N_{\lambda_i}}$ il en résulte que la matrice de ϕ_{λ_i} dans la base B_i est

$$M_{B_i}(\phi_{\lambda_i}) = \lambda_i I_{n_i} + M_{B_i}(u_{\lambda_i}) ,$$

où $n_i := \dim(N_{\lambda_i})$, donc $M_{B_i}(\phi_{\lambda_i})$ est une matrice formée de blocs de Jordan (positionné sur la diagonale) de valeur propre λ_i. Désignons par $\tilde{J}_{\lambda_i}$ cette matrice. Notons que $\dim(N_{\lambda_i}) = \mathrm{m}_a(\lambda_i)$ d'après le corollaire 3.22, donc $\tilde{J}_{\lambda_i}$ est bien une matrice d'ordre $\mathrm{m}_a(\lambda_i)$.

Il suffit de considerer la base concaténée

$$B = B_1 B_2 \dots B_k ,$$

d'appliquer le lemme 3.17, et de remarquer que $\mathrm{m}_g^{u_{\lambda_i}}(0) = \mathrm{m}_g(\lambda_i)$ (exercice) et $\iota(u_{\lambda_i}) = \beta(\lambda_i)$ d'après la proposition 4.12. ■

Important : Les démonstrations des théorèmes 5.2 et 4.11 fournissent un algorithme explicite pour la jordanisations des endomorphismes à polynôme caractéristique scindé. Il suffit d'appliquer l'algorithme pour la réduction des endomorphismes nilpotents (donné par la démonstration du théorème 4.11) à chaque endomorphisme nilpotent $u_\lambda \in \mathrm{End}(N_\lambda)$ et de concaténer les bases obtenues.

5.2 La décomposition de Dunford

Lemme 5.3 *(sous-espaces invariants par un endomorphisme diagonalisable) Soit E un $\mathbb{K}$ espace vectoriel de dimension finie n, $\phi \in \mathrm{End}(E)$ un endomorphisme diagonalisable et $F \subset E$ un sous-espace ϕ-invariant. Alors*

1. *L'endomorphisme $\phi_F \in \mathrm{End}(F)$ induit par ϕ est diagonalisable.*
2. *Le spectre $\mathrm{Spec}(\phi_F)$ de ϕ_F est donné par*
$$\mathrm{Spec}(\phi_F) = \{\lambda \in \mathrm{Spec}(\phi) |\ F \cap E_\lambda \neq \{0_E\}\} \ ,$$
donc il est un sous-ensemble de $\mathrm{Spec}(\phi)$.
3. *Pour tout $\lambda \in \mathrm{Spec}(\phi_F)$ l'espace propre F_λ de ϕ_F associé à λ est*
$$F_\lambda = F \cap E_\lambda \ .$$
4. *On a une décomposition en somme directe*
$$F = \bigoplus_{\substack{\lambda \in \mathrm{Spec}(\phi) \\ F \cap E_\lambda \neq \{0_E\}}} (F \cap E_\lambda) \ .$$

Démonstration: 1. Soit $m_\phi(X)$ le polynôme minimal de ϕ. Puisque $m_\phi(\phi) = 0_{\mathrm{End}(E)}$ on a évidemment $m_\phi(\phi_F) = 0_{\mathrm{End}(F)}$ (pourquoi ?), donc $m_\phi(X)$ est aussi un polynôme annulateur de ϕ_F . Il en résulte que $m_{\phi_F}(X)$ est un diviseur de $m_\phi(X)$ (qui este scindé dans $\mathbb{K}$ et dont toutes les racines sont simples, d'après le Corollaire 3.23), ce qui implique que $m_{\phi_F}(X)$ est aussi un polynôme scindé dans $\mathbb{K}$ sont toutes les racines sont simples. Donc ϕ_F est diagonalisable d'après le Corollaire 3.23.

2., 3. Remarquer que pour *tout* $\lambda \in \mathbb{K}$ on a $\ker(\phi_F - \lambda \mathrm{id}_F) = F \cap \ker(\phi - \lambda \mathrm{id}_E)$. Expliquer cette égalité. Cette égalité implique

$$\begin{aligned}\mathrm{Spec}(\phi_F) := \{\lambda \in \mathbb{K} |\ \ker(\phi_F - \lambda \mathrm{id}_F) \neq \{0_E\}\} &= \{\lambda \in \mathbb{K} |\ F \cap \ker(\phi - \lambda \mathrm{id}_E) \neq \{0_E\}\} \\ &= \{\lambda \in \mathrm{Spec}(\phi) |\ F \cap E_\lambda \neq \{0_E\}\} \ .\end{aligned}$$

4. En appliquant le théorème 2.16 3. à l'endomorphisme diagonalisable ϕ_F on obtient

$$F = \bigoplus_{\lambda \in \mathrm{Spec}(\phi_F)} F_\lambda = \bigoplus_{\substack{\lambda \in \mathrm{Spec}(\phi) \\ F \cap E_\lambda \neq \{0_E\}}} (F \cap E_\lambda) \ .$$

■

Lemme 5.4 *(endomorphismes diagonalisables qui commutent) Soit E un $\mathbb{K}$ espace vectoriel de dimension finie n et ϕ, $\psi \in \mathrm{End}(E)$ qui commutent, i.e. tels que $\phi \circ \psi = \psi \circ \phi$. Alors*

1. *Tout espace propre E_λ de ϕ est ψ-invariant.*
2. *Si ϕ et ψ sont diagonalisables, alors ils sont simultanément diagonalisables, i.e. il existe une base B de E telle que les deux matrices $M_B(\phi)$, $M_B(\psi)$ soient diagonales.*

Démonstration: 1. Soit $x \in E_\lambda$. Alors

$$\phi(\psi(x)) = \psi(\phi(x)) = \psi(\lambda x) = \lambda \psi(x) \ ,$$

donc $\psi(x) \in E_\lambda$ par la définition de ce sous-espace.

2. En appliquant le lemme 5.3 à l'endomorphisme ψ et au sous-espace ψ-invariant E_λ il en résulte que l'endomorphisme $\psi_\lambda \in \mathrm{End}(E_\lambda)$ induit par ψ est diagonalisable. Soit B_λ une base de E_λ qui diagonalise ψ_λ. Les vecteurs de B_λ seront des vecteurs propres de ψ_λ, donc aussi de ψ. En posant $\mathrm{Spec}(\phi) = \{\lambda_1, \lambda_2 \ldots, \lambda_k\}$ on obtient une base concaténée

$$B = B_{\lambda_1} B_{\lambda_2} \ldots B_{\lambda_k}$$

qui est une base formée de vecteurs qui sont simultanément des vecteurs propres de ϕ et ψ, donc cette base diagonalise ϕ et ψ simultanément. ■

Lemme 5.5 *(endomorphismes nilpotents qui commutent) Soient u, $v \in \mathrm{End}(E)$ deux endomorphismes nilpotents qui commutent et a, $b \in \mathbb{K}$. Alors l'endomorphisme $au + bv$ est nilpotent.*

Démonstration: Soient m, $n \in \mathbb{N}$ tels que $u^m = v^n = 0_{\mathrm{End}(E)}$

Puisque u et v commutent, au et bv commutent aussi donc on peut dévelpper $(au+bv)^{m+n}$ en utilisant la formule du binôme de Newton, donc

$$(au + bv)^{m+n} = \sum_{k=0}^{m+n} C_{m+n}^k a^k b^{n+m-k} u^k v^{n+m-k} \ .$$

Remarquons que pour tout $k \in \{0, \ldots, m+n\}$ on a soit $k \geqslant m$ soit $n+m-k \geqslant n$, donc dans chaque terme de cette somme on a soit $u^k = 0$, soit $v^{n+m-k} = 0$. Ceci montre que $(au + bv)^{m+n} = 0$, donc $au + bv$ est nilpotent.

■

Théorème 5.6 *(la décomposition de Dunford) Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\phi \in \mathrm{End}(E)$ avec $P_\phi(X)$ scindé dans $\mathbb{K}$. Alors il existe un et un seul couple (δ, u) d'endomorphismes de E tels que :*

1. *$\phi = \delta + u$.*
2. *δ et u commutent.*
3. *δ est diagonalisable et u est nilpotent.*

De plus, les endomorphismes δ et u satisfaisant les trois propriétés ci-dessus sont des polynômes de ϕ.

Démonstration: A. *L'existence :* Soit

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(\phi)} N_\lambda$$

la décomposition de E en somme directe de sous-espaces caractéristiques de ϕ (voir le corollaire 3.21) et désignons par $p_\lambda \in \mathrm{End}(E)$ la projection sur N_λ par rapport à cette décomposition. Posons

$$\delta := \sum_{\lambda \in \mathrm{Spec}(\phi)} \lambda p_\lambda \ .$$

Plus précisément, en décomposant chaque $x \in E$ comme $x = \sum_{\lambda \in \mathrm{Spec}(\phi)} x_\lambda$ avec $x_\lambda \in N_\lambda$, on a

$$\delta(x) := \sum_{\lambda \in \mathrm{Spec}(\phi)} \lambda x_\lambda \ .$$

C'est facile de voir que δ est diagonalisable. En effet l'espace caractéristique N_λ de ϕ devient espace propre de δ, donc E se décompose en somme directe d'espaces propres de δ.

Pour compléter la démonstration de l'existence, il suffit de montrer que $u := \phi - \delta$ est nilpotent et que δ et u commutent.

La restriction de u à N_λ coïncide avec la restriction de $\phi - \lambda \mathrm{id}$ sur cet espace, et par la définition de N_λ, on a

$$(\phi - \lambda \mathrm{id})^{\beta(\lambda)}\Big|_{N_\lambda} = 0 \ .$$

Donc $u^{\beta(\lambda)}\Big|_{N_\lambda} = 0$ pour tout $\lambda \in \mathrm{Spec}(\phi)$. Il en résulte $u^{\max\{\beta(\lambda) |\ \lambda \in \mathrm{Spec}(\phi)\}} = 0_{\mathrm{End}(E)}$, donc u est bien nilpotent.

Vérifions que δ et u commutent. Puisque N_λ est ϕ invariant (d'après le corollaire 3.21) et évidemment δ invariant, il en résulte que N_λ est aussi u-invariant. Pour $x \in N_\lambda$ on a donc

$$u(\delta(x)) = u(\lambda x) = \lambda u(x) = \delta(u(x)).$$

Donc u et δ commutent sur N_λ pour tout $\lambda \in \mathrm{Spec}(\phi)$, ce qu'implique qu'ils commutent sur la somme de ces sous-espaces, donc sur E.

D'après la 3me conclusion C3 du théorème 3.18 (du théorème des noyaux), les projections $p_\lambda \in \mathrm{End}(E)$ sont des polynômes de ϕ, donc $\delta = \sum_{\lambda \in \mathrm{Spec}(\phi)} \lambda p_\lambda$ sera aussi un polynôme de ϕ. Ceci implique que $u = \phi - \delta$ sera aussi un polynôme de ϕ.

B. *L'unicité :*

Supposons que $\phi = \delta' + u'$ où δ' est diagonalisable, u' est nilpotent et $\delta' \circ u' = u' \circ \delta'$. Puisque δ' commute avec u', il va commuter aussi avec $\phi = \delta' + u'$, donc avec tout polynôme de ϕ, en particulier avec $\delta = \sum_{\lambda \in \mathrm{Spec}(\phi)} \lambda p_\lambda$. Puisque les endomorphismes δ et δ' commutent et sont diagonalisables, il en résulte d'après le lemme 5.4 qu'ils sont simultanément diagonalisables. Mais alors $\delta - \delta'$ sera aussi diagonalisable (plus précisément sera diagonalisé par toute base qui diagonalise simultanément δ et δ'). Similairement u et u' commutent, donc d'après le lemme 5.5 la différence $u' - u$ sera un endomorphisme nilpotent. Mais l'égalité $\delta + u = \delta' + u'$ implique

$$\delta - \delta' = u' - u.$$

Il en résulte que l'endomorphisme $\delta - \delta'$ est à fois diagonalisable et nilpotent, donc $\delta - \delta' = 0$ (voir la proposition 4.2). ■

Remarque 5.7 *Soit $A \in M_{n,n}(\mathbb{K})$ une matrice supérieure triangulaire. Alors A s'écrit comme $A = D + U$, où D est la matrice formée des éléments diagonaux de A et U est la matrice formée des éléments qui se trouvent au-dessus de la diagonale. D est diagonalisable et U est nilpotente (voir la remarque 4.3, mais en en général D et U ne commutent pas, donc, en général, $A = D + U$ n'est* `pas` *la décomposition de Dunford de (l'endomorphisme associé à) A.*

5.3 L'exponentielle d'un endomorphisme

Soient $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ et E un $\mathbb{K}$-espace vectoriel de dimension finie n. Munissons cet espace d'une norme $\|\cdot\| : E \to \mathbb{R}_+$. On va munir $\mathrm{End}(E)$ de la norme d'opérateur associée à $\|\cdot\|$: pour $\phi \in \mathrm{End}(E)$ on pose

$$N(\phi) := \sup_{x \in E \setminus \{0_E\}} \frac{\|\phi(x)\|}{\|x\|} = \sup_{\substack{x \in E \\ \|x\|=1}} \|\phi(x)\|.$$

La norme d'opérateur sur l'espace d'endomorphismes a deux propriétés importantes :

Proposition 5.8 *1. Pour tout $x \in E$ et $\phi \in \mathrm{End}(E)$ on a $\|\phi(x)\| \leqslant N(\phi)\|x\|$.*

2. Pour tous ϕ, $\psi \in \mathrm{End}(E)$ on a $N(\psi \circ \phi) \leqslant N(\psi)N(\phi)$.

Par récurrence on obtient les inégalités

$$\forall \phi \in \mathrm{End}(E)\, \forall k \in \mathbb{N}^*\, \forall x \in E,\ \|\phi^k(x)\| \leqslant N(\phi)^k \|x\|. \tag{23}$$

$$\forall \phi \in \mathrm{End}(E)\, \forall k \in \mathbb{N}^*,\ N(\phi^k) \leqslant N(\phi)^k. \tag{24}$$

La deuxième inégalité montre que

Proposition 5.9 *La série $\sum_{k \geqslant 0} \frac{1}{k!}\phi^k$ est absolument convergente dans l'espace normé* $(\mathrm{End}(E), N)$.

La somme de cette série sera notée $\exp(\phi)$ ou e^ϕ. Dans le cas particulier $E = \mathbb{K}^n$ on obtient l'exponentielle $\exp(A)$ d'une matrice carrée $A \in M_{n,n}(\mathbb{K})$.

Remarque 5.10 *Soient ϕ, $\psi \in \mathrm{End}(E)$ tels que $\phi \circ \psi = \psi \circ \phi$. Alors*

1. $\exp(\phi) \circ \psi = \psi \circ \exp(\phi)$.

2. $\exp(\phi + \psi) = \exp(\phi) \circ \exp(\psi) = \exp(\psi) \circ \exp(\phi)$.

Démonstration: Exercice.

On peut montrer que l'application $\exp : \mathrm{End}(E) \to \mathrm{End}(E)$ ainsi obtenue est de classe $\mathcal{C}^\infty$. Nous aurons besoin seulement la différentiabilité de la composition de cette application avec l'application linéaire $\mathbb{R} \ni t \mapsto t\phi \in \mathrm{End}(E)$ associée à un endomorphisme fixé ϕ.

Proposition 5.11 *Soit $\phi \in \mathrm{End}(E)$. L'application $t \mapsto \exp(t\phi)$ est de classe $\mathcal{C}^\infty$ et*

$$\frac{d^k}{dt^k}\exp(t\phi) = \phi^k \exp(t\phi) = \exp(t\phi)\phi^k.$$

Démonstration: Exercice. Montrer d'abord que la série $\sum_{k\geqslant 0}\frac{1}{k!}(t\phi)^k$ ainsi que les séries obtenues en dérivant terme à terme par rapport à t sont normalement, donc uniformément convergentes sur tout intervalle compact. Utiliser le théorème de dérivation de la somme d'une série convergente de fonctions. ■

L'exponentielle d'un endomorphisme intervient dans la théorie des équations différentielles ordinaires :

Proposition 5.12 *Soient $\phi \in \mathrm{End}(E)$ et $x_0 \in \mathbb{R}$. La solution du problème de Cauchy*

$$\frac{d}{dt}x(t) = \phi(x(t)),\ x(0) = x_0$$

s'écrit

$$x(t) = \exp(t\phi)x_0.$$

Démonstration: Exercice. Utiliser la proposition 5.11.

Question : Comment on calcule explicitement l'exponentielle d'une matrice carrée $A \in M_{n,n}(\mathbb{K})$?

Réponse : 1. Pour une matrice diagonalisable Δ on trouve une base B qui diagonalise ϕ_Δ, donc telle que

$$\Delta = P\begin{pmatrix}\lambda_1 I_{d_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k I_{d_k}\end{pmatrix}P^{-1}$$

où P est la matrice dont les colonnes sont les vecteurs de B. Alors

$$\Delta^j = P\begin{pmatrix}\lambda_1^j I_{d_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k^j I_{d_k}\end{pmatrix}P^{-1},$$

pour tout $j \in \mathbb{N}$ et donc

$$\exp(\Delta) = P\begin{pmatrix}e^{\lambda_1} I_{d_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & e^{\lambda_k} I_{d_k}\end{pmatrix}P^{-1}.$$

2. Pour une matrice nilpotente U on a

$$\exp(U) = \sum_{j=0}^{\iota(U)}\frac{1}{j!}U^j,$$

donc la série qui donne $\exp(U)$ devient une somme finie.

3. Pour une matrice arbitraire A on utilise la décomposition de Dunford $A = \Delta + U$, où Δ est diagonalisable, U est nilpotente et $\Delta U = U\Delta$. Puisque Δ et U commutent, on a

$$\exp(A) = \exp(\Delta + U) = \exp(\Delta)\exp(U) = \exp(U)\exp(\Delta).$$

Exemple : Soit

$$J_\lambda^r := \left.\begin{pmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & (0)\\ & & \ddots & \ddots & \\ & & & \ddots & \ddots\\ & (0) & & \lambda & 1\\ & & & & \lambda\end{pmatrix}\right\} r$$

un bloc de Jordan de taille r et valeur propre λ.

Nous avons

$$tJ_\lambda^r = t\lambda I_r + tJ_0^r.$$

Évidemment $\exp(t\lambda I_r) = e^{t\lambda} I_r$. En utilisant le fait que J_0^r est nilpotente d'indice de nilpotence r, on obtient (exercice) :

$$\exp(tJ_0^r) = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2} \\ & & & \ddots & t \\ 0 & \dots & & & 1 \end{pmatrix},$$

donc

$$\exp(tJ_\lambda^r) = e^{t\lambda} I_r \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2} \\ & & & \ddots & t \\ 0 & \dots & & & 1 \end{pmatrix} = e^{t\lambda} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2} \\ & & & \ddots & t \\ 0 & \dots & & & 1 \end{pmatrix}.$$

Pour $t = 1$ on obtient $\exp(J_\lambda^r)$.

Remarque 5.13 *Supposons que l'endomorphisme $\phi_A \in \mathrm{End}(\mathbb{K}^n)$ associé à A a été jordanisé, donc nous connaissons $P \in \mathrm{GL}(n, \mathbb{K})$ telle que*

$$P^{-1}AP = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & J_k \end{pmatrix},$$

où J_i sont des blocs de Jordan. Alors

$$\exp(A) = P \begin{pmatrix} \exp(J_1) & 0 & \dots & 0 \\ 0 & \exp(J_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \exp(J_k) \end{pmatrix} P^{-1}, \ \exp(tA) = P \begin{pmatrix} \exp(tJ_1) & 0 & \dots & 0 \\ 0 & \exp(tJ_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \exp(tJ_k) \end{pmatrix} P^{-1},$$

donc $\exp(A)$, $\exp(tA)$ s'écrivent facilement en utilisant les formules connues pour $\exp(J_\lambda^r)$, $\exp(tJ_\lambda^r)$.