Réduction des endomorphismes et des matrices

Peter Haïssinsky, Université de Paul Sabatier

2014–2015

Ce chapitre complète le chapitre sur la diagonalisation. On traite de différents théorèmes fondamentaux de réduction. Le texte comporte deux appendices rappelant la notion de matrices des cofacteurs et traitant de l'arithmétique élémentaire des polynômes.

Introduction

On se donne un espace vectoriel E de dimension finie sur un corps commutatif $\mathbb{K}$ (que l'on peut penser comme étant $\mathbb{R}$ ou $\mathbb{C}$) et un endomorphisme $u \in \mathcal{L}(\mathrm{E})$. La question traitée dans ce chapitre est la suivante: quelle est la forme la plus simple pour exprimer u dans une base?

1 Polynômes d'endomorphismes

Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ un polynôme à coefficients dans $\mathbb{K}$.

Considérons un espace vectoriel sur $\mathbb{K}$ et un endomorphisme $u \in \mathcal{L}(\mathrm{E})$. On définit $u^0 = \mathrm{Id}$, et par récurrence, $u^{n+1} = u^n \circ u$ pour $n \geq 0$. Si $\mathrm{P}(\mathrm{X}) = a_d\mathrm{X}^d + \ldots + a_1\mathrm{X} + a_0$, on associe l'endomorphisme $\mathrm{P}(u) \in \mathcal{L}(\mathrm{E})$ définie par

$$\mathrm{P}(u) = a_d u^d + \ldots + a_1 u + a_0\mathrm{Id} = \sum_{j=0}^{d} a_j u^j \,.$$

Remarquons tout de suite que si x est un vecteur propre de u associé à la valeur propre $\lambda \in \mathbb{K}$ alors $u^j(x) = \lambda^j x$ pour tout $j \geq 0$ et

$$\mathrm{P}(u)(x) = a_d\lambda^d x + \ldots + a_1\lambda x + a_0 x = \mathrm{P}(\lambda)x \,. \tag{1.1}$$

Enfin, notons que si P et Q sont deux polynômes, alors $(\mathrm{PQ})(u) = \mathrm{P}(u) \circ \mathrm{Q}(u) = \mathrm{Q}(u) \circ \mathrm{P}(u)$. En effet, si $\mathrm{P}(\mathrm{X}) = \sum a_i\mathrm{X}^i$ et $\mathrm{Q}(\mathrm{X}) = \sum b_j\mathrm{X}^j$ alors

$$\begin{aligned}
\mathrm{P}(u) \circ \mathrm{Q}(u) &= \mathrm{P}(u)\left(\sum_j b_j u^j\right) = \sum_j b_j \mathrm{P}(u)\left(u^j\right) \\
&= \sum_j b_j \left(\sum_i a_i u^i(u^j)\right) = \sum_{i,j} a_i b_j u^{i+j} \\
&= (\mathrm{PQ})(u) \,.
\end{aligned}$$

De même, si $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, $n \geq 1$, on définit $\mathrm{P}(\mathrm{A}) \in \mathcal{M}_n(\mathbb{K})$ en posant

$$\mathrm{P}(\mathrm{A}) = a_d\mathrm{A}^d + \ldots + a_1\mathrm{A} + a_0\mathrm{I}_n = \sum_{j=0}^{d} a_j\mathrm{A}^j \,,$$

où $\mathrm{A}^0 = \mathrm{I}_n$ désigne la matrice identité.

On énonce un résultat fondamental de la théorie:

Lemme 1.1 (des noyaux). — *Soit u un endormorphisme d'un espace vectoriel et soient* P *et* Q *deux polynômes premiers entre eux. Alors*

$$\operatorname{Ker}(\mathrm{PQ})(u) = \operatorname{Ker}\mathrm{P}(u) \oplus \operatorname{Ker}\mathrm{Q}(u)\,.$$

De plus, les projecteurs sur $\operatorname{Ker}\mathrm{P}(u)$ *et* $\operatorname{Ker}\mathrm{Q}(u)$ *parallèlement à* $\operatorname{Ker}\mathrm{Q}(u)$ *et* $\operatorname{Ker}\mathrm{P}(u)$ *respectivement sont des polynômes en u.*

Démonstration. — Comme P et Q sont premiers entre eux, le lemme de Bézout implique l'existence de deux polynômes U et V tels que $\mathrm{PU} + \mathrm{QV} = 1$. Par conséquent, on a $\mathrm{U}(u) \circ \mathrm{P}(u) + \mathrm{V}(u) \circ \mathrm{Q}(u) = \mathrm{Id}$. Donc, si $x \in \mathrm{E}$, alors

$$\mathrm{U}(u) \circ \mathrm{P}(u)(x) + \mathrm{V}(u) \circ \mathrm{Q}(u)(x) = x\,. \tag{1.2}$$

Du coup, si $x \in \operatorname{Ker}\mathrm{P}(u) \cap \operatorname{Ker}\mathrm{Q}(u)$, alors on trouve $x = 0$ donc ces espaces sont en somme directe.

Prenons maintenant $x \in \operatorname{Ker}(\mathrm{PQ})(u)$. On écrit $x_1 = \mathrm{U}(u) \circ \mathrm{P}(u)(x)$ et $x_2 = \mathrm{V}(u) \circ \mathrm{Q}(u)(x)$ de sorte que $x = x_1 + x_2$. On a donc

$$\mathrm{Q}(u)(x_1) = \mathrm{Q}(u) \circ \mathrm{U}(u) \circ \mathrm{P}(u)(x) = \mathrm{U}(u) \circ (\mathrm{PQ})(u)(x) = \mathrm{U}(0) = 0$$

et de même $\mathrm{P}(u)(x_2) = 0$, ce qui montre que $\operatorname{Ker}(\mathrm{PQ})(u) \subset \operatorname{Ker}\mathrm{P}(u) \oplus \operatorname{Ker}\mathrm{Q}(u)$.

Pour l'inclusion réciproque, on prend $x_1 \in \operatorname{Ker}\mathrm{P}(u)$ et $x_2 \in \operatorname{Ker}\mathrm{Q}(u)$ de sorte que

$$\begin{aligned}(\mathrm{PQ})(u)(x_1 + x_2) &= (\mathrm{QP})(u)(x_1) + (\mathrm{PQ})(u)(x_2)\\ &= \mathrm{Q}(u) \circ \mathrm{P}(u)(x_1) + \mathrm{P}(u) \circ \mathrm{Q}(u)(x_2)\\ &= \mathrm{Q}(u)(0) + \mathrm{P}(u)(0) = 0\end{aligned}$$

ce qui montre que $\operatorname{Ker}\mathrm{P}(u) \oplus \operatorname{Ker}\mathrm{Q}(u) \subset \operatorname{Ker}(\mathrm{PQ})(u)$.

On remarque de plus que $\pi : \mathrm{E} \to \operatorname{Ker}\mathrm{P}(u)$ défini par $(\mathrm{VQ})(u)$ est le projecteur sur $\operatorname{Ker}\mathrm{P}(u)$ parallèlement à $\operatorname{Ker}\mathrm{Q}(u)$. D'après (1.2), on a $x = \pi(x) + (\mathrm{Id} - \pi)(x)$ et donc $\pi(x) = \pi^2(x) + 0$. Ceci montre que les projecteurs sont bien des polynômes en u. ∎

On généralise:

Corollaire 1.2. — *Soit u un endormorphisme d'un espace vectoriel et soient* $\mathrm{P}_1, \ldots, \mathrm{P}_n$ *des polynômes premiers entre eux deux à deux. Alors*

$$\operatorname{Ker}(\mathrm{P}_1 \ldots \mathrm{P}_n)(u) = \operatorname{Ker}\mathrm{P}_1(u) \oplus \ldots \oplus \operatorname{Ker}\mathrm{P}_n(u)\,.$$

De plus, chaque projecteur $\pi_j : \mathrm{E} \to \operatorname{Ker}\mathrm{P}_j(u)$ *parallèlement à* $\oplus_{i \neq j} \operatorname{Ker}\mathrm{P}_i$ *est un polynôme en u.*

Démonstration. — On procède par récurrence: le cas $n = 2$ est déjà fait. On suppose le résultat connu jusqu'au rang $n \geq 2$.

Montrons par l'absurde que P_{n+1} est premier avec $\mathrm{P}_1 \ldots \mathrm{P}_n$: si P_{n+1} divise $\mathrm{P}_1 \ldots \mathrm{P}_n$, alors le lemme de Gauss implique que comme P_{n+1} est premier avec P_n alors P_{n+1} divise $\mathrm{P}_1 \ldots \mathrm{P}_{n-1}$. En procédant ainsi, on montre que P_{n+1} doit diviser P_1, ce qui est absurde.

Du coup, le lemme des noyaux implique que

$$\operatorname{Ker}(\mathrm{P}_1 \ldots \mathrm{P}_{n+1})(u) = \operatorname{Ker}(\mathrm{P}_1 \ldots \mathrm{P}_n)(u) \oplus \operatorname{Ker}\mathrm{P}_{n+1}(u)\,.$$

L'hypothèse de récurrence implique

$$\operatorname{Ker}(\mathrm{P}_1 \ldots \mathrm{P}_n)(u) = \operatorname{Ker}\mathrm{P}_1(u) \oplus \ldots \oplus \operatorname{Ker}\mathrm{P}_n(u)\,.$$

Donc

$$\operatorname{Ker}(\mathrm{P}_1 \ldots \mathrm{P}_{n+1})(u) = \operatorname{Ker}(\mathrm{P}_1)(u) \oplus \ldots \oplus \operatorname{Ker}\mathrm{P}_{n+1}(u)\,.$$

La propriété des projecteurs découle facilement de l'hypothèse de récurrence: la projection sur $\operatorname{Ker}\mathrm{P}_{n+1}$ vient du lemme du noyau, et pour les autres les projecteurs s'obtiennent par composition de deux projecteurs qui s'expriment par des polynômes en u. ∎

On suppose dans la suite que E est un espace vectoriel de dimension finie.

1.1 Le théorème de Cayley-Hamilton

On énonce et on démontre le théorème de Cayley-Hamilton avant d'en donner quelques applications.

Théorème 1.3 (Cayley-Hamilton). — *Soit u un endomorphisme de* E *de dimension $n \geq 1$ et notons χ_u son polynôme caractéristique. Alors on a $\chi_u(u) = 0$. Autrement dit, pour tout $x \in \mathrm{E}$, on a $\chi_u(u)(x) = 0$.*

On remarque que si u est diagonalisable, alors (1.1) implique le résultat.

Démonstration. — On se fixe une base $\mathcal{B}$ de E et on note $\mathrm{A} = \mathrm{Mat}(u, \mathcal{B})$. Pour $\lambda \in \mathbb{K}$, on note $\mathrm{B}(\lambda)$ la matrice des cofacteurs de $\mathrm{A} - \lambda \mathrm{I}_n$ de sorte que

$$ {}^t\mathrm{B}(\lambda) \times (\mathrm{A} - \lambda \mathrm{I}_n) = \det(\mathrm{A} - \lambda \mathrm{I}_n)\mathrm{I}_n = \chi_u(\lambda)\mathrm{I}_n \, . $$

Par définition de la matrice de cofacteurs, $\mathrm{B}(\lambda)$ est une matrice à coefficients polynomiaux en λ, de degré $(n-1)$. Par conséquent, on peut trouver des matrices scalaires $\mathrm{B}_0, \ldots, \mathrm{B}_{n-1}$ telles que

$$ {}^t\mathrm{B}(\lambda) = \mathrm{B}_0 + \lambda \mathrm{B}_1 + \ldots + \lambda^{n-1}\mathrm{B}_{n-1} \, . $$

Du coup, on a

$$ (\mathrm{B}_0 + \lambda \mathrm{B}_1 + \ldots + \lambda^{n-1}\mathrm{B}_{n-1})(\mathrm{A} - \lambda \mathrm{I}_n) = \mathrm{B}_0\mathrm{A} + \lambda(\mathrm{B}_1\mathrm{A} - \mathrm{B}_0) + \ldots + \lambda^{n-1}(\mathrm{B}_{n-1}\mathrm{A} - \mathrm{B}_{n-2}) - \lambda^n \mathrm{B}_{n-1} \, . $$

Si on écrit

$$ \chi_u(\lambda) = a_n\lambda^n + \ldots + a_1\lambda + a_0 $$

alors on obtient

$$ \mathrm{B}_0\mathrm{A} + \lambda(\mathrm{B}_1\mathrm{A} - \mathrm{B}_0) + \ldots + \lambda^{n-1}(\mathrm{B}_{n-1}\mathrm{A} - \mathrm{B}_{n-2}) - \lambda^n \mathrm{B}_{n-1} = a_n\lambda^n\mathrm{I}_n + \ldots + a_1\lambda\mathrm{I}_n + a_0\mathrm{I}_n \, . $$

Comme cette identité est valable pour tout $\lambda \in \mathbb{K}$, on peut identifier les coefficients terme à terme:

$$ \left\{ \begin{array}{lcl} \mathrm{B}_0\mathrm{A} & = & a_0\mathrm{I}_n \\ \mathrm{B}_1\mathrm{A} - \mathrm{B}_0 & = & a_1\mathrm{I}_n \\ & \cdots & \\ \mathrm{B}_j\mathrm{A} - \mathrm{B}_{j-1} & = & a_j\mathrm{I}_n \\ & \cdots & \\ -\mathrm{B}_{n-1} & = & a_n\mathrm{I}_n \end{array} \right. $$

Multiplions à droite la $j$ème ligne par A^{j-1}:

$$ \left\{ \begin{array}{lcl} \mathrm{B}_0\mathrm{A} & = & a_0\mathrm{I}_n \\ \mathrm{B}_1\mathrm{A}^2 - \mathrm{B}_0\mathrm{A} & = & a_1\mathrm{A} \\ & \cdots & \\ \mathrm{B}_j\mathrm{A}^{j+1} - \mathrm{B}_{j-1}\mathrm{A}^j & = & a_j\mathrm{A}^j \\ & \cdots & \\ -\mathrm{B}_{n-1}\mathrm{A}^n & = & a_n\mathrm{A}^n \end{array} \right. $$

On somme toutes les lignes afin d'obtenir

$$ \mathrm{B}_0\mathrm{A} + (\mathrm{B}_1\mathrm{A}^2 - \mathrm{B}_0\mathrm{A}) + \ldots + (\mathrm{B}_{n-1}\mathrm{A}^n - \mathrm{B}_{n-2}\mathrm{A}^{n-1}) - \mathrm{B}_{n-1}\mathrm{A}^n = a_n\mathrm{A}^n + \ldots + a_1\mathrm{A} + a_0\mathrm{I}_n $$

soit $0 = \chi_u(\mathrm{A})$. ∎

En particulier, si $\mathrm{A} \in \mathcal{M}_2(\mathbb{K})$, alors $\mathrm{A}^2 - (\mathrm{tr}\,\mathrm{A})\mathrm{A} + (\det \mathrm{A})\mathrm{I}_2 = 0$.

On propose ici quelques applications qui montrent que l'on peut s'en tirer même si la matrice en question n'est pas diagonalisable.

Calcul de puissances. Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $p \in \mathbb{N}$ que l'on suppose très grand. Si $p \geq n$, alors la division euclidienne nous fournit deux polynômes Q et R avec $\deg R \leq n-1$ tels que

$$X^p = \chi_A(X) \times Q(X) + R(X)\,.$$

Par conséquent, en remplaçant X par A, on obtient $A^p = R(A)$.

Une autre approche consiste à dire qu'il existe un polynôme Q de degré au plus $n-1$ tel que $A^n = Q(A)$. Dans ce cas, on obtient la formule de récurrence $A^{p+n} = A^p Q(A)$.

Pour $n = 2$, on a

$$\begin{aligned} A^3 &= A((\operatorname{tr} A)A - (\det A)I_2) \\ &= \operatorname{tr} A((\operatorname{tr} A)A - (\det A)I_2) - (\det A)A \\ &= [(\operatorname{tr} A)^2 - (\det A)]A - (\operatorname{tr} A)(\det A)I_2 \end{aligned}$$

et on trouve par récurrence

$$A^{p+2} = A^p[(\operatorname{tr} A)A - (\det A)I_2]\,.$$

Calcul de l'inverse. Soit $A \in \mathcal{M}_n(\mathbb{K})$ et écrivons $\chi_A(X) = (-1)^n X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + \det A$. Si A est inversible, on a donc $\det A \neq 0$ et

$$\det A \cdot I_n = (-1)^{n+1}A^n + \ldots + a_1 A = A((-1)^{n-1}A^{n-1} + \ldots + a_1 I_n)$$

donc

$$A^{-1} = \frac{1}{\det A}\left((-1)^{n-1}A^{n-1} + \ldots + a_1 I_n\right)\,.$$

1.2 Polynôme minimal

On a vu que si u est un endomorphisme d'un espace vectoriel de dimension finie alors il existe un polynôme P non trivial (le polynôme caractéristique par exemple) tel que $P(u)$ soit l'endomorphisme trivial ($P(u) = 0$).

Dans cette partie on étudie l'annulateur de u:

$$\operatorname{Ann} u = \{P \in \mathbb{K}[X],\ P(u) = 0\}$$

en vue de mieux comprendre la diagonalisabilité de u.

On remarque tout d'abord que $\operatorname{Ann} u$ est un idéal de $\mathbb{K}[X]$. En effet, si $P, Q \in \operatorname{Ann} u$ et $\lambda \in \mathbb{K}$ alors $(P + \lambda Q) \in \operatorname{Ann} u$ et si $R \in \mathbb{K}[X]$ est un polynôme quelconque $P(u) \circ R(u) = 0$ aussi.

Proposition 1.4. — *Il existe un unique polynôme unitaire μ_u de degré minimal dans $\operatorname{Ann} u$. De plus,*

1. *tout autre $P \in \operatorname{Ann} u$ est un multiple de μ_u: il existe $Q \in \mathbb{K}[X]$ tel que $P = Q\mu_u$;*
2. *toute racine de μ_u est valeur propre de u;*
3. *toute valeur propre de u est racine de μ_u.*

Le polynôme μ_u s'appelle *le polynôme minimal* de u. Si $\{\lambda_1, \ldots, \lambda_d\}$ est le spectre de u, alors $(X - \lambda_1)\ldots(X - \lambda_d)$ divise μ_u.

Démonstration. — Soit μ un polynôme unitaire de degré minimal dans $\operatorname{Ann} u$ et montrons que tout autre $P \in \operatorname{Ann} u$ est un multiple de μ: par division euclidienne, il existe Q et R de degré au plus $\deg \mu - 1$ tel que $P = Q\mu + R$. Par conséquent, on a $R(u) = P(u) - Q(u) \circ \mu(u) = 0$, donc $R \in \operatorname{Ann} u$. Par minimalité du degré de μ, on doit avoir $R = 0$. Donc $P = Q\mu$. Montrons maintenant que μ est unique: si ν est un autre tel polynôme, on aurait $\nu = \mu Q$; mais comme $\deg \mu = \deg \nu$, le polynôme Q est constant et comme μ et ν sont unitaires, on trouve $Q = 1$. Du coup, $\mu = \nu$.

Le théorème de Cayley-Hamilton implique que μ_u divise χ_u. Donc les racines de μ_u sont aussi des racines de χ_u, donc des valeurs propres.

Soit λ une valeur propre de u et supposons qu'elle ne soit pas racine de μ_u. Il existe alors des constantes $b_0, \ldots, b_d$, avec $b_0 \neq 0$, telles que

$$\begin{aligned} \mu_u(\mathrm{X}) &= b_0 + b_1(\mathrm{X}-\lambda) + \ldots + b_d(\mathrm{X}-\lambda)^d \\ &= b_0 + (\mathrm{X}-\lambda)[b_1 + \ldots + b_d(\mathrm{X}-\lambda)^{d-1}]\,. \end{aligned}$$

Par conséquent on a

$$0 = \mu_u(u) = b_0\mathrm{Id} + (u-\lambda\mathrm{Id})[b_1\mathrm{Id} + \ldots + b_d(u-\lambda\mathrm{Id})^{d-1}]$$

d'où on tire

$$(u-\lambda\mathrm{Id})\left(\frac{b_1}{b_0}\mathrm{Id} + \ldots + \frac{b_d}{b_0}(u-\lambda\mathrm{Id})^{d-1}\right) = -\mathrm{Id}\,.$$

Mais ceci implique que $(u-\lambda\mathrm{Id})$ est inversible, donc que λ n'est pas valeur propre de u: contradiction. ∎

Le polynôme minimal permet de donner le critère suivant de diagonalisabilité:

Théorème 1.5. — *Soit u un endomorphisme d'un espace vectoriel* E *de dimension finie. Alors u est diagonalisable si et seulement si μ_u est scindé et toutes ses racines sont simples.*

Démonstration. — Supposons d'abord u diagonalisable et considérons une base de vecteurs propres associées aux valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_d$. Notons D la matrice de u dans cette base. On sait que μ_u a pour facteurs $\mathrm{P} = (\mathrm{X}-\lambda_1)\ldots(\mathrm{X}-\lambda_d)$. Comme chaque espace propre est invariant par u, on vérifie que $(\mathrm{D}-\lambda_1\mathrm{I}_{m_1})\ldots(\mathrm{D}-\lambda_d\mathrm{I}_{m_d}) = 0$, donc $\mathrm{P} \in \operatorname{Ann} u$. Comme c'est un facteur de μ_u et qu'il est unitaire, on en déduit que $\mathrm{P} = \mu_u$. Donc les racines sont simples.

Passons à la réciproque. Supposons que μ_u est scindé à racines simples. Les racines de μ_u sont donc l'ensemble des valeurs propres $\lambda_1, \ldots, \lambda_d$:

$$\mu_u(\mathrm{X}) = (\mathrm{X}-\lambda_1)\ldots(\mathrm{X}-\lambda_d)$$

et les facteurs $(\mathrm{X}-\lambda_j)$ sont premiers entre eux. Donc le lemme des noyaux implique

$$\mathrm{E} = \operatorname{Ker}\mu_u(u) = \operatorname{Ker}(u-\lambda_1\mathrm{Id}) \oplus \ldots \oplus \operatorname{Ker}(u-\lambda_d\mathrm{Id}),$$

ce qui signifie que E est somme directe des espaces propres de u, donc que u est diagonalisable. ∎

Corollaire 1.6. — *Soit u un endomorphisme défini sur un espace vectoriel de dimension finie. S'il existe un polynôme scindé à racines simples qui annule u, alors u est diagonalisable.*

Démonstration. — Soit $\mathrm{P} \in \operatorname{Ann} u$ scindé à racines simples. Par définition du polynôme minimal, μ_u divise P, donc μ_u est aussi scindé à racines simples, d'après le théorème de Gauss. Du coup, le théorème précédent montre que u est diagonalisable. ∎

1.3 Polynômes minimaux associés à des vecteurs

Soit $u \in \mathcal{L}(\mathrm{E})$ un endomorphisme d'un espace vectoriel de dimension finie. Soit $x \in \mathrm{E}$, et notons

$$\operatorname{Ann}(x,u) = \{\mathrm{P} \in \mathbb{K}[\mathrm{X}].\ \mathrm{P}(u)(x) = 0\}$$

l'annulateur de x relatif à u. On vérifie comme dans le cas de $\operatorname{Ann} u$ que c'est un idéal, donc il existe un unique polynôme unitaire μ_x, *le polynôme minimal de x relativement à u*, tel que tout $\mathrm{P} \in \operatorname{Ann}(x,u)$ est divisible par μ_x.

Proposition 1.7. — *Il existe $x \in \mathrm{E}$ tel que $\mu_u = \mu_x$.*

DÉMONSTRATION. — Par définition, $\mu_u(u)(x) = 0$ pour tout x, donc μ_x divise μ_u. Or μ_u s'écrit comme un produit de polynômes irréductibles. Donc le nombre de polynômes apparaissant comme polynômes minimaux relativement à u est fini. Notons $x_1, \ldots, x_k$ des représentants. Du coup, pour tout $x \in \mathrm{E}$, il existe $j \in \{1, \ldots, k\}$, tel que $\mu_x = \mu_{x_j}$, ce qui implique d'une part que $\mu_{x_j}(x) = 0$ et donc

$$\mathrm{E} = \bigcup_{1 \leq j \leq k} \operatorname{Ker} \mu_{x_j}(u)\,.$$

Or comme $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, il existe j tel que $\mathrm{E} = \operatorname{Ker} \mu_{x_j}(u)$ (voir le lemme suivant). Du coup, on a $\mu_{x_j}(u)(x) = 0$ pour tout x, donc μ_u divise $\mu_{x_j}(u)$: ils ont donc même degré et, étant unfitaires, ils sont égaux. ∎

Lemme 1.8. — *Soit* E *un espace vectoriel sur un corps* $\mathbb{K}$ *de cardinal infini. S'il existe des sous-espaces* $\mathrm{F}_1, \ldots, \mathrm{F}_k$ *tels que* $\mathrm{E} = \cup \mathrm{F}_j$*, alors il existe un indice* j *tel que* $\mathrm{E} = \mathrm{F}_j$*.*

DÉMONSTRATION. — On peut supposer que la collection $\mathrm{F}_1, \ldots, \mathrm{F}_k$ est minimale et $k \geq 2$. Prenons $x \in \mathrm{F}_1 \setminus \cup_{2 \leq j \leq k} \mathrm{F}_j$ et $y \notin \mathrm{F}_1$. Du coup, pour tout $\lambda \in \mathbb{K}$, le vecteur $\lambda x + y \notin \mathrm{F}_1$, car sinon on aurait $y \in \mathrm{F}_1$ aussi. Donc, pour chaque λ, ce vecteur est dans l'un des F_j, $2 \leq j \leq k$. Or comme $\mathbb{K}$ est infini, on peut trouver deux scalaires $\lambda_1 \neq \lambda_2$ tels que $\lambda_1 x + y, \lambda_2 x + y$ sont dans le même F_j. Du coup leur différence $(\lambda_1 x + y) - (\lambda_2 x + y) = (\lambda_1 - \lambda_2)x$ est aussi dans F_j. Comme $\lambda_1 \neq \lambda_2$, on en déduit $x \in \mathrm{F}_j$, ce qui est contraire à notre hypothèse de travail. Donc $k = 1$. ∎

2 Critères de diagonalisabilité

En reprenant les résultats vus précédemment, on a les critères suivants qui assurent qu'un endomorphisme u défini sur un espace de dimension finie est diagonalisable. Les propriétés suivantes sont équivalentes:

1. L'endomorphisme u est diagonalisable.
2. Il existe une base de vecteurs propres.
3. On a $\sum_{\lambda \in \mathcal{S}(u)} \dim \mathrm{E}_\lambda = \dim \mathrm{E}$.
4. On a $\mathrm{E} = \oplus_{\lambda \in \mathcal{S}(u)} \mathrm{E}_\lambda$.
5. Le polynôme caractéristique χ_u est scindé sur $\mathbb{K}$ et, pour chaque $\lambda \in \mathcal{S}(u)$, on a $\dim \mathrm{E}_\lambda = m_\lambda$.
6. Il existe un polynôme scindé P à racines simples tel que $\mathrm{P}(u) = 0$.

3 Matrices compagnons et espaces monogènes

Soit $\mathrm{P}(\mathrm{X}) = \mathrm{X}^d + a_{d-1}\mathrm{X}^{d-1} + \ldots + a_0$ un polynôme de degré $d \geq 1$. On associe sa *matrice compagnon*

$$\mathrm{C}_\mathrm{P} = \mathrm{C}(\mathrm{P}) = \begin{pmatrix} 0 & \ldots & \ldots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{d-2} \\ 0 & \ldots & 0 & 1 & -a_{d-1} \end{pmatrix}$$

Proposition 3.1. — *Le polynôme caractéristique de la matrice compagnon d'un polynôme* P *est* $(-1)^d \mathrm{P}$*. Le polynôme minimal est* P*.*

Démonstration. — On calcule $\det(\mathrm{C_P}-\lambda \mathrm{I}_d)$ en développant par rapport à la dernière colonne: on écrit les mineurs de $(\mathrm{C_P}-\lambda \mathrm{I}_d)$ de la dernière colonne par blocs, avec le premier bloc de taille $(j-1)\times(j-1)$:

$$\mathrm{A}_{j,d} = \det \begin{pmatrix} -\lambda & 0 & \dots & 0 & & & & \\ 1 & \ddots & 0 & 0 & & & 0 & \\ 0 & \ddots & \ddots & 0 & & & & \\ 0 & & 1 & -\lambda & & & & \\ & & & & 1 & -\lambda & 0 & 0 \\ & & & & 0 & \ddots & \ddots & 0 \\ & & 0 & & \vdots & 0 & \ddots & -\lambda \\ & & & & 0 & \dots & 0 & 1 \end{pmatrix} = (-\lambda)^{j-1}$$

Du coup, en développant par rapport à la dernière colonne, on trouve

$$\chi(\lambda) = \sum_{j=1}^{d-1} (-1)^{j+d}(-a_{j-1})(-\lambda)^{j-1} + (-1)^{2d}(-a_{d-1}-\lambda)(-\lambda)^{d-1} = (-1)^d \mathrm{P}(\lambda)\,.$$

Pour montrer que $\mu_u = \mathrm{P}$, il suffit d'exhiber un vecteur x tel que $(x, u(x), \dots, u^{d-1}(x))$ est libre. Dans ce cas, aucun polynôme de degré inférieur à $d-1$ ne pourra annuler u, donc on aura $\deg \mu_u = d$, et comme il est unitaire et comme il doit être un multiple de χ_u, ce sera forcément P.

D'après la forme de $\mathrm{C_P}$, il suffit de choisir $x = \mathrm{e}_1$. ∎

Définition 3.2. — *Soient* E *un espace vectoriel de dimension finie et* $u : \mathrm{E} \to \mathrm{E}$ *un endomorphisme. On dit qu'un sous-espace* F *de* E *est u-monogène si* F *est invariant et s'il existe* $x \in \mathrm{F}$ *tel que, pour* $y \in \mathrm{F}$*, il existe un polynôme* $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ *tel que* $y = \mathrm{P}(u)(x)$*.*

Proposition 3.3. — *Si* F *est un sous-espace u-monogène de dimension* $k \geq 1$*, alors il existe* $x \in \mathrm{F}$*, tel que* $(x, u(x), \dots, u^{k-1}(x))$ *est une base de* F*. Dans cette base, la matrice de* $u|_{\mathrm{F}}$ *est la matrice compagnon de son polynôme caractéristique.*

Démonstration. — Par définition, tout vecteur de F s'écrit de la forme $\mathrm{P}(u)(x)$ pour un certain x donné. Comme F est de dimension k, la famille $(x, u(x), u^2(x), \dots, u^k(x))$ est liée (car elle contient $k+1$ éléments). Il existe donc $\lambda_0, \dots, \lambda_{k-1}$ tels que

$$u^k(x) = \sum_{j=0}^{k-1} \lambda_j u^j(x)\,.$$

Notons $\mu_x(\mathrm{X}) = \mathrm{X}^k - \sum_{j=0}^{k-1} \lambda_j \mathrm{X}^j$.

Soit $y \in \mathrm{F}$ et soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ tels que $y = \mathrm{P}(u)(x)$. Faisons la division euclidienne de P par μ_x de sorte que $\mathrm{P} = \mathrm{Q}\mu_x + \mathrm{R}$, avec $\deg \mathrm{R} < k$. Du coup, on a

$$y = \mathrm{P}(u)(x) = \mathrm{Q}(u)\mu_x(u)(x) + \mathrm{R}(u)(x) = \mathrm{R}(u)(x)\,.$$

Donc $(x, u(x), u^2(x), \dots, u^{k-1}(x))$ engendre F et c'est donc une base puisque son cardinal vaut $\dim \mathrm{F}$.

On vérifie facilement que la matrice de $u|_{\mathrm{F}}$ dans cette base est la matrice compagnon de μ_x. ∎

Proposition 3.4. — *Un sous-espace invariant* F *par un endomorphisme* $u \in \mathcal{L}(\mathrm{E})$ *est monogène si et seulement si on a* $\chi_{u|_{\mathrm{F}}} = (-1)^{\dim \mathrm{F}} \mu_{u|_{\mathrm{F}}}$.

Démonstration. — On a déjà vu que dans un espace monogène, on avait l'égalité $\chi_{u|_{\mathrm{F}}} = (-1)^{\dim \mathrm{F}} \mu_{u|_{\mathrm{F}}}$. Passons à la réciproque et supposons $\chi_{u|_{\mathrm{F}}} = (-1)^{\dim \mathrm{F}} \mu_{u|_{\mathrm{F}}}$. D'après la proposition 1.7, il existe $x \in \mathrm{F}$ tel que $\mu_{u|_{\mathrm{F}}} = \mu_x$ et $\deg \mu_x = \dim \mathrm{F}$. Or on a $\mu_x(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \ldots + a_0$ et donc

$$u^n(x) = -a_{n-1}u^{n-1}(x) - \ldots - a_0 x\,.$$

De plus, $(x, \dots, u^{n-1}(x))$ est libre, car sinon on aurait un polynôme annulateur de x de degré strictement plus petit que μ_x. Cela suffit pour conclure. ∎

4 Décomposition de Frobenius et invariants de similitude

On démontre ici le résultat central duquel on déduira les autres théorèmes de réductions. Aucune hypothèse n'est demandé sur l'endomorphisme u, en particulier sur ses polynômes caractéristique et minimal. De plus, il permet de caractériser complètement les classes de similitude, cf. le corollaire 4.2.

Théorème 4.1 (de décomposition de Frobenius). — *Soit u un endomorphisme sur un espace vectoriel E de dimension finie sur $\mathbb{K}$. Il existe des sous-espaces $\mathrm{E}_1, \dots, \mathrm{E}_k$ de E invariants par u tels que*

1. *chaque sous-espace est monogène;*
2. *on a $\mathrm{E} = \oplus_{1\leq j\leq k}\mathrm{E}_j$;*
3. *si on note les polynômes minimaux $\mathrm{P}_1, \dots, \mathrm{P}_k$ de $\mathrm{E}_1, \dots, \mathrm{E}_k$, alors P_{j+1} divise P_j et cette suite ne dépend pas de la décomposition.*

En particulier, on peut trouver une base $\mathcal{B}$ de E telle que la matrice de u dans cette base est une matrice diagonale par blocs dont chaque bloc est la matrice compagnon de P_j. De plus, on a $\mu_u = \mathrm{P}_1$ et $\chi_u = \prod \mathrm{P}_j$.

Les polynômes $\mathrm{P}_1, \dots, \mathrm{P}_k$ obtenus dans le théorème sont les *invariants de similitude de u.*

Démonstration. — On montre d'abord l'existence.

On procède par récurrence sur la dimension de E. Le cas de la dimension 1 ne pose pas de problème. Supposons donc le théorème vrai jusqu'à la dimension $n-1$ et supposons E de dimension n. La proposition 1.7 nous fournit $v \in \mathrm{E}$ tel que $\mu_u = \mu_v$. On note E_1 l'espace monogène engendré par v de dimension $d \leq n$. Si $d = n$, alors le théorème est montré grâce à la proposition 3.4. Si $d < n$, il suffit de trouver un supplémentaire à E_1 invariant par u pour montrer l'existence de la décomposition,.

Pour cela, on prend la base $(\mathrm{e}_1, \dots, \mathrm{e}_d) = (v, u(v), \dots, u^{d-1}(v))$ de E_1 que l'on complète en une base $\mathcal{B}$ de E et on considère la base duale $\mathcal{B}^*$. On note

$$\mathrm{F} = \{x \in \mathrm{E},\ \mathrm{e}_d^*(u^k(x)) = 0 \text{ pour tout } k \geq 0\} = \bigcap_{k\geq 0} \operatorname{Ker} \mathrm{e}_d^* \circ u^k\,.$$

Cet ensemble s'exprime comme une intersection de sous-espaces vectoriels, donc c'est un sous-espace de E.

Si $x \in \mathrm{F}$, alors pour tout $k \geq 0$, on a $\mathrm{e}_d^* \circ u^k(u(x)) = \mathrm{e}_d^* \circ u^{k+1}(x) = 0$ donc $u(x) \in \mathrm{F}$, et $u(\mathrm{F}) \subset \mathrm{F}$.

Soit $x \in \mathrm{E}_1 \cap \mathrm{F}$, alors on peut écrire $x = \sum_{j=0}^{d-1} \lambda_j u^j(v)$ donc

$$\lambda_j = \mathrm{e}_d^*(u^{d-1-j}(x)) = 0$$

donc $x = 0$ et $\mathrm{E}_1 \cap \mathrm{F} = \{0\}$.

Enfin, on utilise ici un argument de dualité pour montrer que $\dim \mathrm{F} + \dim \mathrm{E}_1 = \dim \mathrm{E}$. Remarquons d'abord que $\mathbb{K}[u]$ est un sous-espace vectoriel de $\mathcal{L}(\mathrm{E})$ de dimension d, car seul le reste de la division euclidienne d'un polynôme par μ_u importe dans $\mathbb{K}[u]$.

On considère $\mathrm{T} : \mathbb{K}[u] \to \mathrm{E}^*$, $f \mapsto \mathrm{e}_d^* \circ f$. Si $\mathrm{T}(f) = 0$, alors cela signifie que $f(\mathrm{E}) \subset \mathbb{K}u^{d-1}(v)$. En écrivant $f = \sum_{0\leq j\leq d-1} \lambda_j u^j$, on obtient, pour $1 \leq p \leq d-1$,

$$0 = \mathrm{e}_d^* f(u^p(v)) = \mathrm{e}_d^* \left(\sum \lambda_j \mathrm{e}_d^* \circ u^{j+p}(v)\right) = \lambda_{d-1-p}\,.$$

Donc T est injective et $\dim \operatorname{Im} \mathrm{T} = d$. Or on a

$$\mathrm{F} = \{x \in \mathrm{E},\ \varphi(x) = 0 \text{ pour tout } \varphi \in \operatorname{Im} \mathrm{T}\}$$

et cela implique justement $\dim \mathrm{F} + \dim \operatorname{Im} \mathrm{T} = \dim \mathrm{E}$, c'est-à-dire $\dim \mathrm{F} + \dim \mathrm{E}_1 = \dim \mathrm{E}$ puisque $\dim \mathrm{E}_1 = d$.

Comme $\mathrm{E} = \mathrm{E}_1 \oplus \mathrm{F}$ et $u(\mathrm{F}) \subset \mathrm{F}$, on peut considérer le polynôme minimal P_2 de $u|_{\mathrm{F}}$. Comme μ_u annule u globalement, on en déduit que P_2 divise μ_u. L'hypothèse de récurrence permet de conclure à l'existence de la décomposition.

Chaque sous-espace étant monogène, on trouve une base de ces sous-espaces de sorte que la matrice de la restriction de u soit $\mathrm{C}(\mathrm{P}_i)$. On vérifie sans mal le reste des conclusions.

Passons maintenant à l'unicité des invariants de similitude. Faisons quelques remarques préliminaires pour une décomposition donnée.

(R1) Le polynôme P_j est le polynôme minimal de la restriction de u à $\mathrm{E}_j \oplus \ldots \oplus \mathrm{E}_k$, car si $\deg \mathrm{P} < \deg \mathrm{P}_j$, alors $\mathrm{P}(u)|_{\mathrm{E}_j} \neq 0$.

(R2) Si $\varphi \in \mathbb{K}[u]$, alors $\varphi(\mathrm{E}) = \oplus \varphi(\mathrm{E}_j)$ car chaque sous-espace est invariant.

On suppose que l'on a deux décompositions (E_j) et (F_i) avec polynômes (P_j) et (Q_i) qui vérifient les conclusions du théorème. On montre par récurrence l'égalité des polynômes. On a $\mathrm{P}_1 = \mathrm{Q}_1 = \mu_u$ par (R1). Supposons l'égalité vérifiée jusqu'au rang $k-1$.

Par (R1), on a $\mathrm{P}_k(u)(\mathrm{E}_j) = 0$ pour tout $j \geq k$. Donc (R2) donne

$$\dim \mathrm{P}_k(u)(\mathrm{E}) = \sum_{1 \leq j < k} \dim \mathrm{P}_k(u)(\mathrm{E}_j) = \sum_{j \geq 1} \dim \mathrm{P}_k(u)(\mathrm{F}_i)\,. \tag{4.1}$$

Or, pour $j < k$, comme $\mathrm{P}_j = \mathrm{Q}_j$, il existe des bases $\mathcal{B}_j$ et $\mathcal{C}_j$ de E_j et F_j de sorte que les matrices des restrictions de u à E_j et F_j sont les mêmes. Du coup, l'isomorphisme $\varphi_j : \mathrm{E}_j \to \mathrm{F}_j$ qui transforme $\mathcal{B}_j$ en $\mathcal{C}_j$ conjugue $u|_{\mathrm{E}_j}$ à $u|_{\mathrm{F}_j}$: on a $\varphi_j \circ u|_{\mathrm{E}_j} = u|_{\mathrm{F}_j} \circ \varphi_j$, et on peut en déduire $\dim \mathrm{P}_k(u)(\mathrm{E}_j) = \dim \mathrm{P}_k(u)(\mathrm{F}_j)$. Avec (4.1), cela entraîne $\mathrm{P}_k(u)(\mathrm{F}_i) = \{0\}$ pour tout $i \geq k$, donc $\mathrm{P}_k(u)$ annule $\oplus_{i \geq k} \mathrm{E}_i$. Par (R1), on vient d'établir que Q_k divise P_k. Par symétrie, P_k divise aussi Q_k, donc $\mathrm{P}_k = \mathrm{Q}_k$ car ils sont unitaires. ∎

Corollaire 4.2. — *Deux endomorphismes sont conjugués si et seulement si ils ont les mêmes invariants de similitude.*

Démonstration. — Si $u = \varphi \circ v \circ \varphi^{-1}$, alors on peut trouver une base $\mathcal{B}$ tel que $\mathrm{Mat}(u, \mathcal{B})$ soit donnée par le théorène 4.1. Du coup, $\mathrm{Mat}(v, \varphi(\mathcal{B})) = \mathrm{Mat}(u, \mathcal{B})$. Donc les invariants sont identiques.

Réciproquement, notons $\mathcal{B}_u$ et $\mathcal{B}_v$ des bases qui vérifient les conclusions du théorème 4.1 de sorte que $\mathrm{Mat}(u, \mathcal{B}_u) = \mathrm{Mat}(v, \mathcal{B}_v)$. On construit un isomorphisme φ tel que $\varphi(\mathcal{B}_u) = \mathcal{B}_v$. On vérifie que $\varphi \circ u \circ \varphi^{-1} = v$. ∎

5 Autres réductions

Avant de présenter les décompositions de Jordan et de Dunford, on applique le théorème précédent aux endomorphismes nilpotents et on étudie les endomorphismes diagonalisables qui commutent.

5.1 Endomorphismes nilpotents

Définition 5.1. — *Un endomorphisme nilpotent est un endomorphisme u tel que qu'il existe $k \geq 1$ pour lequel $u^k = 0$.*

On en déduit que le polynôme minimal est de la forme $\mu_u(\lambda) = \lambda^k$, avec $1 \leq k \leq \dim \mathrm{E}$. Dans ce cas, les invariants de similitude sont des $\mathrm{P}_i(\lambda) = \lambda^{k_i}$ et les blocs sont de la forme

$$\begin{pmatrix} 0 & \ldots & \ldots & 0 & 0 \\ 1 & 0 & & \vdots & \vdots \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}.$$

Inversement, si $\mu_u(\lambda) = \lambda^k$, alors $u^k = 0$ par définition, donc u est nilpotent. Comme 0 est la seule racine de μ_u, c'est aussi la seule racine de χ_u, et $\chi_u(\lambda) = (-\lambda)^{\dim \mathrm{E}}$.

5.2 Endomorphismes diagonalisables qui commutent

L'objet de cette partie est la proposition suivante:

Proposition 5.2. — *Deux endormorphismes diagonalisables qui commutent sont simultanément diagpnalisables: si u et u' sont diagonalisables et si $u \circ u' = u' \circ u$, alors il existe une base de E telle que les matrices de u et u' sont diagonales.*

On commence par un lemme:

Lemme 5.3. — *Soit u un endomorphisme diagonalisable. Si $\mathrm{F} \subset \mathrm{E}$ est un sous-espace invariant, alors $u|_\mathrm{F}$ est aussi diagonalisable.*

Démonstration. — Si u est diagonalisable, alors son polynôme minimal est scindé et à racines simples. Par restriction à F, on a aussi $\mu_u(u|_\mathrm{F}) = 0$ donc $u|_\mathrm{F}$ admet un polynôme annulateur scindé à racines simples, ce qui implique aussi que $u|_\mathrm{F}$ est diagonalisable. ∎

Démonstration. — (proposition) Soit λ une valeur propre de u; montrons que $u'(\mathrm{E}_\lambda) \subset \mathrm{E}_\lambda$. Si $x \in \mathrm{E}_\lambda$, on veut montrer que $u(u'(x)) = \lambda u'(x)$. On utilise que les endomorphismes commutent: $u(u'(x)) = u'(u(x)) = u'(\lambda x) = \lambda u'(x)$. Par conséquent, u' laisse invariant la décomposition de E en sous-espaces propres de u. D'après le lemme précédent, chaque restriction de u' est diagonalisable donc on obtient, pour chaque espace propre de u, une base de vecteurs propres à la fois pour u et pour u'. ∎

5.3 Décomposition de Jordan

On définit, pour $n \geq 1$ et $\lambda \in \mathbb{K}$, $\mathrm{J}_1 = 0$ et, pour $n \geq 2$,

$$\mathrm{J}_n = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix} \quad \text{et} \quad \mathrm{J}_n(\lambda) = \mathrm{J}_n + \lambda \mathrm{I}_n = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda \end{pmatrix}$$

Théorème 5.4 (de réduction de Jordan). — *Soit $u : \mathrm{E} \to \mathrm{E}$ un endomorphisme sur un espace vectoriel de dimension finie dont le polynôme caractéristique est scindé. Alors il existe une base $\mathcal{B}$ de E telle que la matrice de u soit diagonale par blocs et chaque bloc est de la forme $\mathrm{J}_n(\lambda)$.*

Dans l'énoncé, une valeur propre peut apparaître plusieurs fois. Lorsque u est diagonalisable, on n'obtient que des blocs de dimension 1.

Démonstration. — Si u est nilpotent, c'est-à-dire il existe $k \geq 1$ tel que $u^k = 0$, alors le théorème découle de la décomposition de Frobenius, où on inverse les bases de chaque bloc. Au lieu de prendre des bases de la forme $(x, u(x), \dots, u^\ell(x))$, on prend $(u^\ell(x), \dots, u(x), x)$. On obtient ainsi une matrice diagonale par blocs où chaque bloc est de la forme J_n.

Dans le cas général, on a $\chi_u = \prod(\lambda - \lambda_j)^{m_j}$, donc le lemme des noyaux et le théorème de Cayley-Hamilton impliquent la décomposition

$$\mathrm{E} = \oplus \mathrm{Ker}\,(u - \lambda_j \mathrm{Id})^{m_j}\,.$$

Les espaces $\tilde{\mathrm{E}}(\lambda_j) = \mathrm{Ker}\,(u - \lambda_j \mathrm{Id})^{m_j}$ sont les *espaces caractéristiques* de u. On vérifie qu'ils sont invariants par u: si $x \in \mathrm{Ker}\,(u - \lambda_j \mathrm{Id})^{m_j}$, alors $u^{m_j}(u(x)) = u^{m_j+1}(x) = u(u^{m_j}(x)) = 0$.

Sur chaque $\tilde{\mathrm{E}}(\lambda_j)$, la restriction de $u - \lambda_j \mathrm{Id}$ est nilpotente et sa matrice est diagonale par blocs de blocs de la forme J_n, donc la matrice de u dans $\tilde{\mathrm{E}}(\lambda_j)$ est de la forme $\mathrm{J}_n(\lambda_j)$. ∎

On peut facilement calculer le polynôme minimal à partir de la décomposition de Jordan. Pour chaque $\lambda_j \in \mathcal{S}(u)$, on a en facteur $(\lambda - \lambda_j)^{k_j}$ où k_j est la taille maximale des blocs $\mathrm{J}_k(\lambda_j)$.

5.4 Décomposition de Dunford

Théorème 5.5 (de réduction de Dunford). — *Soit $u : \mathrm{E} \to \mathrm{E}$ un endomorphisme sur un espace vectoriel de dimension finie dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphismes tels que*

1. *$u = d + n$;*
2. *d est diagonalisable et n est nilpotent;*

3. $n \circ d = d \circ n$.

Démonstration. — Rappelons que pour obtenir la décomposition de Jordan, on a appliqué le lemme des noyaux pour décomposer E en sous-espaces caractéristiques $\tilde{\mathrm{E}}_j$, $j \in \{1, \dots, k\}$. Du coup, les projecteurs $\pi_j : \mathrm{E} \to \tilde{\mathrm{E}}_j$ parallèlement à $\oplus_{i \neq j} \tilde{\mathrm{E}}_i$ sont donnés par des polynômes en u. Comme $\mathrm{E} = \oplus \tilde{\mathrm{E}}_j$, on a $\mathrm{Id} = \sum \pi_j$. Posons $d = \sum \lambda_j \pi_j$. C'est un endomorphisme diagonalisable: dans la base donnée par la décomposition de Jordan, la matrice de d est la partie diagonale de la matrice de u. Du coup, si on pose $n = u - d$, alors n est nilpotent: dans cette même base, on peut montrer que le polynôme caractéristique de n est $(-\lambda)^{\dim \mathrm{E}}$, donc le théorème de Cayley-Hamilton permet de conclure.

Les deux endomorphismes commutent car d s'exprime comme un polynôme en u et donc n aussi.

Passons maintenant à l'unicité. On suppose que l'on a deux décompositions $u = n + d = n' + d'$. On remarque que comme n et d commutent, il en est de même de u et d et de u et n; même chose avec d' et n'. Du coup, si d et n proviennent de la décomposition précédente, d et n commutent aussi avec d' et n' car ce sont des polynômes en u.

On écrit $d - d' = n' - n$. Comme n et n' commutent, on peut appliquer la formule du binôme de Newton à $(n' - n)$. Posons $\mathrm{N} = 2 \dim \mathrm{E}$: on a

$$(n' - n)^{\mathrm{N}} = \sum_{k=1}^{\mathrm{N}} \mathrm{C}_{\mathrm{N}}^{k} (n')^k n^{\mathrm{N}-k} = 0$$

car $n^{\dim \mathrm{E}} = (n')^{\dim \mathrm{E}} = 0$ et $\max\{k, \mathrm{N} - k\} \geq \dim \mathrm{E}$ pour tout $k \in \{0, \dots, \mathrm{N}\}$.

Par ailleurs, comme d et d' sont diagonalisables et qu'ils commutent, il existe une base qui diagonalise simultanément d et d' de sorte qu'ils ont pour matrice $\mathrm{diag}(\lambda_j)$ et $\mathrm{diag}(\lambda'_j)$. Du coup, $(d - d')^{\mathrm{N}} = 0$ ce qui signifie que $(\lambda_j - \lambda'_j)^{\mathrm{N}} = 0$ pour tout j c'est-à-dire $d = d'$. Il vient $n = n'$ aussi. ∎

L'application centrale de la décomposition de Dunford concerne le calcul de l'exponentielle d'un endomorphisme. En effet, à partir de la formule

$$\exp u = \sum_{n \geq 0} \frac{u^n}{n!}$$

on obtient, en écrivant $u = d + n$ et en utilisant que $d \circ n = n \circ d$,

$$\exp u = \exp d \sum_{k=0}^{\dim \mathrm{E}} \frac{n^k}{k!} \,.$$

On rappelle que l'exponentielle d'une matrice diagonale est la matrice diagonale dont les éléments de la diagonale sont les exponentielles des coefficients de la matrice initiale.

A Matrice des cofacteurs

Soit $\mathrm{A} = (a_{i,j})$ une matrice carrée d'ordre $n \geq 1$, On désigne par $\mathrm{M}_{i,j}$ la matrice carrée de taille $n - 1$ obtenue en supprimant de A sa $i$ème ligne et sa $j$ème colonne; le déterminant $\det \mathrm{M}_{i,j}$ s'appelle le mineur du coefficient $a_{i,j}$; le mineur "signé" $(-1)^{i+j} \det \mathrm{M}_{i,j} = \mathrm{A}_{i,j}$ est appelé le cofacteur de $a_{i,j}$.

On appelle comatrice de A, ou matrice des cofacteurs de A, la matrice $cof(\mathrm{A})$ dont le terme (i, j) est le cofacteur du terme $a_{i,j}$ de A. Notons que $cof(\mathrm{A})$ est de même format que A.

On a

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

En particulier, lorsque A est inversible, on a

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

B Arithmétique des polynômes

Les polynômes à coefficients dans un corps jouissent de nombreuses propriétés en commun avec les entiers relatifs. Le point de départ est l'analogue de la division euclidienne. On établit ainsi quelques propriétés qui découlent de l'existence de cette division euclidienne.

Théorème B.1. — *Soient* A *un élément de* $\mathbb{K}[X]$ *et* B *un élément* non nul *de* $\mathbb{K}[X]$*. Alors il existe des polynômes* Q *et* R*, déterminés de façon unique, tels que*

$$A = QB + R,\ \deg(R) < \deg(B).$$

On appelle Q *le* quotient *de la* division euclidienne *de* A *par* B*, et* R *le* reste *de cette division.*

Dans $\mathbb{Z}$, cela se lirait $a = bq + r$, avec $0 \leq r < b$.

Démonstration. — Nous commencerons par démontrer l'unicité du couple (Q, R). Soit donc (Q', R') deux autres polynômes vérifiant les conditions du théorème, différents de (Q, R). On a clairement $Q \neq Q'$, puisque sinon on aurait aussi $R = R'$, et l'on obtient, en soustrayant les deux divisions :

$$B(Q - Q') = R' - R.$$

Cela implique comme condition sur les degrés que

$$\deg(B) + \deg(Q - Q') = \deg(R' - R)$$

et donc

$$\deg(R' - R) \geq \deg(B).$$

Mais d'autre part

$$\deg(R' - R) \leq \max\{\deg(R'), \deg(R)\} < \deg(B),$$

ce qui est une contradiction. On a donc unicité de la division euclidienne.

Passons maintenant à la démonstration de l'existence. Si A est nul, on prend $Q = R = 0$. Nous supposerons donc A non nul et raisonnerons par récurrence sur le degré d de A. Posons $A = a_0 + \cdots + a_d X^d$ et $B = b_0 + \cdots + b_p X^p$, avec a_d, b_p non nuls. Pour $d = 0$, A est un polynôme constant et il suffit de poser $Q = 0$ et $R = a_0$. Supposons maintenant la propriété vraie pour tout polynôme de degré plus petit qu'un certain $d > 0$. Soit alors A un polynôme de degré $d + 1$,

$$A = a_0 + \cdots + a_{d+1} X^{d+1}.$$

Posons $Q_1 = (a_{d+1}/b_p) X^{n+1-p}$, et soit $A_1 = A - BQ_1$. C'est un polynôme de degré plus petit que d, et on peut donc lui appliquer l'hypothèse de récurrrence : il existe (Q_2, R_2) avec $\deg(R_2) < p$ tels que $A_1 = BQ_2 + R_2$. Remplaçons alors A_1 par sa définition :

$$A = A_1 + BQ_1 = B(Q_1 + Q_2) + R_2,$$

et le couple $(Q_1 + Q_2, R_2)$ satisfait les conclusions du théorème. ∎

Le calcul pratique de la division euclidienne se fait comme dans le cas de $\mathbb{Z}$: on pose la division de la même manière.

Exemple. Soit à diviser $A = X^5 + 2X^3 - 3X - 2$ par $B = X^3 + X + 1$. Posons la division, comme à l'école primaire :

$$\begin{array}{cccc|c} X^5 & +2X^3 & -3X & -2 & \underline{X^3 + X + 1} \\ X^3 & -X^2 & -3X & -2 & X^2 + 1 \\ & -X^2 & -4X & -3 & \end{array}$$

et donc $Q = X^2 + 1$ et $R = -X^2 - 4X - 3$.

Comme dans le cas de $\mathbb{Z}$, si le reste de la division euclidienne de A par B est nul, nous dirons que B *divise* A ou que B est un *diviseur* de A.

Muni de la division euclidienne, nous pouvons commencer l'étude arithmétique de $\mathbb{K}[X]$.

Théorème B.2. — *Dans* $\mathbb{K}[X]$*, tout idéal est principal. On dit que* $\mathbb{K}[X]$ *est un* anneau principal.

En d'autres termes, si $\mathfrak{a}$ est un idéal de $\mathbb{K}[X]$, il existe un polynôme P, unique à multiplication par un facteur scalaire non nul près, tel que $\mathfrak{a} = (P) = P\mathbb{K}[X]$. On a donc $\mathfrak{a} = \{QP, Q \in \mathbb{K}[X]\}$, l'ensemble des multiples de P.

Démonstration. — Si $\mathfrak{a} = \{0\}$, on prend $P = 0$. Sinon, soit $E = \{\deg(Q), Q \in \mathfrak{a} - \{0\}\}$. C'est un sous-ensemble non-vide de $\mathbb{N}$ et il admet donc un plus petit élément $d \geq 0$. Soit alors $P \in \mathfrak{a}$ tel que $\deg(P) = d$. On a clairement $(P) \subset \mathfrak{a}$. Soit maintenant A un élément de $\mathfrak{a}$. Nous pouvons effectuer la division euclidienne de A par P, qui donne $A = QP + R$, avec $\deg(R) < \deg(P)$. Comme $\mathfrak{a}$ est un idéal, on a $R \in \mathfrak{a}$, et par hypothèse sur P et minimalité de d, on a nécessairement $\deg(R) = -\infty$, *i.e.* $R = 0$. On a donc $A = QP$, et donc $A \in (P)$. L'unicité à un multiple scalaire non nul près vient de la définition plus haut et du fait que $U(\mathbb{K}[X]) = U(\mathbb{K}) = \mathbb{K}^*$. Cela achève la démonstration. ∎

Dans la suite, nous choisirons toujours, comme générateur d'un idéal $\mathfrak{a}$, *l'unique* polynôme unitaire P tel que $\mathfrak{a} = (P)$. Si en effet on trouve un polynôme P tel que $\mathfrak{a} = (P)$, on peut le multiplier par l'inverse de son coefficient dominant, et on obtient un polynôme unitaire.

Soient maintenant deux polynômes P et Q, non tous les deux nuls. Alors l'idéal $\mathfrak{a} = (P, Q)$ est principal, et il existe donc un unique polynôme unitaire S tel que $\mathfrak{a} = (S)$. Il est alors clair que P et Q sont multiples de S, et donc que S divise P et Q.

Définition B.3. — *Soient* P *et* Q *deux polynômes non tous les deux nuls. On appelle* plus grand commun diviseur *de* P *et* Q *l'unique polynôme unitaire* S *tel que* $(P, Q) = (S)$*. On le note* $\mathrm{pgcd}(P, Q)$*. On dit que deux polynômes* P *et* Q *sont* premiers entre eux *si* $\mathrm{pgcd}(P, Q) = 1$*, donc si leurs uniques diviseurs communs sont les polynômes constants non nuls.*

Théorème de Bézout. — *Pour que les polynômes* P *et* Q *soient premiers entre eux, il faut et il suffit qu'il existe des polynômes* U *et* V *tels que*

$$UP + VQ = 1.$$

Démonstration. — Supposons P et Q premiers entre eux. Alors $(P, Q) = (1)$, mais l'idéal engendré par 1 est trivialement $\mathbb{K}[X]$ en entier. Par définition de (P, Q), il existe alors une combinaison linéaire $UP + VQ$ de P et Q telle que $UP + VQ = 1$. Réciproquement, s'il existe de tels U et V, alors par définition de (P, Q), on a $1 \in (P, Q)$, et donc $(P, Q) = \mathbb{K}[X] = (1)$, et P et Q sont premiers entre eux. ∎

Théorème de Gauss. — *Soient* A, B, C *trois polynômes tels que* C *soit premier avec* B *et divise* AB*. Alors* C *divise* A*.*

Démonstration. — Le polynôme C divise AB, donc il existe L tel que $AB = CL$. Il est en revanche premier avec B donc, par le théorème de Bézout, il existe U, V tels que $UB + VC = 1$. On en déduit que

$$UCL + VCA = A, \text{ i.e. } C(UL + AV) = A,$$

ce qui prouve bien que C divise A. ∎