Chapitre 3

Théorie des invariants de similitude

Dans tout ce chapitre E est un $\mathbb{F}$-espace vectoriel de dimension finie $d \geq 1$ et $u \in \mathcal{L}(E)$. On note χ_u le polynôme caractéristique de u et μ_u le polynôme minimal de u.

3.1 Endomorphismes cycliques

Définition 3.1.1. *Pour $x \in E$ on pose*

$$E_{u,x} = \mathrm{Vect}(u^k(x))_{k\in\mathbb{N}}.$$

C'est le plus petit sous-espace de E stable par u qui contient x. On l'appelle le sous-espace u-cyclique de E engendré par x.

Proposition 3.1.1. *Si on note $d_{u,x}$ la dimension de $E_{u,x}$, alors $(x, u(x), \dots, u^{d_{u,x}-1}(x))$ est une base de $E_{u,x}$.*

Démonstration. Il suffit de montrer que la famille $(x, u(x), \dots, u^{d_{u,x}-1}(x))$ est libre (ça sera une base car elle possède $d_{u,x}$ vecteurs). Si elle était liée il existerait $l \leq d_{u,x} - 1$ tel que $u^l(x) \in \mathrm{Vect}(x, \dots, u^{l-1}(x))$. En appliquant u on en déduirait

$$u^{l+1}(x) \in \mathrm{Vect}(u(x), \dots, u^l(x)) \subset \mathrm{Vect}(x, u(x), \dots, u^l(x)) \subset \mathrm{Vect}(x, u(x), \dots, u^{l-1}(x))$$

car $u^l(x) \in \mathrm{Vect}(x, \dots, u^{l-1}(x))$. Par récurrence on en déduirait

$$u^{l+k}(x) \in \mathrm{Vect}(x, \dots, u^{l-1}(x))$$

pour tout $k \geq 0$, et ainsi $E_{u,x} = \mathrm{Vect}(x, \dots, u^{l-1}(x))$ ce qui est absurde car E est de dimension $d_{u,x} > l$. Donc $(x, u(x), \dots, u^{d_{u,x}-1}(x))$ est libre. $\square$

Définition 3.1.2. *On dit que u est cyclique s'il existe un vecteur $x \in E$ (nécessairement non nul) tel que $E = E_{u,x}$. On dit alors que x est u-cyclique, ou encore que x est un vecteur cyclique pour u.*

Ces endomorphismes jouent un rôle important dans la théorie des invariants de similitude.

Définition 3.1.3. *Soit $P = X^d + a_{d-1}X^{d-1} + \cdots + a_1X + a_0 \in \mathbb{F}[X]$ un polynôme unitaire de degré d, on appelle la matrice compagnon associée à P la matrice*

$$C(P) := \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{d-1} \end{pmatrix}.$$

On a une première caractérisation des endomorphismes cycliques :

Proposition 3.1.2. *Les propriétés suivantes sont équivalentes :*

1. *u est cyclique.*
2. *Il existe une base B de E telle que $\mathrm{Mat}_B(u)$ est une matrice compagnon.*

Démonstration. 1. ⇒ 2. : D'après la proposition 3.1.1 si x est u-cyclique, alors $B = (x, u(x), \dots, u^{d-1}(x))$ est une base de E, et en décomposant $u^d(x)$ dans B sous la forme $u^d(x) = -\sum_{k=0}^{d-1} a_k u^k(x)$, on a $\mathrm{Mat}_B(u) = C(P)$ pour $P = \sum_{k=0}^{d-1} a_k X^k + X^d$.

2. ⇒ 1. : Si $\mathrm{Mat}_B(u)$ est compagnon, en notant x le premier vecteur de la base B, on a $B = (x, u(x), \dots, u^{d-1}(x))$ et donc u est bien cyclique. □

Les matrices compagnon vérifient la propriété suivante.

Proposition 3.1.3. *Soit $A = C(P) \in \mathcal{M}_d(\mathbb{F})$, alors $\chi_A = \mu_A = P$. En particulier $C(P)$ est semblable à $C(Q)$ si et seulement si $P = Q$, si et seulement si $C(P) = C(Q)$.*

Démonstration. Calculons

$$\chi_A(X) = \det \begin{pmatrix} X & 0 & \dots & 0 & a_0 \\ -1 & X & \dots & 0 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & X & a_{d-2} \\ 0 & 0 & \dots & -1 & X + a_{d-1} \end{pmatrix}.$$

On fait sur $XI_d - A$ les opérations $L_1 := L_1 + XL_2 + X^2L_3 + \cdots + X^{d-1}L_d$ et on se retrouve avec

$$A'(X) = \begin{pmatrix} 0 & 0 & \dots & 0 & P(X) \\ -1 & X & \dots & 0 & a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & X & a_{d-2} \\ 0 & 0 & \dots & -1 & X + a_{d-1} \end{pmatrix}$$

dont le déterminant se calcule en développant selon la première ligne :

$$\chi_A(X) = \det(A'(X)) = (-1)^{1+d} P(X) \det \begin{pmatrix} -1 & X & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & X \\ 0 & 0 & \dots & -1 \end{pmatrix} = (-1)^{1+d} P(X)(-1)^{d-1} = P(X).$$

Ainsi $\chi_A = P$. Ensuite notons $(e_1, \dots, e_d)$ la base canonique de $\mathbb{F}^d$, on a : $A^k(e_1) = e_k$ pour $k = 0, \dots, d-1$, donc si $Q = \sum_{i=0}^{d-1} c_i X^i \in \mathbb{F}_{d-1}[X] - \{0\}$ on aura $Q(A)(e_1) = \sum_{i=0}^{d-1} c_i e_{i+1} \neq 0$. Donc μ_A est de degré au moins d. On constate ensuite que $P(A)(e_1) = 0$. On en déduit que $A^k P(A)(e_1) = P(A)A^k(e_1) = P(A)(e_k) = 0$ pour tout $k = 1, \dots, d$ et donc $P(A) = 0$, donc $\mu_A = P$.

En particulier si $C(P)$ et $C(Q)$ sont semblables, elles ont même polynôme caractéristique donc $P = Q$. □

Des propositions 3.1.2 et 3.1.3 on déduit le résultat suivant.

Corollaire 3.1.1. *Si u est cyclique, alors $\pi_u = \mu_u$.*

On rappelle que si $F < E$ est un sous-espace u-stable, alors

$$\chi_{u_{|F}} | \chi_u$$

et

$$\mu_{u_{|F}} | \mu_u.$$

Pour $x \in E$ on pose

$$u_x := u_{|E_{u,x}} \in \mathcal{L}(E_{u,x}).$$

Comme u_x est cyclique on a

$$\mu_{u_x} = \chi_{u_x} | \chi_u.$$

Avec cette étude préliminaire des endomrophismes cycliques il est maintenant facile de démontrer le résultat suivant, appelé théorème de Cayley-Hamilton.

Théorème 3.1.1. *On a $\mu_u | \chi_u$, i.e. $\chi_u(u) = 0$.*

Démonstration. En effet $\chi_u(u)(x) = 0$ pour tout x dans E puisque $\chi_u(u)$ est nul sur $E_{u,x}$. On en déduit que $\chi_u(u) = 0$. □

Exercice 3.1.1. *On suppose u diagonalisable. Montrer que u est cyclique si et seulement si u est à spectre simple (sepctre simple : d valeurs propres distinctes).*

3.2 Sous-espaces cycliques de dimension maximale

La théorie des invariants de similitude permet de caractériser les classes de similitudes dans $\mathcal{M}_d(\mathbb{F})$, autrement dit les orbites de l'action de $\mathrm{GL}(E)$ sur $\mathcal{L}(E)$ par conjugaison. Dans cette théorie telle qu'on la présente ici, les sous-espaces cycliques de E de dimension maximale jouent un rôle crucial. On rappelle qu'un sous-espace cylique de E est de la forme $E_{u,x}$, et que sa dimension $d_{u,x}$ qui est aussi le degré du polynôme caractéristique χ_{u_x} est donc aussi le degré du polynôme minimal μ_{u_x} puisque $\chi_{u_x} = \mu_{u_x}$. Comme $\mu_{u_x} | \mu_u$ car μ_u annule u_x, on en déduit que $d_{u,x} \leq d^\circ(\mu_u)$. On va maintenant voir qu'il existe au moins un sous-espaces cycliques de E de dimension maximale $d^\circ(\mu_u)$.

Proposition 3.2.1. *Il existe x dans E tel que $\mu_{u_x} = \mu_u$.*

Avant d'en débuter la démonstration, on constate que si $Q \in \mathbb{F}[X]$ et $x \in E$, alors

$$Q(u)_{|E_{u,x}} = 0 \Leftrightarrow Q(u)(x) = 0.$$

Le sens direct est immédiat puisque $x \in E_{u,x}$, et inversement si $Q(u)(x) = 0$ alors $Q(u)u^k(x) = 0$ pour tout $k \geq 0$ car $Q(u)$ et u^k commutent. On en déduit que $Q(u)$ est nul sur $E_{u,x}$. On constate aussi que si F est un sous-espace u-stable de E qui contient x, alors $E_{u,x} \subset F$.

Démonstration du lemme 3.2.1. Il se démontre en trois étapes. On décompose μ_u en irréductibles :

$$\mu = P_1^{m_1} \dots P_r^{m_r}$$

avec $P_i \in \mathbb{F}[X]$ irréducitble, et les $m_i \geq 1$. On pose $E_i = \mathrm{Ker}(P_i^{m_i}(u))$, c'est un sous-espace u-stable de E. D'après le lemme des noyaux on a $E = \oplus_{i=1}^r E_i$.

Première étape : on montre qu'il existe $x_i \in E_i$ tel que $\mu_{u,x_i} = P_i^{m_i}$.
On remarque que $P_i^{m_i}(u)$ est nul sur E_i, donc il annule tous les x dans E_i. Comme E_i est u-stable, si $x \in E$ alors $E_{u,x} \subset E_i$ et donc $P_i^{m_i}(u)$ est nul sur $E_{u,x}$ dès que $x \in E_i$. Ainsi μ_{u_x} divise

$P_i^{m_i}$ pour tout x dans E_i. Si cette division était stricte pour tout x, comme P_i est irréductible, on aurait μ_{u_x} divise $P_i^{m_i-1}$ pour tout x dans E_i, i.e. $P_i^{m_i-1}(u)$ serait nul sur $E_{u,x}$ por tout x dans E_i, et donc $P_i^{m_i-1}(u)$ serait nul sur E_i ($x \in E_{u,x}$). Mais alors

$$(P_1^{m_1} \dots P_{i-1}^{m_{i-1}} P_i^{m_i-1} P_{i+1}^{m_{i+1}} \dots P_r^{m_r})(u)$$

serait nul sur $E = \oplus_{i=1}^r E_i$, ce qui contredirait la définition de μ_u. Donc il existe x_i tel que $\mu_{u,x_i} = P_i^{m_i}$.

Deuxième étape : Si x et y dans E vérifient $\mu_{u_x} \wedge \mu_{u_y} = 1$, alors $\mu_{u_{x+y}} = \mu_{u_x}\mu_{u_y}$.
Montrons que μ_{u_x} divise $\mu_{u_{x+y}}$, ce qui revient à montrer que $\mu_{u_{x+y}}(u)$ annule x. On remarque d'abord que

$$\mu_{u_{x+y}}(u)(x+y) = 0 \Longrightarrow \mu_{u_{x+y}}(u)(x) = -\mu_{u_{x+y}}(u)(y).$$

D'après le lemme de Bézout il existe deux polynômes A et B tels que $A\mu_{u_x} + B\mu_{u_y} = 1$, et donc $\mathrm{Id}_E = A(u)\mu_{u_x}(u) + B(u)\mu_{u_y}(u)$, ce qui implique que

$$\mathrm{Ker}(\mu_{u_x}(u)) \cap \mathrm{Ker}(\mu_{u_y}(u)) = \{0\}.$$

De plus comme $x \in \mathrm{Ker}(\mu_{u_x}(u))$ (par définition de μ_{u_x}) qui est u-stable donc $Q(u)$-stable pour tout polynôme Q, on a $\mu_{u_{x+y}}(u)(x) \in \mathrm{Ker}(\mu_{u_x}(u))$. Pour les mêmes raisons $-\mu_{u_{x+y}}(u)(y) \in \mathrm{Ker}(\mu_{u_y}(u))$, et donc $\mu_{u_{x+y}}(u)(x) = 0$ car il est dans l'intersection $\mathrm{Ker}(\mu_{u_x}(u)) \cap \mathrm{Ker}(\mu_{u_y}(u))$. Nous avons démontré que μ_{u_x} divise $\mu_{u_{x+y}}$ et par symétrie μ_{u_y} divise $\mu_{u_{x+y}}$. Le lemme de Gauss implique alors que

$$\mu_{u_x}\mu_{u_y} | \mu_{u_{x+y}}.$$

Réciproquement

$$(\mu_{u_x}\mu_{u_y})(u)(x+y) = \mu_{u_x}(u)\mu_{u_y}(u)(x) + \mu_{u_x}(u)\mu_{u_y}(u)(y)$$

$$= \mu_{u_y}(u)\mu_{u_x}(u)(x) + \mu_{u_x}(u)\mu_{u_y}(u)(y) = 0 + 0 = 0$$

donc

$$\mu_{u_{x+y}} | \mu_{u_x}\mu_{u_y},$$

ce qui termine la démonstration de la deuxième étape.

Troisième étape : il existe $x \in E$ tel que $\mu_{u_x} = \mu_u$.
Posons $x = x_1 + \cdots + x_r$. D'après l'étape 1 on a $\mu_{u,x_i} = P_i^{m_i}$. D'après une récurrence immédiate et l'étape 2 on a $\mu_{u,x_1+\cdots+x_k} = \prod_{i=1}^k P_i^{m_i}$ pour $k = 1, \dots, r$ car $P_k^{m_k}$ est premier à $\prod_{i=1}^{k-1} P_i^{m_i}$. On en déduit que

$$\mu_{u,x_1+\cdots+x_r} = \mu_u.$$

□

On a le corollaire immédiat suivant.

Corollaire 3.2.1. *L'endomorphisme u est cyclique si et seulement si $\mu_u = \chi_u$.*

Démonstration. Le sens direct a été vu dans le corollaire 3.1.1. Réciproquement si $\mu_u = \chi_u$, comme il existe $x \in E$ tel que $\chi_{u_x} = \mu_{u_x} = \mu_u$ d'après la proposition 3.2.1, on en déduit que $\chi_{u_x} = \chi_u$ est de degré d. Or χ_{u_x} a pour degré la dimension $d_{u,x}$ de $E_{u,x}$, donc $E_{u,x}$ et E ont la même dimension, ils sont donc égaux. □

Exercice 3.2.1. *Donner une démonstration plus directe de l'exercice 3.1.1.*

Le deuxième résultat crucial sur les sous-espaces cycliques de E de dimension maximale, et qu'ils admettent automatiquement un supplémentaire stable.

Proposition 3.2.2. *Soit $E_{u,x}$ un sous-espace cyclique de E de dimension maximale, i.e. $d_{u,x} = d^\circ(\mu_u)$, alors $E_{u,x}$ admet un supplémentaire stable par u.*

Démonstration. On pose $d_{\mu_u} = d_{u,x} = d^\circ(\mu_u)$. On rappelle que $(x, \dots, u^{d_{\mu_u}-1}(x))$ est une base de $E_{u,x}$. Il existe alors une unique forme linéaire ϕ_0 dans $E^*_{u,x}$ telle que $\phi_0(x) = \cdots = \phi_0(u^{d_{\mu_u}-2}(x)) = 0$ et $\phi_0(u^{d_{\mu_u}-1}(x)) = 1$. On étend ϕ_0 en une forme linéaire $\phi \in E^*$ (en lui imposant par exemple d'être nulle sur un supplémentaire de $E_{u,x}$), de telle sorte que ϕ vérifie $\phi(x) = \cdots = \phi(u^{d_{\mu_u}-2}(x)) = 0$ et $\phi(u^{d_{\mu_u}-1}(x)) = 1$. Alors la famille $(\phi, u^*(\phi), \dots, u^{*d_{\mu_u}-1}(\phi))$ est libre dans E^*. En effet si

$$0 = \sum_{k=0}^{d_{\mu_u}-1} \lambda_k u^{*k}(\phi) = \sum_{k=0}^{d_{\mu_u}-1} \lambda_k \phi \circ u^k,$$

alors en évaluant en x on obtient $\lambda_{d_{\mu_u}-1} = 0$, puis en évaluant en $u(x)$ on obtient $\lambda_{d_{\mu_u}-2} = 0$ et ainsi de suite. On remarque que $\mathrm{Vect}(\phi, u^*(\phi), \dots, u^{*d_{\mu_u}-1}(\phi))$ est u^*-stable car $u^{*d_{\mu_u}} \in \mathrm{Vect}(\mathrm{Id}_{E^*}, u^*, \dots, u^{*d_{\mu_u}-1})$ puisque $u^{d_{\mu_u}} \in \mathrm{Vect}(\mathrm{Id}_E, u, \dots, u^{d_{\mu_u}-1})$. En fait

$$\mathrm{Vect}(\phi, u^*(\phi), \dots, u^{*d_{\mu_u}-1}(\phi)) = E^*_{u^*,\phi},$$

et il est de dimension d_μ. Son orthogonal $F = (E^*_{u^*,\phi})^\perp$ est donc u-stable et de dimension $d - d_{\mu_u}$. On aura terminé si on montre que $F \cap E_{u,x} = \{0\}$. Soit $y \in F \cap E_{u,x}$, alors $y = \sum_{k=0}^{d_{\mu_u}-1} a_k u^k(x)$ pour $a_k \in \mathbb{F}$, et il est annulé par $\phi \circ u^i$ pour tout i. On a donc $0 = \phi(y) = a_{d_{\mu_u}-1}$, puis on en déduit que $0 = \phi(u(y)) = a_{d_{\mu_u}-2}$ et ainsi de suite, donc tous les a_k sont nuls et y aussi. $\square$

3.3 Invariants de similitude

On commence par établir un résultat qui nous sera utile dans la démonstration de l'unicité des invariants de similitude.

Proposition 3.3.1. *Soient x et y dans E tels que $\mu_{u_x} = \mu_{u_y}$. Alors il existe un unique $g \in \mathrm{Iso}(E_{u,x}, E_{u,y})$ qui envoie x sur y et qui vérifie $g \circ u_x = u_y \circ g$.*

Démonstration. On remarque que $d_{u,x} = d_{u,y}$ (car $\chi_{u,x} = \mu_{u_x} = \mu_{u_y} = \chi_{u,y}$), et on note δ cette valeur commune. Si g existe elle est unique car elle envoie $u^k(x)$ sur $u^k(y)$ pour $k = 0, \dots, \delta - 1$, et automatiquement bijective car elle envoie base sur base. Soit donc g l'unique élément de $\mathrm{Iso}(E_{u,x}, E_{u,y})$ qui envoie $u^k(x)$ sur $u^k(y)$ pour $k = 0, \dots, \delta - 1$. Clairement $g \circ u_x$ et $u_y \circ g$ coïncident en $x, u(x), \dots, u^{\delta-2}(x)$, mais aussi en $u^{\delta-1}(x)$ car $\mu_{u_x} = \mu_{u_y}$. Donc $g \circ u_x = u_y \circ g$. $\square$

Corollaire 3.3.1. *Soient x et y dans E tels que $\mu_{u_x} = \mu_{u_y}$, alors pour tout $R \in \mathbb{F}[X]$ on a $\dim(R(u)(E_{u,x})) = \dim(R(u)(E_{u,y}))$.*

Démonstration. On remarque que

$$g(R(u)(E_{u,x})) = g(R(u_x)(E_{u,x})) = g \circ R(u_x)(E_{u,x}) = R(u_y) \circ g(E_{u,x})$$

$$= R(u_y)(g(E_{u,x})) = R(u_y)(E_{u,y}) = R(u)(E_{u,y}),$$

et le résultat en découle. $\square$

On note $\sim$ la relation de similitude sur $\mathcal{M}_d(\mathbb{F})$. Si $d = d_1 + \cdots + d_r$ avec $d_i > 0$ et $M_i \in \mathcal{M}_d(\mathbb{F})$ on pose

$$\operatorname{diag}(M_1, \ldots, M_r) = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_r \end{pmatrix} \in \mathcal{M}_d(\mathbb{F}).$$

Théorème 3.3.1. *Soit $M \in \mathcal{M}_d(\mathbb{F})$, alors il existe une unique suite finie de polynômes unitaires non constants $P_1, \ldots, P_r$ vérifiant $P_{i+1}|P_i$ pour $i = 1, \ldots, r-1$, telle que*

$$M \sim \operatorname{diag}(C(P_1), \ldots, C(P_r)).$$

C'est une traduction matricielle du théorème suivant qui est celui qu'on va démontrer.

Théorème 3.3.2. *Soit $u \in \mathcal{L}(E)$, alors il existe une unique suite finie de polynômes unitaires non constants $P_1, \ldots, P_r$ tels que $P_{i+1}|P_i$ pour $i = 1, \ldots, r-1$, telle qu'il existe une décomposition $E = \oplus_{i=1}^r E_i$ avec les $E_i = E_{u,x_i}$ cycliques tels que $\mu_{u_{x_i}} = P_i$.*

Démonstration. Pour l'existence, on choisit grâce à la proposition 3.2.1 un vecteur $x_1 \in E$ tel que $\mu_{u_{x_1}} = \mu_u$, et on pose $E_1 = E_{u,x_1}$ et $P_1 := \mu_u$. D'après la proposition 3.2.2 E_1 possède un supplémentaire u-stable F. Par récurrence F se décompose comme voulu, et on conclut en constatant que $\mu_{u_x}|\mu_u = P_1$ pour n'importe quel $x \in E$.
Supposons maintenant qu'il existe une autre telle suite $Q_1, \ldots, Q_s$ qui convient, et soient $F_1, \ldots, F_s$ les sous-espaces cycliques associés (avec $F_i = \mu_{u_{y_i}}$). On a $Q_1(u)(E) = \oplus_{i=1}^s Q_1(u)(F_i)$ (la somme est directe car les F_i sont u-stables) mais $Q_1(u)(F_i) = \{0\}$ car $Q_i|Q_1$ donc $Q_1(u) = \{0\}$, i.e. $\mu_u|Q_1$. Comme $Q_1 = \mu_{u_{y_1}}$ il divise μ_u, et donc $Q_1 = \mu_u = P_1$ et $d_{u,x_1} = d_{u,y_1}$. Ensuite

$$Q_2(u)(E) = \oplus_{i=1}^s Q_2(u)(F_i) = Q_2(u)(F_1),$$

mais aussi

$$Q_2(u)(E) = \oplus_{i=1}^r Q_2(u)(E_i).$$

Or $Q_2(u)(F_1)$ et $Q_2(u)(E_1)$ ont la même dimension d'après le corollaire 3.3.1. On en déduit que $Q_2(u)(E_i) = \{0\}$ pour $i \geq 2$, en particulier pour $i = 2$, et donc $P_2|Q_2$. Par symétrie $Q_2|P_2$ et on répète le même argument (un utilisant à nouveau le corollaire 3.3.1) pour aboutir $s = r$ et $P_i = Q_i$ pour $i = 1, \ldots, r$. □

On remarque que $\mu_u = P_1$ et $\chi_u = P_1 \ldots P_r$ dans le théorème ci-dessus.

3.4 Réduction de Jordan des endomorphismes trigonalisables

Pour $\lambda \in \mathbb{F}$ et $k \geq 1$ on pose

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathcal{M}_k(\mathbb{F}).$$

On remarque que

$$N_k = J_k(0)$$

est nilpotente d'ordre k et on en déduit le résultat suivant.

Lemme 3.4.1. *La matrice $J_k(\lambda)$ est cyclique avec $\chi_{J_k(\lambda)} = \mu_{J_k(\lambda)} = (X - \lambda)^k$ et elle est semblable à sa transposée.*

Démonstration. Comme $\mu_{J_k(\lambda)} | \chi_{J_k(\lambda)}$ on a $\mu_{J_k(\lambda)} = (X - \lambda)^a$ avec $a \leq k$, mais si on avait $a < k$ on en déduirait que N_k est nilpotente d'ordre a, ce qui n'est pas le cas. De plus on a ${}^tJ_k(\lambda) = w_k J_k(\lambda) w_k^{-1}$ avec

$$w_k = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} \in \mathrm{GL}_k(\mathbb{F}).$$

□

On a alors la première étape vers notre théorème principal.

Proposition 3.4.1. *Soit P un polynôme scindé et notons $\lambda_1, \dots, \lambda_t$ ses racines distinctes, ce qui revient à dire que C_P est trigonalisable de valeurs propres $\lambda_1, \dots, \lambda_t$. Notons m_i la multiplicité de λ_i comme racine de P. Alors*

$$C_P \sim \mathrm{diag}(J_{m_1}(\lambda_1), \dots, J_{m_t}(\lambda_t)).$$

Démonstration. Il suffit de montrer que si on pose $A = \mathrm{diag}(J_{m_1}(\lambda_1), \dots, J_{m_t}(\lambda_t))$, alors A a les mêmes invariants de similitude que C_P. or C_P étant cyclique, ses invariants de similitudes sont au nombre de 1 : $\mu_{C_P} = P$ (qui vaut aussi χ_{C_P}). Il suffit donc de montrer que $\mu_A = \chi_A = P$. La relation $\chi_A = P$ est claire et μ_A divise P. Mais $\mu_{J_{m_i}(\lambda_i)}$ doit diviser μ_A pour tout i (raisonner en termes de sous-espaces stables), or $\mu_{J_{m_i}(\lambda_i)} = (X - \lambda_i)^{m_i}$ d'après le lemme 3.4.1, et comme les $(X - \lambda_i)^{m_i}$ sont premiers entre eux, on en déduit que leur produit divise μ_A donc $\mu_A = \chi_A$. □

Soit $m \in \mathbb{N}^*$, on dit que $(m_1, \dots, m_l)$ est une *partition* de m si $m_1 \geq \dots \geq m_l \geq 1$ sont des entiers tels que $\sum_{i=1}^r m_i = m$. Si $\overline{m} := (m_1, \dots, m_l)$ est une partition de m on pose

$$J_{\overline{m}}(\lambda) = \mathrm{diag}(J_{m_1}(\lambda), \dots, J_{m_l}(\lambda)).$$

Le théorème fondamental de ce paragraphe est le suivant.

Théorème 3.4.1. *Soit $M \in \mathcal{M}_d(\mathbb{F})$ trigonalisable de valeurs propres distinctes $\lambda_1, \dots, \lambda_t$ (avec $t \leq d$), alors il existe un unique élément $(d_{\lambda_1}, \dots, d_{\lambda_t}) \in (\mathbb{N}^*)^t$ avec $\sum_{i=1}^t d_{\lambda_i} = d$ et pour chaque d_{λ_i} une unique partition $\overline{d_{\lambda_i}}$, tels que*

$$M \sim \mathrm{diag}(J_{\overline{d_{\lambda_1}}}(\lambda_1), \dots, J_{\overline{d_{\lambda_t}}}(\lambda_t)).$$

(on remarque que d_{λ_i} est clairement la multiplicité de λ_i dans χ_M).

Démonstration. Pour l'existence, notons $P_1 | \dots | P_r$ la suite des invariants de similitude de M. Alors on a $P_j = \prod_{i=1}^t (X - \lambda_i)^{d_{i,j}}$ avec $d_{i,1} \geq \dots \geq d_{i,r} \geq 0$ et $\sum_{j=1}^r d_{i,j} = d_{\lambda_i}$ pour $i = 1, \dots, t$. Pour chaque $i = 1, \dots, t$ on pose l_i le plus grand k entre 1 et r tel que $d_{i,k} \geq 1$, de sorte que $\overline{d_{\lambda_i}} := (d_{i,1,d} ots, d_{i,l_i})$ est une partition de d_{λ_i}. D'après le théorème 3.3.1 et la proposition 3.4.1, en utilisant que $\mathrm{diag}(A, B) \sim \mathrm{diag}(B, A)$ autant de fois que c'est nécessaire on en déduit

$$M \sim \mathrm{diag}(J_{\overline{d_{\lambda_1}}}(\lambda_1), \dots, J_{\overline{d_{\lambda_t}}}(\lambda_t)).$$

Pour l'unicité si

$$M \sim \mathrm{diag}(J_{\overline{d_{\lambda_1}}'}(\lambda_1), \dots, J_{\overline{d_{\lambda_t}}'}(\lambda_t)),$$

avec $\overline{d_{\lambda_i}}'$ une partition quelconque de d_{λ_i}. On écrit $\overline{d_{\lambda_i}}' = (d'_{i,1}, \ldots, d'_{i,l'_i})$ et on pose $r' = \max_{i=1,\ldots,t} l'_i$. On pose ensuite pour chaque i tel que $l'_i < r$, $d'_{i,l'_i+1} = \cdots = d'_{i,r'} = 0$, et $P'_i = \prod_{j=1}^t (X-\lambda_i)^{d'_{i,l'_i+1}}$. Alors clairement $P'_i | P'_{i+1}$ pour $i = 1, \ldots, r'-1$ et de plus en utilisant à nouveau la proposition 3.4.1 et $\mathrm{diag}(A,B) \sim \mathrm{diag}(B,A)$ on obtient que

$$M \sim \mathrm{diag}(P'_1, \ldots, P'_{r'}).$$

Par unicité des invariants de similitude on en déduit $r' = r$ et $P'_i = P_i$, puis $\overline{d'_{\lambda_i}} = \overline{d_{\lambda_i}}$, ce qui entraîne l'unicité de chaque $\overline{d_{\lambda_i}}$. □

On en déduit immédiatement une partie du théorème suivant.

Théorème 3.4.2. *Soit $u \in \mathcal{L}(E)$ trigonalisable, alors il existe un unique couple $(d,n) \in \mathcal{L}(E)^2$ avec d diagonalisable, n nilpotent et $d \circ n = n \circ d$ tel que $u = d + n$. De plus d et n sont des polynômes en u.*

Démonstration. L'existence d'un tel couple (d,n) découle immédiatement du théorème 3.4.1. De plus ce même théorème décompose E en une somme directe de sous-espaces u-stables

$$E = E_1 \oplus \cdots \oplus E_t$$

où E_i admet une base B_i telle que $\mathrm{Mat}_{B_i}(u_{|E_i}) = J_{\overline{d_{\lambda_i}}}(\lambda_i)$. Posons $u_i := u_{|E_i}$, alors $(u_i - \lambda_i \mathrm{Id}_{E_i})^{d_{\lambda_i}} = 0$, c'est à dire que $E_i \subset \ker((u - \lambda_i \mathrm{Id}_{E_i})^{d_{\lambda_i}})$. Comme d'après le théorème de Cailey-Hamilton combiné au lemme des noyaux on a aussi

$$E = \ker((u - \lambda_1 \mathrm{Id}_{E_1})^{d_{\lambda_1}}) \oplus \cdots \oplus \ker((u - \lambda_t \mathrm{Id}_{E_t})^{d_{\lambda_t}}),$$

on en déduit que les $E_i = \ker((u - \lambda_i \mathrm{Id}_{E_i})^{d_{\lambda_i}})$ pour des raisons de dimension. Autrement dit les E_i sont les sous-espaces caractéristiques de u. On note p_i la projection de E sur E_i selon $\oplus_{j \neq i} E_j$, alors on verra en exercice 3.4.3 que $p_i \in \mathbb{F}[u]$, or $d = \sum_{i=1}^t \lambda_i p_i \in \mathbb{F}[u]$, donc $n = u - d \in \mathbb{F}[u]$ aussi. Si (d', n') est un autre couple qui convient, alors $d' - d = n - n'$ et le côté gauche est diagonalisable car d et d' commutent pusique d' commute à n donc à u donc à $\mathbb{F}[u]$, et le côté droit est nilpotent car n' commute à n. On en déduit que chaque côté est nul. □

Exercice 3.4.1. *Soit $u \in \mathcal{L}(E)$. Montrer que μ_u et χ_u ont les mêmes facteurs irréductibles (unitaires) dans $\mathbb{F}[X]$.*

Solution. Notons $P_r | P_{r-1} | \ldots | P_2 | P_1$ les invariants de similitude de u, alors $\mu_u = P_1$ et $\chi_u = P_1 \ldots P_r$. Comme chaque P_i divise P_1, on en déduit que $P_1 = \mu_u$ et χ_u ont les mêmes facteurs irréductibles.

Exercice 3.4.2. *On décompose μ_u et χ_u en produit d'irréductibles uniraires : $\mu_u = \prod_{i=1}^t R_i^{a_i}$ et $\chi_u = \prod_{i=1}^t R_i^{b_i}$ avec $1 \leq a_i \leq b_i$ d'après l'exercice 3.4.1 et le théorème de Cayley-Hamilton. Alors $\ker(R_i(u)^{a_i}) = \ker(R_i(u)^{b_i})$ pour tout $i = 1, \ldots, t$.*

Solution. si $x \in \ker(R_i(u)^{a_i})$, alors $R_i(u)^{b_i}(x) = R_i(u)^{b_i - a_i} \circ R_i(u)^{a_i}(x) = 0$, donc

$$\ker(R_i(u)^{a_i}) \subset \ker((R_i(u)^{b_i})).$$

En particulier si on pose $d_{1,i} = \dim(\ker(R_i(u)^{a_i}))$ et $d_{2,i} = \dim(\ker(R_i(u)^{b_i}))$ on a $d_{1,i} \leq d_{2,i}$. Mais d'après le lemme des noyaux on a aussi $\ker(\mu_u(u)) = E = \oplus_{i=1}^t \ker(R_i(u)^{a_i})$ et $\ker(\chi_u(u)) = E = \oplus_{i=1}^t \ker(R_i(u)^{a_i})$. En particulier $\sum_{i=1}^t d_{1,i} = d = \sum_{i=1}^t d_{2,i}$, et comme $d_{1,i} \leq d_{2,i}$ on en déduit $d_{1,i} = d_{2,i}$ et donc $\ker(R_i(u)^{a_i}) = \ker((R_i(u)^{b_i}))$.

Exercice 3.4.3. *Avec les notations de l'exercice 3.4.2, on note p_i la projection de E sur $C_i := \ker(R_i(u)^{b_i})$ par rapport à $\oplus_{j \neq i} C_j$. Alors $p_i \in \mathbb{F}[u]$.*

Solution. On pose $Q_i = \prod_{j \neq i} R_j^{b_j}$ det sorte que $\mathrm{pgcd}(Q_1, \ldots, Q_t) = 1$. Ainsi d'après le théorème de bézout il existe $U_1, \ldots, U_r$ dans $\mathbb{F}[X]$ tels que $\sum_{k=1}^{t} U_k Q_k = 1$ et donc

$$\sum_{k=1}^{t} U_k(u) \circ Q_k(u) = \mathrm{Id}_E. \tag{3.1}$$

Il est clair que $Q_i(u)$ annule C_j dès que $j \neq i$ car $Q_i = A_{i,j} R_j^{b_i}$ pour $A_{i,j} \in \mathbb{F}[X]$ par définition et donc $Q_i(u) = A_{i,j}(u) \circ R_j(u)^{b_i}$. On en déduit que $U_i(u) \circ Q_i(u)$ annule aussi C_j pour $j \neq i$. De plus d'après (3.1) pour $x \in C_i$, on a $x = \sum_{k=1}^{t} U_k(u) \circ Q_k(u)(x) = U_i(u) \circ Q_i(u)(x)$, et donc $p_i = U_i(u) \circ Q_i(u) \in \mathbb{F}[u]$.