Algèbre Linéaire Avancée 2 pour physiciens

Marc Troyanov

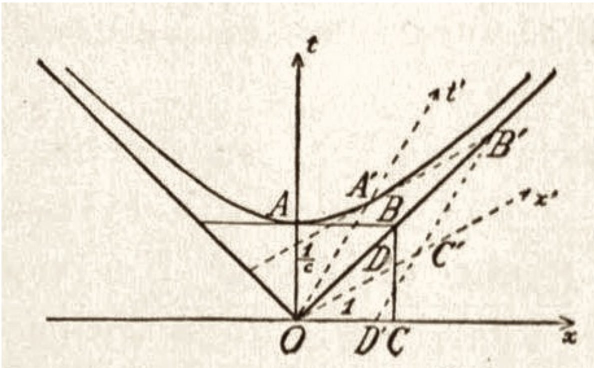

Fig. 1.

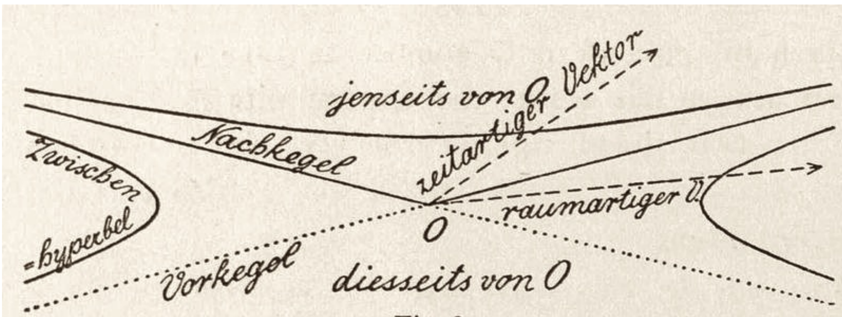

Fig. 2.

Ce document contient les notes du cours d'algèbre linéaire avancée 2 enseigné à l'EPFL, printemps 2024 (document en cours de révision).

Table des matières

9 Structure des endomorphismes **3**

9.1 Triangulation des matrices et des endomorphismes 4
9.2 Polynômes d'endomorphismes et de matrices 7
9.3 Polynômes annulateurs et polynôme minimal d'un endomorphisme 8
9.4 Le théorème de Cayley-Hamilton 10
9.5 Une autre preuve du théorème de Cayley-Hamilton 13
9.6 Vecteurs propres généralisés et théorème de réduction primaire 15
9.7 Complément sur les multiplicités 19
9.8 Lemme des noyaux et preuve du théorème de réduction primaire 20
9.9 Décomposition de Dunford 22
9.10 Sous-espaces cycliques d'un endomorphisme 24
9.11 La forme normale de Jordan d'un endomorphisme 26
9.12 Conséquences du théorème de réduction de Jordan 29
9.13 Réduction pratique d'une matrice à sa forme normale de Jordan 31
9.14 Sur les endomorphismes d'espaces vectoriels réels. 36

Chapitre 9

Structure des endomorphismes

Position du problème

Etant donné un endomorphisme $f \in \mathcal{L}(V)$ d'un espace vectoriel V de dimension finie sur un corps K, il est naturel de chercher à en analyser la structure. Le mot analyser vient du grec $\grave{\alpha}\nu\grave{\alpha}\lambda\upsilon\omega$ (*analuô*), qui signifie *délier, décomposer.* Pour analyser la structure d'un endomorphisme on cherche à le réduire en une somme directe d'endomorphismes les plus simples possibles.

De façon plus précise, nous allons chercher à décomposer l'espace V en somme directe de sous-espaces vectoriels qui sont invariants par f :

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_q, \qquad f(W_i) \subset W_i \quad \text{pour tout } i.$$

de façon telle que $f_i = f|_{W_i} \in \mathcal{L}(W_i)$ soit un endomorphisme aussi simple que possible (noter que l'invariance de W_i est nécessaire pour que l'endomorphisme f_i soit bien défini).

Supposons une telle décomposition donnée, on peut alors choisir une base $\mathcal{B}_i$ de chaque sous-espace W_i et la réunion $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_q$ de ces bases forme une base de V (car V est somme directe des W_i). Dans cette base la matrice de f prend la forme d'une matrice diagonale par blocs

$$M_{\mathcal{B}}(f) = A = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_q} \end{pmatrix},$$

où chaque A_i est la matrice de f_i dans la base $\mathcal{B}_i$.

Remarque. Nous nous permettons de noter parfois une telle matrice sous l'une des deux formes suivantes :

$$A = \mathrm{Diag}(A_1, A_2, \cdots, A_q) \quad \text{ou} \quad A = A_1 \oplus A_2 \oplus \cdots \oplus A_q.$$

La première notation nous rappelle que A est une matrice "diagonale par bloc" et la seconde notation nous rappelle que A est la matrice d'un endomorphisme qui laisse invariante une décomposition de V en somme directe dans une base adaptée.

Par exemple

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \oplus (5) \oplus (6) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \oplus \begin{pmatrix} 5 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}$$

Observons que deux endomorphismes conjugués ont la même structure dans le sens suivant : Supposons que $f, f' \in \mathcal{L}(V)$ sont conjugués par un automorphisme g, i.e. $f' = g \circ f \circ g^{-1}$ et que V admet une décomposition comme somme directe de sous-espaces $W_1, \dots, W_q$ invariants par f, alors les sous-espaces $W_i' = g(W_i)$ sont invariants par f' et l'espace V est aussi somme directe des W_i'. De plus si $\mathcal{B}_i$ est une base de W_i, alors $g(\mathcal{B}_i)$ est une base de W_i' et f et f' ont même matrice dans les bases respectives $\mathcal{B} = \mathcal{B}_1 \cup \cdots \mathcal{B}_q$ et $\mathcal{B}' = \mathcal{B}_1' \cup \cdots \mathcal{B}_q'$:

$$M_{\mathcal{B}'}(f') = M_{\mathcal{B}}(f) = A = A_1 \oplus A_2 \oplus \cdots \oplus A_q, \quad \text{avec} \quad A_i = M_{\mathcal{B}_i'}(f_i') = M_{\mathcal{B}_i}(f_i).$$

Ces considérations s'étendent aux matrices carrées. Analyser la structure d'une matrice $B \in M_n(K)$ revient à analyser l'endomorphisme correspondant $L_B : K^n \to K^n$, défini par $L_B(X) = BX$, et à décomposer K^n en somme directe de sous-espaces vectoriels invariants par L_B, puis choisir une base $\mathcal{B}$ de K^n adaptée à cette décomposition de K^n et finalement à faire le changement de base pour obtenir une matrice diagonale par blocs $A = PBP^{-1}$.

Noter que dans ce cas, la matrice de changement de bases P est la matrice dont la $j^{\text{ème}}$ colonne est donnée par les composantes dans la base canonique du $j^{\text{ème}}$ vecteur de la base $\mathcal{B}$.

De même que deux endomorphismes conjugués ont la même structure, deux matrices semblables ont aussi la même structure, et donc les même écritures comme matrices diagonales par blocs (après changement de base).

Le cas le plus simple est celui d'un endomorphisme (ou d'une matrice) diagonalisable. La structure d'un tel endomorphisme est simplement donnée par la décomposition de l'espace V en somme directe de sous-espaces invariants de dimension 1 (la base obtenue étant une base formée de vecteurs propres).

Dans ce qui suit nous étudions la structure des endomorphismes dont le polynôme caractéristique est scindé (c'est toujours le cas si $K = \mathbb{C}$). Les invariants déjà connus pour analyser un tel endomorphisme (ou une telle matrice) sont : son *polynôme caractéristique*, son *spectre* ainsi que les *multiplicités algébrique* et *géométrique* de chaque valeur propre. Nous verrons dans ce chapitre d'autres invariants tels que le *polynôme minimal* et le multiplicités généralisées.

9.1 Triangulation des matrices et des endomorphismes

Rappelons qu'une matrice carrée $A = (a_{ij})$ est dite *triangulaire supérieure* si $a_{ij} = 0$ pour $i > j$, une matrice triangulaire est donc de la forme

$$A = \begin{pmatrix} a_{11} & * & * & \cdots & * \\ 0 & a_{22} & * & & \vdots \\ \vdots & 0 & a_{33} & & \vdots \\ \vdots & \vdots & & \ddots & * \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \tag{9.1}$$

La matrice $A = (a_{ij})$ est *triangulaire inférieure* si $a_{ij} = 0$ pour $i < j$. La transposée d'une matrice triangulaire inférieure est une matrice triangulaire supérieure et, dans la suite, on dira simplement qu'une matrice est *triangulaire* lorsqu'elle est triangulaire supérieure.

Définitions. La matrice $A \in M_n(K)$ est dite *triangulable* [1] si elle est semblable à une matrice triangulaire, i.e. s'il existe $P \in \mathrm{GL}_n(K)$ tel que $P^{-1}AP$ est triangulaire.

On dit qu'un endomorphisme $f \in \mathcal{L}(V)$ d'un K-espace vectoriel de dimension finie est *triangulable* si sa matrice dans une base adéquate est triangulaire.

La proposition suivante est une reformulation de cette définition :

Proposition 9.1.1. *Soit V un K-espace vectoriel de dimension finie. L'endomorphisme $f \in \mathcal{L}(V)$ est triangulable si et seulement s'il existe une base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V et des scalaires $a_{ij} \in K$ tels que*

$$f(v_j) = \sum_{i \leqslant j} a_{ij} v_i = \sum_{i=1}^{j} a_{ij} v_i. \tag{9.2}$$

Noter que la condition (9.2) peut aussi s'écrire

$$f(v_j) \in \mathrm{Vec}\left(\{v_1, v_2 \ldots, v_j\}\right).$$

Théorème 9.1.2. *Soit V un K-espace vectoriel de dimension finie. Un endomorphisme $f \in \mathcal{L}(V)$ est triangulable si et seulement si son polynôme caractéristique $\chi_f(t)$ est scindé.*

Rappelons qu'un polynôme est dit *scindé* s'il est produit de polynômes de degré 1.

Preuve. Si f est triangulable, alors il existe une base dans laquelle la matrice de f prend la forme (9.1). Le polynôme caractéristique de f est donc

$$(t - a_{11})(t - a_{22}) \ldots (t - a_{nn}),$$

en particulier ce polynôme est scindé.

On démontre la réciproque par récurrence sur la dimension n du K-espace vectoriel V. Si $n = 1$, il n'y a rien à démontrer car toute matrice de taille 1×1 est triangulaire. Soit $n = \dim(V) > 1$ et supposons le théorème démontré pour tout espace vectoriel de dimension $n - 1$.

Par hypothèse, le polynôme caractéristique $\chi_f(t)$ est scindé. En particulier il existe au moins une racine $\lambda_1 \in K$ de $\chi_f(t)$. Soit $v_1 \in V$ un vecteur propre associé à cette valeur propre, i.e. $v_1 \neq 0$ et $f(v_1) = \lambda_1 v_1$. Choisissons maintenant $v_2, \ldots, v_n \in V$ tels que $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ est une base de V. Notons $W_1 = \mathrm{Vec}(v_1) = Kv_1$ et $W_2 = \mathrm{Vec}(v_2, \ldots, v_n)$. Alors $W_2 \subset V$ est un sous-espace vectoriel de dimension $n-1$ tel que $V = W_1 \oplus W_2$. Notons encore $\pi_1 : V \to W_1$ et $\pi_2 : V \to W_2$ les projections canoniques, alors $f = f_1 + f_2$, avec $f_i = \pi_i \circ f$. La matrice de f dans la base $\mathcal{B}$ prend la forme

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & S & \\ 0 & & & \end{pmatrix}$$

1. On dit parfois que la matrice est *trigonalisable*.

où S est la matrice de l'endomorphisme $g_2 = f_2|_{W_2} \in \mathcal{L}(W_2)$ dans la base $\{v_2, \ldots, v_n\}$. Le polynôme caractéristique de f est égal au polynôme caractéristique de la matrice ci-dessus, donc

$$\chi_f(t) = (t - \lambda_1) \cdot \chi_S(t) = (t - \lambda_1) \cdot \chi_{g_2}(t)$$

Ceci entraîne en particulier que le polynôme $\chi_{g_2}(t)$ est également scindé. Par hypothèse de récurence, g_2 est triangulable et on peut donc trouver une (nouvelle) base $\{v_2', \ldots, v_n'\}$ de W_2 dans laquelle la matrice S' de g_2 est triangulaire.
Observons que pour $j \geq 2$ on a $f(v_j') = f_1(v_j') + f_2(v_j') = f_1(v_j') + g_2(v_j')$. Or $f_1(v_j')$ est un multiple de v_1, donc la matrice de f dans la base $\{v_1, v_2', \ldots, v_n'\}$ prend la forme

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & S' & \\ 0 & & & \end{pmatrix}$$

où $S' \in M_{n-1}(K)$ est triangulaire. Il s'agit donc d'une matrice triangulaire de $M_n(K)$.

□

Corollaire 9.1.3. *Toute matrice $A \in M_n(\mathbb{C})$ est triangulable.*

Preuve. Sur $\mathbb{C}$ tout polynôme est scindé par le théorème fondamental de l'algèbre.

□

Corollaire 9.1.4. *Le coefficient d'ordre $n-1$ du polynôme caractéristique d'une matrice $A \in M_n(\mathbb{C})$ est égal à l'opposé de sa trace.*

Preuve. On sait que deux matrices semblables ont la même trace et le même polynôme caractéristique. Par le corollaire précédent, on peut donc supposer que la matrice A est de la forme (9.1). On a alors

$$\begin{aligned} \chi_A(t) &= \prod_{i=1}^{n}(t - a_{ii}) = t^n - \left(\sum_{i=1}^{n} a_{ii}\right) t^{n-1} + \ldots + (-1)^n \prod_{i=1}^{n} a_{ii} \\ &= t^n - \mathrm{Tr}(A) t^{n-1} + \ldots + (-1)^n \det(A). \end{aligned}$$

□

Remarque. On peut aussi prouver ce corollaire directement en examinant la définition du polynôme caractéristique.

Corollaire 9.1.5. *(A) La trace d'une matrice $A \in M_n(\mathbb{C})$ est la somme de ses valeurs propres comptées selon leur multiplicité algébrique :*

$$\mathrm{Tr}(A) = \sum_{\lambda \in \sigma(A)} \lambda \cdot \mathrm{multalg}_A(\lambda).$$

(B) Le déterminant d'une matrice $A \in M_n(\mathbb{C})$ est le produit de ses valeurs propres comptées selon leur multiplicité algébrique :

$$\det(A) = \prod_{\lambda \in \sigma(A)} \lambda^{\mathrm{multalg}_A(\lambda)}.$$

Preuve. (A) Cette formule est évidente pour une matrice triangulaire T. Le théorème précédent nous dit que tout matrice $A \in M_n(\mathbb{C})$ est triangulable. Il existe donc $P \in GL_n(\mathbf{C})$ telle que $T = P^{-1}AP$ est triangulaire. Cela termine la preuve car A et T ont les mêmes valeurs propres et $\mathrm{Tr}(A) = \mathrm{Tr}(P^{-1}AP)$.

(B) Le raisonnement est le même pour le déterminant.

□

Remarque. Le résultat du corollaire précédent reste valable, avec la même preuve, pour toute matrice carrée A à coefficients dans un corps K quelconque, à condition que son polynôme caractéristique soit scindé.

9.2 Polynômes d'endomorphismes et de matrices

La philosophie pour la suite de ce chapitre est la suivante : pour analyser la structure d'un endomorphisme $f \in \mathcal{L}(V)$, on essaye de décomposer V en somme directe de sous-espaces invariants et on analyse la structure de f sur les sous-espaces invariants.

Polynôme d'un endomorphisme

Une opération qui jouera un rôle fondamental est la suivante : soit $f \in \mathcal{L}(V)$ un endomorphisme de V et $p(t) = a_0 + a_1 t + \cdots + a_k t^k$ un polynôme à coefficients dans le corps K. Alors on note $p(f) \in \mathcal{L}(V)$ l'endomorphisme obtenu en substituant l'endomorphisme f à l'indéterminée t :

$$p(f) = a_0 \cdot \mathrm{Id}_V + a_1 \cdot f + \cdots + a_k \cdot f^k \in \mathcal{L}(V),$$

où par définition f^m signifie $f \circ f \circ \cdots \circ f$ (m fois) et $f^0 = \mathrm{Id}_V$. De même, si $A \in M_n(K)$ alors on définit $p(A) \in M_n(K)$ par

$$p(A) = a_0 \mathbf{I}_n + a_1 A + \cdots + a_k A^k.$$

On vérifie alors facilement les résultats suivants :

Théorème 9.2.1. *(a) Pour un endomorphisme $f \in \mathcal{L}(V)$ donné, l'application $K[t] \to \mathcal{L}(V)$ donnée par $p(t) \mapsto p(f)$ est un homomorphisme de K-algèbres. En particulier si $p, q \in K[t]$, alors*

$$(p \cdot q)(f) = p(f) \circ q(f).$$

(b) Si $W \subset V$ est un sous espace invariant *par f, alors ce sous-espace est aussi invariant par $p(f)$.*

(c) Si $f \in \mathcal{L}(V)$ et $g \in GL(V)$, alors pour tout $p \in K[t]$, on a

$$p(g^{-1}fg) = g^{-1}p(f)g.$$

(d) Si $v \in V$ est un vecteur propre de f et λ est la valeur propre associée, alors v est aussi un vecteur propre de $p(f)$ et la valeur propre associée est $p(\lambda)$

Remarquons en particulier que la propriété (a) implique que pour tous polynômes $p(t), q(t) \in K[t]$, l'endomorphisme $p(f)$ commute avec $q(f)$:

$$p(f) \circ q(f) = q(f) \circ p(f).$$

La preuve de cette proposition consiste simplement à vérifier les définitions. Démontrons par exemple l'assertion (d). Observons d'abord que si $f(v) = \lambda v$, alors pour tout entier k on a $f^k(v) = \lambda^k v$. Soit maintenant $p(t) = \sum_{k=0}^m a_k t^k$ un polynôme quelconque, alors on a

$$p(f)(v) = \sum_{k=0}^{m} a_k f^k(v) = \sum_{k=0}^{m} a_k \lambda^k v = p(\lambda)v.$$

Les propriétés correspondantes sont aussi vraies pour les matrices, en remplaçant la composition par la multiplication matricielle.

Remarque. La réciproque de la propriété (d) est fausse. Voici un contre-exemple : considérons la matrice $A = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$ et le polynôme $p(t) = t^4$. Alors $p(A) = A^4 = \begin{pmatrix} 16 & 0 \\ 0 & 16 \end{pmatrix}$, donc $16 = 2^4 \in \sigma(A)$, mais 2 n'est pas valeur propre de A (la matrice A n'a aucune valeur propre réelle et ses valeurs propres complexes sont $\pm 2i$).

9.3 Polynômes annulateurs et polynôme minimal d'un endomorphisme

Définition 9.3.1. On dit qu'un polynôme $p(t)$ *annule* la matrice $A \in M_n(K)$, ou que c'est une *polynôme annulateur de* A si $p(A) = 0$.

De même, si $f \in \mathcal{L}(V)$ est un endomorphisme d'un K-espace vectoriel, on dit que $p \in K[t]$ *annule* f, ou que c'est un *polynôme annulateur* de f si $p(f) = 0 \in \mathcal{L}(V)$, i.e. $p(f)$ est l'endomorphisme nul. On remarque que $p(t)$ est un polynôme annulateur de f si et seulement si $p(f)(v) = 0$ pour tout $v \in V$; de façon équivalente $\text{Ker}(p(f)) = V$.

Exemples 1. L'endomorphisme f est dit *nilpotent* s'il existe m tel que $f^m = 0$. Dans ce cas le polynôme t^m est un polynôme annulateur de f.

2. La matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ vérifie $A^4 = I_2$. Par conséquent $t^4 - 1$ est un polynôme annulateur de A.

3. L'opérateur de dérivation $D = \frac{d}{dx}$ est un endomorphisme de $\mathbb{R}[x]$ qui n'admet aucun polynôme annulateur non nul.

Proposition 9.3.2. *(a) Si $f \in \mathcal{L}(V)$ et $g \in GL(V)$, alors tout polynôme qui annule f annule aussi $g^{-1} \circ f \circ g$.*

(b) Deux matrices semblables ont les mêmes polynômes annulateurs.

(c) Si A est la matrice de f dans une base quelconque de V, alors $p(t)$ est un polynôme annulateur de f si et seulement si c'est un polynôme annulateur de A.

Preuve. Exercice. □

Lemme 9.3.3. *Tout endomorphisme $f \in \mathcal{L}(V)$ d'un espace vectoriel de dimension finie admet des polynômes annulateurs non nuls.*

Preuve. Puisque $\dim \mathcal{L}(V) < \infty$, il existe un entier k tel que la famille d'endomorphismes

$$\{\mathrm{Id}_V, f, f^2, \dots f^k\} \subset \mathcal{L}(V)$$

est liée. Soit k le plus petit entier avec cette propriété, il existe alors des scalaires $\alpha_0, \alpha_1, \dots, \alpha_{k-1}$, uniquement définis et tels que

$$f^k + \alpha_{k-1} f^{k-1} + \cdots + \alpha_1 f + \alpha_0 \,\mathrm{Id}_V = 0 \in \mathcal{L}(V),$$

ce qui signifie que le polynôme

$$\mu_f(t) = t^k + \sum_{i=0}^{k-1} \alpha_i t^i \tag{9.3}$$

annule f.

□

Définition. Le polynôme construit en (9.3) s'appelle le *polynôme minimal* de f. On définit le polynôme minimal d'une matrice de la même manière.

Le polynôme minimal d'un endomorphisme possède les propriétés importantes suivantes :

Proposition 9.3.4. *Soit V un K-espace vectoriel de dimension finie.*

(a) Le polynôme minimal $\mu_f(t)$ d'un endomorphisme f est l'unique polynôme unitaire de plus petit degré qui annule f.

(b) Ce polynôme divise tout polynôme qui annule f.

(c) Deux endomorphismes conjugués ont le même polynôme minimal, i.e. si $f \in \mathcal{L}(V)$ et $g \in \mathrm{Aut}(V)$ alors $\mu_{g^{-1}\circ f\circ g} = \mu_f(t)$.

Preuve. La propriété (a) vient de la définition de $\mu_f(t)$. Observer que l'unicité vient de l'indépendance linéaire des endomorphismes $\mathrm{Id}_V, f, f^2, \dots f^{k-1} \in \mathcal{L}(V)$.

Pour prouver (b), on considère un autre polynôme $p(t)$ annulant f. Si $p(t)$ est non nul, alors $\deg(p) \geq \deg(\mu)$. En appliquant la division polynomiale, il existe deux polynômes $q(t)$ et $r(t)$ tels que

$$p(t) = q(t)\mu_f(t) + r(t) \qquad \text{et} \qquad \deg r(t) < \deg \mu_f(t).$$

Observons que $r(f) = p(f) - q(f) \circ \mu_f(f) = 0$, donc $r(t)$ est le polynôme nul par minimalité du degré de $\mu_f(t)$. On a donc montré que tout polynôme annulateur de f est un multiple de $\mu_f(t)$.

La propriété (c) est conséquence du fait que les endomorphismes f et $g^{-1} \circ f \circ g$ ont les mêmes polynômes annulateurs, ils ont donc le même polynôme minimal.

□

Les mêmes définitions s'appliquent aux matrices, et on peut énoncer en particulier le résultat suivant :

Proposition 9.3.5. *Soient $A, B \in M_n(K)$. Si B est semblable à A alors $\mu_B(t) = \mu_A(t)$.*

9.4 Le théorème de Cayley-Hamilton

Le théorème de Cayley-Hamilton dit que tout endomorphisme d'un espace vectoriel de dimension finie est annulé par son polynôme caractéristique.

Théorème 9.4.1 (Cayley-Hamilton)**.** *Pour tout endomorphisme $f \in \mathcal{L}(V)$ d'un espace vectoriel de dimension finie, on a*

$$\chi_f(f) = 0.$$

De même, pour toute matrice $A \in M_n(K)$ on a $\chi_A(A) = 0$. Autrement dit le polynôme caractéristique de A est un polynôme annulateur de A.

Exemple. On rappelle que le polynôme caractéristique de la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $\chi_A(t) = t^2 - (a+d)t + (ad - bc)$, en appliquant ce polynôme à la matrice elle-même, on calcule que

$$\begin{aligned}
\chi_A(A) &= A^2 - (a+d)A + (ad-bc)\,\mathrm{I}_2 \\
&= \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix} - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Remarquons qu'on pourrait penser que le théorème de Cayley-Hamilton est trivial et être tenté de le prouver en posant simplement $\chi_A(A) = \det(AI_n - A) = \det(A - A) = \det(0) = 0$. Cet argument n'est pas valide car le théorème de Cayley-Hamilton dit que $\chi_A(A) = 0$ en tant que matrice ou en tant qu'endomorphisme, alors que $\det(AI_n - A) = \det(0) = 0$ est une information de type scalaire.

Pour la preuve du théorème de Cayley-Hamilton, nous aurons besoin du lemme suivant, dont nous laissons la preuve en exercice :

Lemme 9.4.2. *Si f est un endomorphisme d'un espace vectoriel V de dimension finie, et si $W \subset V$ est un sous-espace vectoriel invariant par f, alors le polynôme caractéristique de la restriction $f|_W$ de f à W divise le polynôme caractéristique de f.*

Preuve. Nous laissons la preuve de ce lemme en exercice.

Démonstration du théorème de Cayley-Hamilton. Nous allons prouver que pour tout $v \in V$ on a $\chi_f(f)(v) = 0$. Si $v = 0$ il n'y a rien à montrer. On suppose donc que $v \neq 0$ et on note ce vecteur par $v_1 = v$. On considère ensuite le plus grand entier k tel que les vecteurs

$$v_1,\ v_2 = f(v_1),\ v_3 = f(v_2) = f^2(v_1), \ldots, v_k = f(v_{k-1}) = f^{k-1}(v_1)$$

sont linéairement indépendants.

Notons $\mathcal{B} = \{v_1, \ldots, v_k\}$ et $W = \mathrm{Vec}(\mathcal{B}) \subset V$ le sous-espace vectoriel engendré par ces vecteurs. L'ensemble $\mathcal{B}$ est une base de W puisque ces vecteurs sont supposés linéairement indépendants et qu'ils engendrent W.
Par construction, on sait que $f(v_k)$ est combinaison linéaire des éléments de $\mathcal{B}$, il existe donc $\alpha_1, \ldots, \alpha_k \in K$ tels que
$$f(v_k) = \alpha_1 v_1 + \ldots + \alpha_k v_k,$$
et comme $f(v_j) = v_{j+1}$ pour $j < k$, on conclut que le sous-espace W est invariant par f.
Notons $g = f|_W$ la restriction de f au sous-espace W. Alors g est un endomorphisme de W et sa matrice dans la base $\mathcal{B}$ est donnée par

$$A = \mathrm{M}_{\mathcal{B}}(g) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_1 \\ 1 & 0 & & 0 & \alpha_2 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & \alpha_{k-1} \\ 0 & 0 & \cdots & 1 & \alpha_k \end{pmatrix}.$$

On a vu aux exercices que le polynôme caractéristique de cette matrice est

$$\chi_A(t) = \det(t\mathbf{I}_n - A) = -\alpha_1 - \alpha_2 t - \ldots - \alpha_k t^{k-1} + t^k.$$

Par conséquent $\chi_g(f) = \chi_A(f)$ est l'endomorphisme de V défini par

$$\chi_g(f) = -\alpha_1 \mathrm{Id} - \alpha_2 f - \ldots - \alpha_k f^{k-1} + f^k.$$

Si on applique cet endomorphisme à v_1, on trouve

$$\begin{aligned} \chi_g(f)(v_1) &= -\alpha_1 v_1 - \alpha_2 f(v_1) - \alpha_3 f^2(v_1) - \cdots - \alpha_k f^{k-1}(v_1) + f^k(v_1) \\ &= -\alpha_1 v_1 - \alpha_2 v_2 - \cdots - \alpha_k v_k + f(v_k) \\ &= 0. \end{aligned}$$

Il reste à prouver que $\chi_f(f)(v_1) = 0$. Or g est la restriction de f au sous-espace invariant $W \subset V$. Le lemme précédent implique alors que $\chi_g(t)$ est un facteur de $\chi_f(t)$, i.e. $\chi_f(t) = q(t) \cdot \chi_g(t)$ pour un certain polynôme $q(t) \in K[t]$. Par conséquent

$$\chi_f(f)(v_1) = q(f) \circ \chi_g(f)(v_1) = q(f)(0) = 0.$$

On a ainsi montré que pour tout vecteur $v_1 \in V$ non nul on a $\chi_f(f)(v_1) = 0$. Cela signifie que χ_f est l'endomorphisme nul.

□

Corollaire 9.4.3. *Soit f un endomorphisme d'un espace vectoriel de dimension finie V. Alors le polynôme minimal $\mu_f(t)$ divise le polynôme caractéristique $\chi_f(t)$. De plus ces deux polynômes ont exactement les mêmes racines (qui sont les valeurs propres de f).*

Démonstration. On sait que le polynôme minimal divise tout polynôme annulateur de f. En particulier $\mu_f(t)$ divise $\chi_f(t)$ puisque $\chi_f(f) = 0$.

Cela signifie que $\chi_f(t)$ est un multiple de $\mu_f(t)$, par conséquent toute racine de $\mu_f(t)$ est aussi une racine de $\chi_f(t)$ (car $\mu_f(\lambda) = 0 \Rightarrow \chi_f(\lambda) = 0$).

Pour prouver le sens inverse, on suppose que $\lambda \in K$ est une racine de $\chi_f(t)$, c'est donc une valeur propre de f. On a vu que pour tout polynôme $p(t)$, si λ est une valeur propre de f, alors $p(\lambda)$ est valeur propre de $p(f)$ (théorème 9.2.1). En particulier $\mu_f(\lambda)$ est une valeur propre de $\mu_f(f)$, donc $\mu_f(\lambda) = 0$ car $\mu_f(f) = 0 \in V$.

□

Une conséquence de ce corollaire est que si $\chi_f(t)$ est scindé, alors $\mu_f(t)$ est aussi scindé. Plus précisément, si $\chi_f(t) = \prod_{i=1}^r (t-\lambda_i)^{m_i}$, avec $\lambda_1, \dots, \lambda_r$ distincts, alors le polynôme minimal est du type $\mu_f(t) = \prod_{i=1}^r (t-\lambda_i)^{k_i}$, où les exposants $k_i \in \mathbb{N}$ vérifient $1 \leq k_i \leq m_i$ pour tout i.

Ceci nous conduit à une *méthode effective* pour trouver le polynôme minimal d'un endomorphisme f ou d'une matrice A, dont le polynôme caractéristique est scindé :

(1) On calcule le polynôme caractéristique $\chi_f(t)$ et on le factorise.

(2) Si ce polynôme est scindé, il s'écrit $\chi_f(t) = \prod_{i=1}^n (t-\lambda_i)^{m_i}$, où $\sigma(f) = \{\lambda_1, \dots, \lambda_r\}$ est l'ensemble des valeurs propres. Noter que $1 \leq r \leq n = \dim(V)$.

(3) On considère tous les polynômes de type $p(t) = \prod_{i=1}^r (t-\lambda_i)^{s_i}$ avec $1 \leq s_i \leq m_i$ en commençant par $s_j = 1$, et on vérifie si $p(f) = 0$.

(4) Le polynôme $p(t)$ de l'étape précédente dont le degré est minimal est le polynôme minimal $\mu_f(t)$.

Exemple 9.4.4. On considère la matrice

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

Le polynôme caractéristique de cette matrice est $\chi(t) = (t-3)(t-2)^3$. On constate que ce polynôme est scindé. Le polynôme minimal est donc l'un des polynômes suivants :

$$p_1(t) = (t-3)(t-2), \quad p_2(t) = (t-3)(t-2)^2, \quad \text{ou} \quad p_3(t) = \chi(t) = (t-3)(t-2)^3.$$

Pour décider lequel de ces polynômes est le polynôme minimal, on calcule $p_1(A)$ et $p_2(A)$ (on sait déjà par Cayley-Hamilton que $p_3(A) = 0$). Le calcul nous donne :

$$p_1(A) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad p_2(A) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Par conséquent le polynôme minimal est $\mu_A(t) = p_2(t) = (t-3)(t-2)^2$.

Il sera commode de considérer, en plus des polynômes caractéristique et minimal, un troisième polynôme :

Définition 9.4.5. Le *polynôme spectral*[2] d'un endomorphisme f de l'espace vectoriel V de dimension finie est défini par $\nu_f(t) = 1$ si f n'a aucune valeur propre et

$$\nu_f(t) = \prod_{\lambda \in \sigma(f)} (t - \lambda) = \prod_{i=1}^{r} (t - \lambda_i)$$

si le spectre $\sigma(f) = \{\lambda_1, \ldots, \lambda_r\}$ de f est non vide.

Par le corollaire 9.4.3, on sait que le polynôme spectral de f divise le polynôme minimal et le polynôme minimal divise le polynôme caractéristique, ce qu'on peut noter par

$$\nu_f(t) \mid \mu_f(t) \mid \chi_f(t),$$

de plus ces trois polynômes ont les mêmes racines, qui sont les valeurs propres de F.
Dans l'exemple précédent, on a

$$\nu_A(t) = (t-3)(t-2), \quad \mu_A(t) = (t-3)(t-2)^2 \quad \text{et} \quad \chi_A(t) = (t-3)(t-2)^3.$$

9.5 Une autre preuve du théorème de Cayley-Hamilton

Dans ce paragraphe, on propose une autre preuve du théorème de Cayley-Hamilton. Cette preuve se place dans le cadre du calcul matriciel et utilise la notion de *polynôme matriciel.*

Définition 9.5.1. Un *polynôme matriciel* de taille n sur un corps K, est une expression formelle

$$P(t) = A_0 + A_1 t + \cdots A_k t^k,$$

où $A_j \in M_n(K)$ pour tout $j \in \{0, \ldots, k\}$. Le symbole t s'appelle l'*indéterminée* du polynôme et on note $M_n(K)[t]$ l'ensemble de ces polynômes matriciels.

Étant donné $P(t) \in M_n(K)[t]$, on peut ou bien substituer un scalaire $x \in K$ à l'indéterminée t, ou bien une matrice $X \in M_n(K)$. Dans les deux cas on obtient une matrice $P(x)$ ou $P(X)$. Par exemple si

$$P(t) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} t + \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} t^2,$$

alors

$$P(2) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} + 4 \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 20 \end{pmatrix},$$

et

$$\begin{aligned} P\begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 20 \end{pmatrix}^2 \\ &= \begin{pmatrix} 0 & 0 \\ 10 & 0 \end{pmatrix} \end{aligned}$$

2. Ce polynôme ne semble pas avoir de nom particulier dans la littérature, on peut aussi l'appeler le *polynôme des valeurs propres.*

Si $P(t) = A_0 + A_1t + \cdots A_kt^k$ et $Q(t) = B_0 + B_1t + \cdots B_mt^m$, sont deux polynômes matriciels, leur produit est défini par la formule usuelle :

$$(P \cdot Q)(t) = \sum_{r=0}^{k+m} \left(\sum_{i=0}^{r} A_i B_{r-i} \right) t^r,$$

où on considère que $A_i = 0$ si $i > k$ et $B_j = 0$ si $j > m$. En travaillant avec des polynômes matriciels, il y a lieu de prendre certaines précautions. En particulier, si $P(t), Q(t) \in M_n(K)[t]$ et $X \in M_n(K)$, alors, généralement on a

$$(P \cdot Q)(X) \neq P(X) \cdot Q(X).$$

Par exemple si

$$P(t) = A_0 + A_1t \text{ et } Q(t) = B_0 + B_1t,$$

alors

$$(P \cdot Q)(t) = A_0B_0 + (A_0B_1 + A_1B_0)t + A_1B_1t^2.$$

On voit donc que pour tout $X \in M_n(K)$ on a

$$(P \cdot Q)(X) = A_0B_0 + (A_0B_1 + A_1B_0)X + A_1B_1X^2,$$

et

$$P(X)Q(X) = A_0B_0 + A_0B_1X + A_1XB_0 + A_1XB_1X.$$

Si X ne commute pas avec B_1 ou B_0, alors, généralement on aura $(P \cdot Q)(X) \neq P(X) \cdot Q(X)$.

Lemme 9.5.2. *Si $P(t) = A_0 + A_1t + \cdots A_kt^k$ et $Q(t) = B_0 + B_1t + \cdots B_mt^m$, sont deux polynômes matriciels, et $X \in M_n(K)$ est une matrice qui commute avec chaque B_j, alors, la relation*

$$(P \cdot Q)(X) = P(X) \cdot Q(X)$$

est vérifiée.

Preuve. En utilisant que X commute avec chaque B_j, on a

$$\begin{aligned}
P(X)Q(X) &= \left(\sum_{i=0}^{k} A_i X^i \right) \cdot \left(\sum_{j=0}^{m} B_j X^j \right) = \sum_{i=0}^{k} \sum_{j=0}^{m} A_i X^i B_j X^j \\
&= \sum_{i=0}^{k} \sum_{j=0}^{m} A_i B_j X^{i+j} \\
&= \sum_{r=0}^{k+m} \left(\sum_{i=0}^{r} A_i B_{r-i} \right) X^r \\
&= (P \cdot Q)(X).
\end{aligned}$$

□

Preuve alternative du théorème de Cayley-Hamilton.

La preuve de la formule de Laplace du §7.5 (chapitre 7 du polycopié 1) s'applique non seulement à une matrice à coefficients dans un corps K mais aussi, et sans changement, à une matrice à coefficients dans l'anneau des polynômes $K[t]$ (ou dans un autre anneau commutatif quelconque). Donc pour tout polynôme matriciel $Q(t) \in M_n(K[t])$ on a

$$\det(Q(t)) \cdot \mathrm{I}_n = \mathrm{Cof}(Q(t))^\top \cdot Q(t).$$

On applique ce qui précède au polynôme matriciel $Q(t) = (t\, I_n - A)$, et on note

$$P(t) = \mathrm{Cof}(t\,\mathrm{I}_n - A)^\top = C_0 + C_1 t + \cdots + C_{n-1}t^{n-1},$$

où les C_j sont des matrices de taille $n \times n$. On a donc l'identité

$$\chi_A(t) \cdot \mathrm{I}_n = P(t) \cdot (t\,\mathrm{I}_n - A),$$

et on sait par le lemme précédent que dans une telle égalité on peut substituer à t toute matrice $X \in M_n(K)$ qui commute avec A. On a donc pour une telle matrice

$$\chi_A(X) = P(X) \cdot (X - A).$$

Or il est trivial que A commute avec A et on a donc prouvé que

$$\chi_A(A) = P(A) \cdot (A - A) = 0 \in M_n(K).$$

$\square$

9.6 Vecteurs propres généralisés et théorème de réduction primaire

Les notions suivantes joueront un rôle central dans la suite de ce chapitre :

Définition 9.6.1. Soit $f \in \mathcal{L}(V)$ un endomorphisme d'un $K-$espace vectoriel et $\lambda \in K$.

(i) On dit qu'un vecteur $v \in V$ est un *vecteur propre généralisé* de f associé à λ si $v \neq 0$ et s'il existe un entier $m \in \mathbb{N}$ tel que $v \in \mathrm{Ker}\left((\lambda\,\mathrm{Id}_V - f)^m\right)$, i.e.

$$(\lambda\,\mathrm{Id}_V - f)^m v = 0.$$

(ii) Le plus petit entier m tel que[3] $(f - \lambda)^m v = 0$ s'appelle l'*ordre* du vecteur propre généralisé.

(iii) Pour tout $k \in \mathbb{N}$, l'entier

$$\delta_{f,\lambda}(k) = \dim\left(\mathrm{Ker}(f - \lambda)^k\right)$$

s'appelle la *multiplicité généralisée d'ordre* m de λ pour f (lorsque $k = 0$, on convient que $\delta_{f,\lambda}(k) = 0$).

Si $A \in M_n(K)$, on notera de même $\delta_{A,\lambda}(k) = \dim\left(\mathrm{Ker}(A - \lambda\mathrm{Id}_n)^k\right)$. Lorsque l'endomorphisme f (ou la matrice A) été fixé, on notera simplement $\delta_\lambda(k)$.

3. Dans la suite, on s'autorisera pour simplifier à noter l'endomorphisme $(f - \lambda\mathrm{Id}_V)$ par $(f - \lambda)$.

Exemples.

(i) Tout vecteur propre est un vecteur propre généralisé d'ordre 1.

(ii) Si f est nilpotent, i.e. s'il existe m tel que f^m est nul, alors tout élément non nul de V est un vecteur propre généralisé (associé à la valeur propre $\lambda = 0$).

(iii) Le vecteur $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ n'est pas un vecteur propre de la matrice $\begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$, mais c'est un vecteur propre généralisé associé à la valeur propre α.

(iv) La fonction $\varphi \in C^\infty(\mathbb{R})$ définie par $\varphi(x) = x^{m-1} \cdot \mathrm{e}^{\lambda x}$ est un vecteur propre généralisé de l'opérateur $\frac{d}{dx}$ car

$$\left(\frac{d}{dx} - \lambda\right)^m \left(x^{m-1} \cdot \mathrm{e}^{\lambda x}\right) = 0.$$

Nous laissons la vérification de ces exemples en exercice.

Lemme 9.6.2. *S'il existe un vecteur propre généralisé pour $f \in \mathcal{L}(V)$ associé à $\lambda \in K$, alors λ est une valeur propre de f.*

Ce lemme nous dit que s'il existe une notion de vecteur propre généralisé, il n'y a pas de notion de *valeur propre généralisée*, qui serait différente des valeurs propres usuelles.

Preuve. Si v est un vecteur propre associé à λ il n'y a rien à montrer. Sinon, il existe $m \geq 2$ tel que $(f - \lambda)^m v = 0$. Supposons m minimal avec cette propriété et posons $w = (f - \lambda)^{m-1} v$. Alors $w \neq 0$ par hypothèse et $(f - \lambda)w = (f - \lambda)^m v = 0$. Donc w est vecteur propre et la valeur propre associée est λ.

$\square$

Remarque 9.6.3. On montre facilement que deux endomorphismes conjugués (ou deux matrices semblables) ont les mêmes multiplicités généralisées pour chaque valeur propre. De plus, si A est la matrice de f dans une base quelconque, alors $\delta_{f,\lambda}(m) = \delta_{A,\lambda}(m)$ pour toute valeur propre λ et tout entier m.

Les multiplicités généralisées d'une matrice $A \in M_n(K)$ peuvent se calculer au moyen de la formule du rang :

$$\delta_{A,\lambda}(m) = n - \operatorname{rang}(A - \lambda)^m.$$

Rappelons que le rang d'une matrice est le nombre maximal de colonne (ou de lignes) qui sont linéairement indépendantes. Il peut se calculer avec le méthode de Gauss-Jordan.

Le théorème suivant est d'une importance majeure, il nous dit en particulier que si f est un endomorphisme d'un espace vectoriel de dimension finie qui admet un polynôme annulateur scindé, alors tout vecteur de V se décompose de façon unique comme somme de vecteurs propres généralisés.

Théorème 9.6.4 (Théorème de réduction primaire)**.** *Soit f un endomorphisme du K-espace vectoriel de dimension finie V. Considérons le polynôme*

$$p(t) = \prod_{i=1}^{r}(t-\lambda_i)^{s_i},$$

où $\{\lambda_1, \ldots, \lambda_r\} = \sigma(f)$ est l'ensemble des valeurs propres et $s_i \in \mathbb{N}$.
Alors on a les propriétés suivante :

(i) Le sous-espace $U_i = \mathrm{Ker}(\lambda - f)^{s_i}$ est invariant par f pour tout $i \in \{1, \ldots, r\}$.

(ii) La restriction de $(\lambda - f)$ à U_i est un endomorphisme nilpotent.

(iii) Le noyau de $p(f)$ est somme directe des U_i :

$$\mathrm{Ker}(p(f)) = U_1 \oplus \cdots \oplus U_r, \tag{9.4}$$

et sa dimension est

$$\dim \mathrm{Ker}(p(f)) = \sum_{i=1}^{r} \delta_{f,\lambda_i}(s_i). \tag{9.5}$$

Une première conséquence intéressante de ce théorème est le résultat suivant :

Corollaire 9.6.5. *Soit $f \in \mathcal{L}(V)$ un endomorphisme d'un K-espace vectoriel de dimension n. Notons $\sigma(f) = \{\lambda_1, \ldots, \lambda_r\}$ l'ensemble de ses valeurs propres et $\nu_f(t) = \prod_{\lambda\in\sigma(f)}(t-\lambda) \in K[t]$ son polynôme spectral. Alors le noyau de $\nu_f(f)$ est le sous-espace vectoriel de V engendré par les vecteurs propres de f. Plus précisément, on a*

$$\mathrm{Ker}(\nu_f(f)) = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_r}(f). \tag{9.6}$$

où $E_{\lambda_i}(f)$ est l'espace propre associé à la valeur propre λ_i.

Preuve. C'est une application directe du théorème 9.6.4. □

Corollaire 9.6.6. *L'endomorphisme f est diagonalisable si et seulement si le polynôme spectral $\nu_f(t)$ annule f.*

Preuve. C'est une conséquence immédiate de la proposition précédente, puisqu'un endomorphisme f d'un espace vectoriel V est diagonalisable si et seulement si V est somme directe des espaces propres de f. □

Remarque. Le corollaire précédent implique que pour un endomorphisme f d'un espace vectoriel V de dimension finie, les conditions suivantes sont équivalentes :

(i) f est diagonalisable.

(ii) $\nu_f(t)$ est un polynôme annulateur de f.

(iii) Le polynôme minimal coïncide avec la polynôme spectral : $\mu_f(t) = \nu_f(t)$.

(iv) Le polynôme minimal $\mu_f(t)$ est scindé et toutes ses racines sont simples.

(v) Il existe un polynôme scindé à racines simples qui annule f.

En particulier, *si $\mu_f(t)$ admet une racine multiple, alors f n'est pas diagonalisable.*

Nous laissons la preuve de cette remarque en exercice.

Exemples.
1. Le polynôme caractéristique de la matrice

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

est $\chi_A(t) = t^2 + 1$. Ce polynôme n'a pas de racines réelles, donc A n'a aucun vecteur propre dans $\mathbb{R}^2$ et n'est donc pas diagonalisable dans $M_2(\mathbb{R})$. Par contre si on regarde A comme matrice complexe, alors $\chi_A(t) = (t+i)(t-i)$ et A est donc diagonalisable dans $M_2(\mathbb{C})$. .
2. Le polynôme caractéristique de la matrice

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

est $\chi_B(t) = (t-1)^2$. Donc le polynôme minimal est ou bien $(t-1)$ ou bien $(t-1)^2$. Or on vérifie immédiatement que $(t-1)$ n'annule pas la matrice B, par conséquent $\mu_B(t) = (t-1)^2$ possède une racine double et B n'est pas diagonalisable dans $M_2(\mathbb{C})$.
3. Le polynôme minimal de la matrice

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

est $(t-3)(t-2)^3$. (cf. exemple 9.4.4). Cette matrice n'est donc pas diagonalisable.

Pour énoncer la seconde conséquence importante du théorème de décomposition primaire, on introduit la notion de sous-espace caractéristique par rapport à un endomorphisme.

Définition 9.6.7. Soit f un endomorphisme d'un espace vectoriel de dimension finie V et $\lambda \in \sigma(f)$ une valeur propre de f de multiplicité algébrique $m_\lambda = \mathrm{multalg}_\lambda(f)$ (rappelons qu'il s'agit du plus grand entier m tel que $(t-\lambda)^m$ divise le polynôme caractéristique $\chi_f(t)$). Le sous-espace vectoriel

$$N_\lambda(f) = \mathrm{Ker}(f-\lambda)^{m_\lambda} \subset V$$

s'appelle le *sous-espace caractéristique*, ou *sous-espace propre généralisé* associé à la valeur propre λ.

Nous avons alors le corollaire suivant du théorème de réduction primaire

Corollaire 9.6.8. *Soit f un endomorphisme d'un espace vectoriel de dimension finie V dont le polynôme caractéristique est scindé, alors*

(i) On a la décomposition suivante de V en somme directe :

$$V = N_{\lambda_1}(f) \oplus \cdots \oplus N_{\lambda_r}(f), \tag{9.7}$$

cela signifie que tout vecteur $v \in V$ peut s'écrire de façon unique comme somme de vecteurs propres généralisés.

(ii) Si $\lambda \in \sigma(f)$ une valeur propre de f, alors l'ensemble des vecteurs propres généralisés associés à λ est l'ensemble des vecteurs non nuls de $N_\lambda(f)$.

Preuve. Dans le cas où le polynôme caractéristique de f est scindé, le théorème de Cayley-Hamilton nous dit alors que $\mathrm{Ker}(\chi_f(f)) = V$. On peut donc appliquer le théorème de réduction primaire 9.6.4 au polynôme $p(t) = \chi_f(t)$ et on conclut que V est somme directe des sous-espaces propres généralisés, ce qui prouve la première affirmation. La seconde affirmation est une conséquence immédiate de la première.

$\square$

La seconde affirmation de ce corollaire peut se reformuler en disant que si $m > m_\lambda = \mathrm{multalg}_f(\lambda)$, alors $\mathrm{Ker}(f - \lambda\,\mathrm{Id}_V)^m = \mathrm{Ker}(f - \lambda)^{m_\lambda}$, de façon équivalente, l'*ordre de tout vecteur propre généralisé est au plus égal à la multiplicité algébrique de la valeur propre associée.*

9.7 Complément sur les multiplicités

Nous démontrons ici deux résultats complémentaires : le premier concerne les multiplicités algébriques d'un endomorphisme :

Proposition 9.7.1. *Soit f un endomorphisme d'un espace vectoriel de dimension finie. Supposons que le polynôme caractéristique de f est scindé, alors*

(a) La somme des multiplicités algébriques de toutes les valeurs propres est égale à la dimension de l'espace vectoriel V :

$$\sum_{\lambda\in\sigma(f)} \mathrm{multalg}_\lambda(f) = \dim(V)$$

(b) La multiplicité algébrique de toute valeur propre $\lambda \in \sigma(f)$ est égale à la dimension de l'espace caractéristique associé :

$$\mathrm{multalg}_\lambda(f) = \dim(N_\lambda(f)).$$

(c) Pour tout $\lambda \in \sigma(f)$ on a

$$1 \leq \mathrm{multgeom}_\lambda(f) \leq \mathrm{multalg}_\lambda(f) \leq n.$$

Rappelons que le sous-espace caractéristique de f associé à la valeur propre λ est le noyau $N_\lambda(f) = \mathrm{Ker}(f - \lambda)^{m_\lambda}$, où $m_\lambda = \mathrm{multalg}_\lambda(f)$.

Preuve. a.) Le polynôme caractéristique de f est scindé, il s'écrit donc $\chi_f(t) = \prod_{i=1}^r (t - \lambda_i)^{m_i}$, où on a noté $m_i = \mathrm{multalg}_{\lambda_i}(f)$ et on a

$$\sum_{i=1}^r m_i = \deg(\chi_f(t)) = \dim(V).$$

b.) Notons $N_i = N_{\lambda_i}(f)$ le $i^{\text{ème}}$ sous-espace caractéristique de f et $m_i = \mathrm{multalg}_{\lambda_i}(f)$. Le sous-espace N_i est invariant par f, donc la restriction de f à N_i définit un endomorphisme $f_i \in \mathcal{L}(N_i)$. Le théorème de réduction primaire 9.6.4 avec le théorème de Cayley-Hamilton nous dit que les sous-espaces N_i sont invariants et $V = N_1 \oplus \cdots \oplus N_r$. Le polynôme caractéristique admet donc la factorisation suivante :

$$\chi_f(t) = \chi_{f_1}(t) \cdots \chi_{f_r}(t).$$

Le lemme 9.6.2 entraîne que $f_i \in \mathcal{L}(N_i)$ n'a qu'une valeur propre qui est λ_i, car tout vecteur non nul de N_i est un vecteur propre généralisé associé à λ_i donc le polynôme $\chi_{f_i}(t)$ est du type $(t-\lambda_i)^{k_i}$ pour tout $i = 1, \ldots, r$. On a donc

$$\prod_{i=1}^{r}(t-\lambda_i)^{k_i} = \chi_f(t) = \prod_{i=1}^{r}(t-\lambda_i)^{m_i},$$

et par unicité de la décomposition d'un polynôme en facteurs irréductibles, on en déduit que

$$m_i = k_i = \deg(\chi_{f_i}(t)) = \dim(N_i).$$

pour tout i.

c.) Rappelons que par définition $\mathrm{multgeom}_\lambda(f) = \dim(E_\lambda(f))$. Pour toute valeur propre λ de f, on a

$$\{0\} \neq E_\lambda(f) = \mathrm{Ker}(f-\lambda) \subset \mathrm{Ker}(f-\lambda)^{m_\lambda} \subset V,$$

d'où les inégalités voulues. □

Le second résultat concerne les multiplicités généralisées :

Proposition 9.7.2. *Si le polynôme caractéristique de $f \in \mathcal{L}(V)$ est scindé et $\lambda \in \sigma(f)$, alors*

(i) Les multiplicités généralisées associées à chaque valeur propre forment une suite monotone : $\delta_{f,\lambda}(k) \leq \delta_{f,\lambda}(k+1)$ pour tout $k \in \mathbb{N}$.

(ii) $\delta_{f,\lambda}(1) = \mathrm{multgeom}_\lambda(f)$.

(iii) $\delta_{f,\lambda}(k) = \mathrm{multalg}_\lambda(f)$, pour tout $k \geq \mathrm{multalg}_\lambda(f)$.

Preuve. Les deux premières propriétés sont immédiates à partir de la définition $\delta_{f,\lambda}(k) = \dim \mathrm{Ker}(f-\lambda)^{m_k}$. La troisième propriété se déduit du point (ii) du Corollaire 9.6.8.

□

Cette proposition implique en particulier qu'il existe $m \leq \mathrm{multalg}_\lambda(f)$ tel que

$$\mathrm{multgeom}_\lambda(f) = \delta_\lambda(1) \leq \delta_\lambda(2) \leq \cdots \leq \delta_\lambda(m) = \delta_\lambda(m+1) = \mathrm{multalg}_\lambda(f).$$

Le plus petit m ayant cette propriété est l'ordre maximal d'un vecteur propre associé à λ.

9.8 Lemme des noyaux et preuve du théorème de réduction primaire

Dans cette section, nous démontrons le théorème de décomposition primaire. La preuve repose sur le résultat suivant, qui en est une généralisation, et qui s'appelle le *lemme des noyaux.*

Théorème 9.8.1 (Lemme des noyaux)**.** *Soit V un K-espace vectoriel, $f \in \mathcal{L}(V)$ un endomorphisme de V et $p(t) \in K[t]$ un polynômes Supposons que $p(t)$ se factorise sous la forme d'un produit*

$$p(t) = q_1(t)q_2(t)\cdots q_r(t)$$

où les polynômes $q_i(t)$ sont deux-à-deux premiers entre eux. Notons $W_i = \mathrm{Ker}(q_i(f))$, pour $i = 1, \ldots, r$. Alors les W_i sont invariants par f et le noyau de $p(f)$ se décompose comme somme directe

$$\mathrm{Ker}(p(f)) = W_1 \oplus \cdots \oplus W_r.$$

Rappelons que les polynômes $q_i(t), q_j(t), \cdots, q_r(t)$ sont *premiers entre eux* s'ils n'admettent pas de facteur commun non constant.

Démonstration. La preuve se fait par récurrence sur r. Pour $r = 1$ il n'y a rien à démontrer, démontrons le théorème pour $r = 2$, i.e. $p(t) = q_1(t)q_2(t)$. Comme q_1 et q_2 sont supposés premiers entre eux, l'identité de Bézout (cf. Appendice A du polycopié du premier semestre) nous dit qu'il existe $s_1(t), s_2(t) \in K[t]$ tels que

$$q_1(t)s_1(t) + q_2(t)s_2(t) = 1.$$

On a aussi en substituant f à l'indéterminée la relation

$$q_1(f)s_1(f) + q_2(f)s_2(f) = \mathrm{Id}_V\,. \tag{9.8}$$

- On montre d'abord que la somme est directe. Soit donc $v \in \mathrm{Ker}(q_1(f)) \cap \mathrm{Ker}(q_2(f))$. On veut montrer que $v = 0$. En utilisant la relation (9.8), on peut écrire

  $$v = (\mathrm{Id}_V)(v) = \big(q_1(f)s_1(f) + q_2(f)s_2(f)\big)(v).$$

  Or ce vecteur est nul car $q_1(f)s_1(f)(v) = s_1(f)q_1(f)(v) = s_1(f)(0) = 0$ puisque on a supposé $v \in \mathrm{Ker}(q_1(f))$. De même $q_2(f)s_2(f)(v) = 0$. Remarquons qu'on a utilisé (et qu'on réutilisera dans la suite) que deux polynômes en f commutent.

- On prouve maintenant que $\mathrm{Ker}(p(f)) = \mathrm{Ker}(q_1(f)) + \mathrm{Ker}(q_2(f))$. Il s'agit de montrer deux inclusions.
  - Supposons d'abord $v \in \mathrm{Ker}(q_1(f)) + \mathrm{Ker}(q_2(f))$. Cela signifie que v s'écrit $v = v_1 + v_2$ où $v_i \in \mathrm{Ker}\,(q_i(f))$ (pour $i = 1, 2$). Montrons que dans ce cas on $v \in \mathrm{Ker}(p(f))$:

    $$\begin{aligned} p(f)(v) &= q_1(f)q_2(f)(v) \\ &= q_1(f)q_2(f)(v_1 + v_2) \\ &= q_2(f)q_1(f)(v_1) + q_1(f)q_2(f)(v_2) \\ &= 0. \end{aligned}$$

  - Supposons maintenant que $v \in \mathrm{Ker}(p(f))$, i.e. $p(f)(v) = 0$. A l'aide de l'expression (9.8), on peut écrire v sous la forme

    $$v = \big(q_1(f)s_1(f) + q_2(f)s_2(f)\big)(v).$$

    Notons alors $v_1 = q_2(f)s_2(f)(v)$ et $v_2 = q_1(f)s_1(f)(v)$ et montrons que cette écriture permet de voir v comme élément de $\mathrm{Ker}(q_1(f)) + \mathrm{Ker}(q_2(f))$. En effet

    $$q_1(f)(v_1) = q_1(f)q_2(f)s_2(f)(v) = s_2(f)p(f)(v) = s_2(f)(0) = 0.$$

    On montre de même que $q_2(f)(v_2) = 0$.

Pour conclure la preuve par récurrence on se ramène au cas $r = 2$ en écrivant

$$p(t) = q_1(t)\varphi(t)$$

avec $\varphi(t) = q_2(t) \cdots q_r(t)$. □

La démonstration du théorème de décomposition primaire est maintenant très courte. Rappelons d'abord l'énoncé :

Théorème. *Soit f un endomorphisme du K-espace vectoriel de dimension finie V et $p(t)$ le polynôme*

$$p(t) = \prod_{i=1}^{r} (t - \lambda_i)^{s_i},$$

où $\{\lambda_1, \ldots, \lambda_r\} = \sigma(f)$ est l'ensemble des valeurs propres et $s_i \in \mathbb{N}$. Alors on a les propriétés suivantes :

(i) Le sous-espace $U_i = \mathrm{Ker}(\lambda_i - f)^{s_i}$ est invariant par f.

(ii) La restriction de $(f - \lambda_i)$ à U_i est un endomorphisme nilpotent.

(iii) Le noyau de $p(f)$ est somme directe des U_i :

$$\mathrm{Ker}(p(f)) = U_1 \oplus \cdots \oplus U_r,$$

et sa dimension est

$$\dim\left(\mathrm{Ker}(p(f))\right) = \sum_{i=1}^{r} \delta_{f,\lambda_i}(s_i).$$

Démonstration. Observons d'abord que $x \in U_i$ si et seulement si $(\lambda_i - f)^{s_i}(x) = 0$, donc

$$(\lambda_i - f)^{s_i}(f(x)) = f\left((\lambda_i - f)^{s_i})(x)\right) = 0,$$

ce qui signifie que $f(x) \in U_i$ et prouve l'affirmation (i).
La preuve de (ii) est évidente puisque la restriction de $(\lambda_i - f)^{s_i}$ à $U_i = \mathrm{Ker}(\lambda_i - f)^{s_i}$ est nulle.
Pour prouver (iii), on remarque que les polynômes $(t - \lambda_i)^{s_i}$ sont premiers entre eux car on suppose que $\lambda_i \neq \lambda_j$ pour $i \neq j$. Un argument par récurrence basé sur le lemme de noyaux entraîne alors immédiatement que

$$\mathrm{Ker}(p(f)) = \mathrm{Ker}(\lambda - f)^{s_1} \oplus \cdots \oplus \mathrm{Ker}(\lambda - f)^{s_r},$$

La dernière équation se déduit maintenant du fait que, par définition, $\dim(U_i) = \delta_{f,\lambda_i}(s_i)$.

□

9.9 Décomposition de Dunford

Dans ce paragraphe, nous allons prouver que tout endomorphisme d'un espace vectoriel complexe de dimension finie est somme d'un endomorphisme diagonalisable et d'un endomorphisme nilpotent. Commençons par un résultat particulièrement simple :

Lemme 9.9.1. *Soit $f \in \mathcal{L}(V)$ un endomorphisme d'un espace vectoriel V de dimension finie dont le polynôme caractéristique est scindé. Supposons que f n'admet qu'une valeur propre λ, alors $(f - \lambda \mathrm{Id}_V)$ est nilpotent.*

Preuve. Les hypothèses entraînent que $\chi_f(t) = (t - \lambda)^n$ où $n = \dim(V)$. Par le théorème de Cayley-Hamilton, on a alors $\chi_f(f) = (f - \lambda \mathrm{Id}_V)^n = 0$, ce qui signifie précisément que $(f - \lambda \mathrm{Id}_V)$ est nilpotent.

□

Théorème 9.9.2. *Toute matrice $A \in M_n(K)$ dont le polynôme caractéristique est scindé peut s'écrire sous la forme $A = D + N$, où D est une matrice diagonalisable et N est une matrice nilpotente qui commute avec D, i.e. $DN = ND$.*

On peut montrer que cette décomposition est unique, on l'appelle la *décomposition de Dunford* de A. Cette décomposition s'appelle aussi la décomposition de *Jordan-Chevalley.*

Preuve. Soit $A \in M_n(K)$ et supposons que le polynôme caractéristique $\chi_A(t)$ est scindé, avec racines $\lambda_1, \ldots, \lambda_r$. Notons $N_i = N_{\lambda_i}(A) = \mathrm{Ker}(A - \lambda_i \, \mathrm{I}_n)^{m_i}$ le sous-espace caractéristique associé à la valeur propre λ_i (où m_i est la multiplicité algébrique de λ_i) Le théorème de réduction primaire implique que N_i est invariant par A et

$$K^n = N_1 \oplus \cdots \oplus N_r.$$

Choisissons une base de K^n dont les m_1 premiers vecteurs forment une base de N_1, puis les m_2 vecteurs suivants forment une base de N_2 etc. Alors la matrice de l'opérateur A prend la forme par blocs suivante dans cette base :

$$B = \begin{pmatrix} B_1 & 0 & \ldots & 0 \\ 0 & B_2 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & B_r \end{pmatrix},$$

où le bloc B_i est une matrice de taille $m_i \times m_i$. Plus précisément on a $B = P^{-1}AP$ où P est la matrice de passage de la base canonique dans la nouvelle base.
En utilisant le théorème de triangulation, on peut au moyen d'un changement de base supplémentaire se ramener au cas où chaque bloc B_i est une matrice triangulaire supérieure dont les coefficients diagonaux sont tous égaux à la valeur propre λ_i.
Chaque sous-matrice B_i peut alors s'écrire $B_i = \lambda_i \mathrm{I}_{m_i} + T_i$ où T_i est une matrice strictement triangulaire (i.e. avec 0 sur la diagonale). La matrice T_i est nilpotente et commute avec $\lambda_i \mathrm{I}_{m_i}$, ce qui complète la preuve du théorème.

□

Remarques. 1. La décomposition de Dunford permet de calculer les puissances de toute matrice carrée à coefficient complexe, car le polynôme caractéristique d'une telle matrice est scindé et on a

$$A^m = (D + N)^m = \sum_{j=0}^{m} \binom{m}{j} D^j N^{m-j}.$$

Les puissances de la matrice diagonale D et de la matrice nilpotente N sont évidemment faciles à calculer. Notons que le développement binomial ci-dessus pour $(D + N)^m$ est valide car les deux matrices commutent.

2. On peut démontrer que la décomposition de Dunford d'une matrice est unique. Dans le paragraphe suivant on est effective (calculable). De plus D et N s'obtiennent comme polynômes de A.

9.10 Sous-espaces cycliques d'un endomorphisme

La structure d'un endomorphisme d'un espace vectoriel de dimension finie dont le polynôme caractéristique est scindé est décrite d'une manière très complète par sa *forme normale de Jordan*[4]. Nous abordons ce thème par la proposition suivante, qui jouera un rôle fondamental dans la suite.

Proposition 9.10.1. *Soit $f \in \mathcal{L}(V)$ un endomorphisme d'un espace vectoriel sur le corps K, et $u \in V$ un vecteur propre généralisé d'ordre m de f pour la valeur propre λ. Alors les vecteurs $u_1, \dots, u_m \in V$ définis par $u_j = (f-\lambda)^{m-j}(u)$, i.e.*

$$u_1 = (f-\lambda)^{m-1}(u),\ u_2 = (f-\lambda)^{m-2}(u),\ \dots, u_{m-1} = (f-\lambda)(u),\ u_m = u, \tag{9.9}$$

sont linéairement indépendants. De plus, le sous-espace vectoriel $U \subset V$ engendré par $\{u_1, \dots, u_m\}$ est invariant par f.

Définition 9.10.2. Un sous-espace vectoriel U de l'espace vectoriel V est dit *cyclique* pour l'endomorphisme $f \in \mathcal{L}(V)$ s'il contient un vecteur propre généralisé u d'ordre $m = \dim(U)$. Dans ce cas, les vecteurs $\{u_1, \dots, u_m\} \subset U$ définis par (9.9) forment la *base cyclique* de U associée au vecteur $u = u_m$. On dit aussi que $u = u_m$ est une *racine* (ou un *générateur*) du cycle.

Remarque 9.10.3. Les vecteurs définis par (9.9) vérifient $(f-\lambda)(u_1) = (f-\lambda)^m(u) = 0$ et $(f-\lambda)(u_j) = u_{j-1}$ pour $j \geq 2$. En appliquant $(f-\lambda)^k$ à ces vecteurs, on trouve inductivement que

$$(f-\lambda)^k(u_j) = \begin{cases} u_{j-k}, & \text{si } j > k, \\ 0, & \text{si } j \leq k. \end{cases} \tag{9.10}$$

Preuve de la proposition. Rappelons que $u \in V$ est un vecteur propre généralisé d'ordre $m \geq 2$ de f pour la valeur propre λ si $(f-\lambda)^m(u) = 0$ et $(f-\lambda)^{m-1}(u) \neq 0$.

Si $m = 1$, nous avons $u_1 = u_m = u$ qui est non nul car c'est un vecteur propre ; il n'y a donc rien à démontrer dans ce cas et on suppose pour la suite de la preuve que $m \geq 2$.

Observons que par définition $(f-\lambda)(u_1) = 0$ et $(f-\lambda)(u_j) = u_{j-1}$ pour $j \geq 2$, par conséquent

$$f(u_1) = \lambda u_1 \quad \text{et} \quad f(u_j) = u_{j-1} + \lambda u_j, \quad \text{pour } j = 2, \dots, m, \tag{9.11}$$

ce qui entraîne en particulier que le sous-espace $U \subset V$ est invariant par f.

Pour montrer que les vecteurs $\{u_1, \dots, u_m\}$ sont linéairement indépendants, on observe d'abord que ces vecteurs sont non nuls en raison de la condition $(f-\lambda)^{j-1}(u_j) = u_1 \neq 0$.

Supposons maintenant que $\sum_{j=1}^m \alpha_j u_j = 0$ avec $\{\alpha_1, \dots, \alpha_m\} \subset K$, et appliquons $(f-\lambda)^{m-1}$ à cette relation. On trouve à partir de (9.10) que

$$0 = \big(f-\lambda\big)^{m-1}\left(\sum_{j=1}^m \alpha_j u_j\right) = \sum_{j=1}^m \alpha_j \big(f-\lambda\big)^{m-1}(u_j) = \alpha_m u_1,$$

4. On dit aussi forme *canonique* de Jordan, ou forme *réduite* de Jordan ; ces expressions sont synonymes.

par conséquent $\alpha_m = 0$. En appliquant $(f-\lambda)^{m-2}$, on trouve maintenant

$$0 = (f-\lambda)^{m-2}\left(\sum_{j=1}^{m}\alpha_j u_j\right) = \sum_{j=1}^{m}\alpha_j (f-\lambda)^{m-2}(u_j) = \alpha_{m-1}u_1 + \alpha_m u_2 = \alpha_{m-1}u_1,$$

ce qui implique que par conséquent $\alpha_{m-1} = 0$. En répétant l'argument, on trouve que $\alpha_j = 0$ pour tout j.

$\square$

Rappelons la relation (9.11) qui dit que $f(u_1) = \lambda u_1$ et $f(u_j) = u_{j-1} + \lambda u_j$ pour $k = 2, \ldots, m$. La matrice de la restriction de f au sous-espace cyclique $U \subset V$ dans la base cyclique prend la forme

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \tag{9.12}$$

Une telle matrice s'appelle un *bloc de Jordan de taille m*. Par exemple

$$J_1(\lambda) = (\lambda), \quad J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad \text{et} \quad J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Remarquons aussi qu'un bloc de Jordan $J_m(0)$ de valeur propre $\lambda = 0$ est une matrice nilpotente et que, d'une manière générale, tout bloc de Jordan est somme d'une matrice scalaire et d'une matrice nilpotente puisque

$$J_m(\lambda) = \lambda I_m + J_m(0).$$

Lorsque $U = V$ dans la définition précédente, on dit que V est un espace vectoriel cyclique pour l'endomorphisme f. C'est donc le cas si et seulement s'il existe un vecteur propre généralisé dont l'ordre est égale à la dimension de V. Le lemme suivant nous donne une information sur ces endomorphismes qui sera très utile dans la suite.

Lemme 9.10.4. *Soit $f \in \mathcal{L}(U)$ un endomorphisme λ-cyclique d'ordre $m = \dim(U)$, alors pour tout $k \in \{1, \ldots, m\}$ on a*

(a) *$\{u_1, \ldots, u_k\}$ est une base de* $\operatorname{Ker}(f-\lambda)^k$.

(b) *$\{u_1, \ldots, u_{m-k}\}$ est une base de* $\operatorname{Im}(f-\lambda)^k$.

(c) *Les multiplicités généralisées de λ valent*

$$\delta_{f,\lambda}(k) = \min\{k, m\}.$$

Preuve. Les affirmations (a) et (b) découlent immédiatement des équations (9.10). L'affirmation (c) se déduit de (a) et de la définition $\delta_{f,\lambda}(k) = \dim(\operatorname{Ker}(f-\lambda)^k)$.

$\square$

Remarque. On peut aussi prouver ce lemme par le calcul matriciel. La matrice de $(f-\lambda)$ dans la base cyclique est un bloc de Jordan $J_m(0)$ pour $\lambda = 0$; il suffit donc de calculer les puissances de $J_m(0)$ pour voir quel est le rang. Il est facile de voir que si $1 \leq k \leq m$, alors $J_m(0)^k$ est la matrice qui a le coefficient 1 en position $(i, i+k)$ et 0 partout ailleurs. Les k premières colonnes de cette matrice sont nulles et les $m-k$ dernières colonnes sont linéairement indépendantes. Donc le rang de $J_m(0)^k$ est égale à $m-k$.

Voyons quelques conséquences immédiates du lemme précédent : Soit $f \in \mathcal{L}(V)$ un endomorphisme λ-cyclique d'ordre m, alors

(i) $(f-\lambda)$ est nilpotent d'ordre m.

(ii) Tout les vecteurs de V sont des vecteurs propres généralisés de f.

(iii) λ est l'unique valeur propre de f et sa multiplicité géométrique est 1.

(iv) On a $\chi_f(t) = \mu_f(t) = (t-\lambda)^m$, en particulier f n'est pas diagonalisable si $m \geq 2$.

9.11 La forme normale de Jordan d'un endomorphisme

Théorème 9.11.1 (Théorème de réduction de Jordan.)**.** *Soient V un espace vectoriel de dimension finie sur un corps K et $f \in \mathcal{L}(V)$ un endomorphisme de V. Si le polynôme caractéristique $\chi_f(t)$ est scindé, alors V peut se décomposer en somme directe de sous-espaces vectoriels cycliques qui sont invariants par f. De plus, le nombre de sous-espaces cycliques associés à une valeur propre λ est égale à la multiplicité géométrique de cette valeur propre.*

Ce théorème dit qu'il existe des sous-espaces vectoriels $V_1, \dots, V_q \subset V$ tels que

(i) $V = V_1 \oplus \cdots \oplus V_q$.

(ii) Les sous-espaces V_j sont invariants par f, i.e. $f(V_j) \subset V_j$ pour $j = 1, \dots, q$.

(iii) La restriction $f_j = f|_{V_i}$ est un endomorphisme cyclique de V_j pour une valeur propre λ_j.

(iv) Pour chaque valeur propre λ_k il y a m_k sous-espaces cycliques, où m_k est la multiplicité géométrique de λ_k.

En particulier V_j est un sous-espace vectoriel d'un sous-espace caractéristique $N_{\lambda_k}(f) \subset V$ (car chaque élément de V_j est un vecteur propre généralisé pour une valeur propre λ_k).

Preuve. La preuve est assez longue et se décompose en plusieurs étapes.

Première étape : réduction au cas d'un endomorphisme n'ayant qu'une valeur propre. Soit $\sigma(f) = \{\lambda_1, \dots, \lambda_r\}$ le spectre de f. Pour tout i, on note $N_{\lambda_i}(f) \subset V$ le sous-espace caractéristique associé à λ_i. Le théorème de décomposition primaire, avec le théorème de Cayley-Hamilton, nous dit que l'espace V est somme directe des $N_{\lambda_i}(f)$ et que ces espaces sont invariants par f. Pour démontrer le théorème de Jordan, nous pouvons donc supposer que f n'a qu'une seule valeur propre λ (i.e. $V = N_\lambda(f)$).

Deuxième étape : réduction au cas d'un endomorphisme nilpotent.

On suppose donc que $f \in \mathcal{L}(V)$ est un endomorphisme qui n'a qu'une valeur propre λ et dont le polynôme caractéristique est scindé. On sait par le lemme 9.9.1 que ces hypothèses impliquent que $g = (f - \lambda \mathrm{Id}_V)$ est nilpotent. Il est clair que $W \subset V$ est un sous-espace vectoriel invariant par f si et seulement si W est invariant par g. De plus, W est λ-cyclique pour f si et seulement si ce sous-espace est cyclique pour g (associé à l'unique valeur propre de g, qui est 0). Pour démontrer le théorème de Jordan, nous pouvons donc supposer sans perte de généralité que f est nilpotent.

Troisième étape : le cas nilpotent.

Il suffit donc de démontrer que *si $f \in \mathcal{L}(V)$ est un endomorphisme nilpotent d'un espace vectoriel V de dimension finie, alors V est somme directe de q sous-espaces invariants cycliques, où $q = \dim(\mathrm{Ker}(f))$.*

(On rappelle qu'un endomorphisme nilpotent n'a qu'une valeur propre, qui est 0, la multiplicité géométrique de cette valeur propre est la dimension de $\mathrm{Ker}(f)$).

La preuve se fait par récurrence sur l'ordre de nilpotence m de f. Si f est nilpotent d'ordre 1, alors f est l'endomorphisme nul. On peut choisir une base quelconque $\{v_1, \dots, v_n\}$ de V et noter $V_j = Kv_j = \mathrm{Vec}(v_i)$ le sous-espace vectoriel de dimension 1 engendré par v_j. Alors chaque V_j est trivialement invariant par f et cyclique d'ordre 1, et on a $V = V_1 \oplus \cdots \oplus V_n$. De plus $n = \dim(V) = \dim(\mathrm{Ker}(f))$, la preuve est donc complète pour le cas $m = 1$.

On suppose maintenant que l'affirmation est démontrée pour les endomorphismes nilpotents d'ordre $m-1$ avec $m \geq 2$ et on considère le cas d'un endomorphisme $f \in \mathcal{L}(V)$ nilpotent d'ordre $m \geq 2$. Notons $W = \mathrm{Im}(f)$ et choisissons un sous-espace de $\mathrm{Ker}(f)$ complémentaire à $W \cap \mathrm{Ker}(f)$ qu'on note U. On a donc

$$\mathrm{Ker}(f) = (\mathrm{Ker}(f) \cap W) \oplus U.$$

Les sous-espaces U et W de V sont invariants par f (car le noyau et l'image d'un endomorphisme sont toujours des sous-espaces invariants). La restriction de f à U est l'endomorphisme nul et la restriction de f à W est un endomorphisme nilpotent d'ordre $(m-1)$.

Par hypothèse de récurrence, il existe une décomposition de W en somme directe de sous-espaces invariants cycliques :

$$W = \mathrm{Im}(f) = W_1 \oplus \cdots \oplus W_q, \qquad f(W_j) \subset W_j \text{ et } W_j \text{ est cyclique pour } f,$$

de plus $q = \dim(\mathrm{Ker}(f|_W)) = \dim(\mathrm{Ker}(f) \cap W)$.

Chaque sous-espace W_j admet donc une base cyclique $\mathcal{C}_j = \{w_{j,1}, \dots, w_{j,m_j}\}$, où $m_j = \dim W_j$. Rappelons que cela signifie que $f(w_{j,1}) = 0$ et $f(w_{j,i}) = w_{j,i-1}$ pour $i > 1$.

Puisque $W_j \subset W = \mathrm{Im}(f)$, il existe un vecteur $v_j \in V$ tel que $f(v_j) = w_{j,m_j}$; on note alors

$$\mathcal{B}_j = \mathcal{C}_j \cup \{v_j\} = \{w_{j,1}, \dots w_{j,m_j}, v_j\} \quad \text{et} \quad V_j = W_j + Kv_j = \mathrm{Vec}(\mathcal{B}_j).$$

Nous affirmons que les sous-espaces V_j et les familles $\mathcal{B}_j$ possèdent les propriétés suivantes :

(a) La réunion $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_q$ est une famille libre de V.

(b) $\mathcal{B}_j = \{w_{j,1}, \dots w_{j,m_j}, v_j\}$ est une base de V_j.

(c) V_j est invariant par f.

(d) V_j est cyclique d'ordre $m_j + 1$ pour f.

Pour prouver (a), il est commode de renoter par w_{j,m_j+1} le vecteur v_j. Supposons que

$$\sum_{j=1}^{q} \sum_{i=1}^{m_j+1} \alpha_{j,i} \cdot w_{j,i} = 0,$$

en appliquant f à cette relation de dépendance linéaire, on obtient que

$$\sum_{j=1}^{q} \sum_{i=2}^{m_j+1} \alpha_{j,i} \cdot w_{j,i-1} = \sum_{j=1}^{q} \sum_{i=1}^{m_j+1} \alpha_{j,i} \cdot f(w_{j,i}) = 0,$$

(on utilise que $f(w_{j,1}) = 0$). En décalant l'indice i d'une unité on peut écrire

$$\sum_{j=1}^{q}\sum_{i=1}^{m_j} \alpha_{j,i+1} \cdot w_{j,i} = 0.$$

Or cette identité implique que chaque $\alpha_{j,i+1} = 0$ pour tout $i \geq 1$ car la réunion des $\mathcal{C}_j$ est une base de W. Mais alors on a aussi

$$\sum_{j=1}^{q} \alpha_{j,1} \cdot w_{j,1} = 0,$$

et donc chaque $\alpha_{j,1} = 0$ car $\{w_{1,1}, \dots w_{q,1}\}$ est une partie libre (c'est un sous-ensemble de $\mathcal{C}_j$).

L'affirmation (b) est maintenant immédiate puisque $\mathcal{B}_j$ est une famille libre qui engendre V_j.

Les affirmations (c) et (d) découlent du fait que $f(\mathcal{B}_j) \subset \mathcal{B}_j \cup \{0\}$ et que $\mathcal{B}_j$ est une base cyclique de V_j par construction.

Nous pouvons maintenant conclure la preuve du théorème. L'affirmation (a) et la définition de U entraînent que $U \oplus V_1 \oplus \cdots \oplus V_q \subset V$ est une somme directe. Nous affirmons que

$$V = U \oplus V_1 \oplus \cdots \oplus V_q. \tag{9.13}$$

En effet, par définition de U on a

$$\dim(\mathrm{Ker}(f)) = \dim(U) + \dim(\mathrm{Ker}(f) \cap W) = \dim(U) + q,$$

et d'autre part $\dim(V_j) = \dim(W_j) + 1$ pour tout $j = 1, \dots, q$. Par conséquent

$$\begin{aligned}
\dim(U \oplus V_1 \oplus \cdots \oplus V_q) &= (\dim(\mathrm{Ker}(f)) - q) + \sum_{j=1}^{q} (\dim(W_j) + 1) \\
&= \dim(\mathrm{Ker}(f)) + \sum_{j=1}^{q} \dim(W_j) \\
&= \dim(\mathrm{Ker}(f)) + \dim(W) \\
&= \dim(V),
\end{aligned}$$

puisque $W = \mathrm{Im}(f)$. L'égalité de ces dimensions impliquent la somme directe (9.13). Finalement, la restriction de f à U est l'endomorphisme nul. On peut donc décomposer U en sous-espaces de dimension 1 (en choisissant une base quelconque), disons $U = U_1 \oplus \cdots \oplus U_p$ avec $p = \dim(U)$. On a donc la décomposition

$$V = U_1 \oplus \cdots \oplus U_p \oplus V_1 \oplus \cdots \oplus V_q,$$

en $(p+q)$-sous-espaces cycliques invariants, et

$$p + q = \dim(U) + \dim(\mathrm{Ker}(f) \cap W) = \dim(\mathrm{Ker}(f)).$$

La preuve du théorème est complète.

□

9.12 Conséquences du théorème de réduction de Jordan

On peut également énoncer le théorème 9.11.1 sous la forme suivante :

Théorème 9.12.1. *Soit $f \in \mathcal{L}(V)$ un endomorphisme d'un espace vectoriel V de dimension finie. Si le polynôme caractéristique de f est scindé, alors il existe une base $\mathcal{B}$ de V dans laquelle la matrice de f est diagonale par blocs et chaque bloc est une matrice de Jordan.*

$$J_{m_1}(\lambda_{i_1}) \oplus J_{m_2}(\lambda_{i_2}) \oplus \cdots \oplus J_{m_q}(\lambda_{i_q}) = \begin{pmatrix} \boxed{J_{m_1}(\lambda_{i_1})} & & & \\ & \boxed{J_{m_2}(\lambda_{i_2})} & & \\ & & \ddots & \\ & & & \boxed{J_{m_q}(\lambda_{i_q})} \end{pmatrix}. \tag{9.14}$$

Le nombre de blocs de Jordan associé à chaque valeur propre est égal à la multiplicité géométrique de cette valeur propre.

Une telle base $\mathcal{B}$ de V s'appelle une *base de Jordan* pour l'endomorphisme f.

Le résultat suivant calcule le nombre de blocs de Jordan de chaque taille :

Proposition 9.12.2. *Soit $f \in \mathcal{L}(V)$ comme dans le théorème précédent et λ une valeur propre de f. On note $\alpha_\lambda(m)$ le nombre de blocs de Jordan de taille m dans la matrice* (9.14) *et $\delta_\lambda(k) = \dim\left(\operatorname{Ker}(f-\lambda)^k\right)$. Alors pour tout $m \geq 1$ on a*

$$\boxed{\alpha_\lambda(m) = 2\delta_\lambda(m) - \delta_\lambda(m+1) - \delta_\lambda(m-1),} \tag{9.15}$$

où les $\delta_\lambda(k) = \dim\left(\operatorname{Ker}(f-\lambda)^k\right)$ sont les multiplicités généralisées de f.

Preuve. On rappelle que si $J = J_m(\lambda)$ est un bloc de Jordan de taille m, alors $\dim\left(\operatorname{Ker}(J-\lambda)^k\right) = \min\{m,k\}$. On a donc pour tout k

$$\delta_\lambda(k) = \sum_{j=1}^{n} \alpha_\lambda(j) \min\{j,k\}, \qquad \text{où } n = \dim(V), \tag{9.16}$$

et par conséquent :

$$2\delta_\lambda(m) - \delta_\lambda(m+1) - \delta_\lambda(m-1) = \sum_{j=1}^{n} \alpha_\lambda(j) \left(2\min\{j,m\} - \min\{j,m-1\} - \min\{j,m+1\}\right).$$

L'égalité 9.15 découle maintenant de l'identité suivante, valide pour des entiers naturels j, m quelconque et dont nous laissons la vérification en exercice :

$$2\min\{j,m\} - \min\{j,m-1\} - \min\{j,m+1\} = \begin{cases} 1, & \text{si } j = m, \\ 0, & \text{si } j \neq m. \end{cases}$$

$\square$

Remarque. En écrivant la formule (9.16) pour $k = 1$, on trouve que pour toute valeur propre λ on a

$$\delta_\lambda(1) = \sum_{q=1}^{n} \alpha_\lambda(q).$$

Le membre de gauche dans cette identité est la multiplicité géométrique de λ et le membre de droite est le nombre total de blocs de Jordan pour cette valeur propre. Ceci redémontre que le nombre de blocs de Jordan associé à chaque valeur propre est égal à la multiplicité géométrique de cette valeur propre.

Corollaire 9.12.3. *Toute matrice $A \in M_n(\mathbb{C})$ est semblable à une matrice de type (9.14). Cette matrice est unique à l'ordre des blocs près.*

Preuve. L'existence d'une forme canonique (9.14) pour toute matrice $A \in M_n(\mathbb{C})$ se déduit immédiatement du théorème de réduction de Jordan. L'unicité à l'ordre des blocs près provient de la proposition précédente qui calcule le nombre de bloc de chaque ordre associé à chaque valeur propre en fonction des dimensions des noyaux $\mathrm{Ker}(A - \lambda)^k$ pour tout k, en en observant que ces dimensions sont les mêmes pour deux matrices semblables.

□

Définition. Si $A \in M_n(\mathbb{C})$ et si $A' = P^{-1}AP$ est de type (9.14), alors on dit que A' est la *forme canonique de Jordan* de la matrice A. On notera parfois $A' = J[A]$ la forme de Jordan de la matrice A, bien qu'elle ne soit unique qu'à permutation près des blocs de Jordan.

La proposition suivante synthétise les informations principales concernant les blocs de Jordan :

Proposition 9.12.4. *Supposons que*

$$\chi_f(t) = \prod_{i=1}^{r} (t - \lambda_i)^{m_i} \quad \text{et} \quad \mu_f(t) = \prod_{i=1}^{r} (t - \lambda_i)^{s_i},$$

alors on a les propriétés suivantes de la forme normale de Jordan :

(i) La taille de chaque bloc $J_p(\lambda_i)$ est au plus égale à s_i

(ii) Pour tout i, il existe au moins un bloc de Jordan $J_{s_i}(\lambda_i)$ de taille s_i.

(iii) Le nombre total de blocs de Jordan pour λ_i est égal à la multiplicité géométrique de λ_i.

(iv) La somme des tailles des blocs de Jordan pour λ_i est égale à la multiplicité algébrique m_i de λ_i.

(v) La dimension de V est la somme des tailles de tous les blocs de Jordan.

(vi) Le nombre de blocs de Jordan de chaque taille pour la valeur propre λ_i est déterminé par les multiplicités généralisées $\delta_{\lambda_i}(k)$ $(1 \le k \le s_i)$, selon l'équation (9.15).

Chacune de ces propriétés a déjà été vue ou est une conséquence assez-simple des résultats précédents. Nous laissons la vérification en exercice.

Exemple. Le proposition précédente implique immédiatement que si $A \in M_3(K)$ est une matrice telle que $\mu_A(t) = (t - \lambda)^2$, alors sa forme normale de Jordan est

$$J[A] = J_2(\lambda) \oplus J_1(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

Une conséquence des résultats précédents est le

Théorème 9.12.5. *Soit V un espace vectoriel de dimension finie sur un corps K quelconque et $f, g \in \mathcal{L}(V)$ deux endomorphismes de V dont les polynômes caractéristiques sont scindés. Alors les conditions suivantes sont équivalentes :*

(i) f et g sont conjugués.

(ii) On a égalité de toutes les multiplicités généralisées : $\delta_{f,\lambda}(k) = \delta_{g,\lambda}(k)$ pour toute valeur propre λ et tout entier k.

(iii) f et g ont la même forme de Jordan (à l'ordre des blocs près).

De même, deux matrices $A, B \in M_n(K)$ dont les polynômes caractéristiques sont scindés sont semblables si et seulement si elles ont les mêmes multiplicités généralisées et donc la même forme de Jordan (à l'ordre des blocs près).

La condition (ii) de ce théorème s'exprime parfois en disant que le spectre $\sigma(f)$, avec la famille de toutes les multiplicités généralisées $\delta_{f,\lambda}(k)$ forment un *système complet d'invariants* pour la classe de conjugaison d'un endomorphisme f dont le polynôme caractéristique est scindé. Une application intéressante de ce théorème est donnée dans l'exercice suivant :

Exercice. Prouver que toute matrice $A \in M_n(\mathbb{C})$ est semblable à sa transposée $A^\top$.

9.13 Réduction pratique d'une matrice à sa forme normale de Jordan

Examinons comment réduire concrètement une matrice A à sa forme normale de Jordan. On parle parfois de "jordanisation" de la matrice, comprise comme une généralisation de la diagonalisation.

On se donne donc une matrice $A \in M_n(K)$ où K est un corps quelconque, et on suppose que le polynôme caractéristique $\chi_A(t)$ est scindé [5] :

$$\chi_A(t) = \prod_{i=1}^{r} (t - \lambda_i)^{m_i}, \qquad \sigma(A) = \{\lambda_1, \dots \lambda_r\} \subset K, \quad \{m_1, \dots, m_r\} \subset \mathbb{N}.$$

L'entier m_i est la multiplicité algébrique de la valeur propre λ_i.

La *forme normale de Jordan* de A est alors une matrice $J[A] \in M_n(K)$ qui vérifie les trois conditions suivantes :

(i) $J[A]$ est semblable à A, i.e. il existe $P \in GL_n(K)$ tel que $J[A] = P^{-1}AP$.

(ii) $J[A]$ est une matrice diagonale par bloc.

(iii) Chaque bloc est un bloc de Jordan $J_m(\lambda)$ associé à une valeur propre $\lambda \in \sigma(A)$.

L'existence et l'unicité de la matrice $J[A]$ (à l'ordre des blocs de Jordan près) ont été démontrées au paragraphe précédent. Les colonnes de la matrice P forment une *base de Jordan* pour A.

Remarque. Un bloc de Jordan de taille 1 est simplement un scalaire car $J_1(\lambda)$ est la 1×1 matrice (λ). Lorsqu'une matrice est diagonalisable, sa jordanisation n'est rien d'autre que sa diagonalisation (et chaque bloc de Jordan est de taille 1).

5. Rappelons que c'est toujours le cas si $K = \mathbb{C}$.

D'un point de vue pratique, la réduction d'une matrice A à sa forme normale de Jordan se décompose en deux problèmes :

Problème 1. Déterminer la forme de Jordan $J[A]$ de A.

Problème 2. Trouver la matrice de changement de base P telle que $J[A] = P^{-1}AP$.

La solution du problème 1 révèle la structure de l'endomorphisme associé à A. Pour jordaniser une matrice, il est préférable de résoudre le problème 1 *avant* de résoudre le problème 2. Pour trouver la matrice P il suffit de construire une base de Jordan de A.

Problème 1. Déterminer la forme normale de Jordan d'une matrice.

Rappelons la proposition 9.12.2 nous permet de déduire la forme normale de Jordan $J[A]$ d'une matrice A à partir des multiplicités généralisées. Toutefois l'équation (9.15) est assez lourde à utiliser et la proposition 9.12.4 nous donne des informations qui sont parfois suffisantes lorsqu'on connaît le polynôme caractéristique et le polynôme minimal d'une matrice A (que l'on suppose scindés).

Exemple 1. (a) On veut déterminer toutes les formes normales de Jordan possibles d'une matrice A dont le polynôme caractéristique est $\chi_A(t) = (t-2)^4$ et le polynôme minimal est $\mu_A(t) = (t-2)^2$.

Pour répondre à cette question, on observe que A est une matrice de taille 4×4 car $\deg(\chi_A(t)) = 4$). Il n'y a qu'une valeur propre, qui est $\lambda = 2$. On sait aussi $(A - \lambda I_4)$ est nilpotent d'ordre 2 car $\mu_A(t) = (t-2)^2$. Donc la forme normale de Jordan de A peut ou bien contenir deux blocs de Jordan $J_2(2)$ ou un bloc $J_2(2)$ et deux blocs $J_1(2)$. Les formes normales de Jordan possibles pour A sont donc

$$J_2(2) \oplus J_2(2) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad (\text{si } \mathrm{multgeom}_2(A) = 2)$$

et

$$J_2(2) \oplus J_1(2) \oplus J_1(2) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad (\text{si } \mathrm{multgeom}_2(A) = 3)$$

(b) Supposons que les polynômes caractéristique et minimal de la matrice A sont respectivement $\chi_A = (t-3)^3(t+1)^4$ et $\mu_A = (t-3)^2(t+1)^3$. Alors la seule forme canonique de Jordan possible pour A s'écrit, à permutation des blocs près, sous la forme suivante :

$$J[A] = J_1(3) \oplus J_2(3) \oplus J_1(-1) \oplus J_3(-1) = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

En effet par la condition (ii) énoncée plus haut il existe au moins un bloc de Jordan $J_2(3)$ et un bloc $J_3(-1)$ car $\mu_A = (t-3)^2(t+1)^3$. Mais la condition (iii) dit que la somme des tailles des blocs de Jordan pour la valeur propre 3 est égale à 3 (i.e. la multiplicité algébrique) et la somme des tailles des blocs de Jordan pour la valeur propre -1 est égale à 4. Donc il n'y qu'une façon de compléter, et c'est d'ajouter un bloc $J_1(3)$ et un bloc $J_1(-1)$.

(c) Considérons une variante où le polynôme caractéristique et minimal sont $\chi_A = (t-3)^3(t+1)^4$ et $\mu_A = (t-3)^2(t+1)^2$. Alors il y a deux formes de Jordan possibles :

$$A' = J_1(3) \oplus J_2(3) \oplus J_2(-1) \oplus J_2(-1) \quad \text{et} \quad A'' = J_1(3) \oplus J_2(3) \oplus J_2(-1) \oplus J_1(-1) \oplus J_1(-1)$$

La multiplicité géométrique de la valeurs propre -1 est égale à 2 dans le premier cas et à 3 dans le second cas.

Remarque. Il n'est pas toujours possible de déterminer la forme de Jordan uniquement à partir des polynômes caractéristique et minimal et des multiplicités géométriques. Dans un tel cas il faut calculer quelques multiplicités généralisées (au pire les calculer toutes).

Problème 2. Trouver une base de Jordan.

Pour trouver une base de Jordan d'une matrice $A \in M_n(K)$, il faut d'abord déterminer sa forme normale de Jordan (sinon on ne sait pas ce qu'on cherche). Une base de Jordan est une base formée de vecteurs propres généralisé qui forment une famille de cycles. Il y a autant de cycles que de blocs de Jordan et la longueur de chaque cycle correspond à la taille du bloc de Jordan correspondant. Un cycle associé à la valeur propre λ est une suite de vecteurs $\{u_1, \dots, u_m\}$ telle que

$$(A-\lambda)(u_m) = u_{m-1}, (A-\lambda)(u_{m-1}) = u_{m-2}, \dots, (A-\lambda)(u_2) = u_1, \ (A-\lambda)(u_1) = 0.$$

En particulier u_1 est vecteur propre. Dans une base de Jordan, chaque cycle doit être maximal, i.e. $u_m \notin \operatorname{Im}(A-\lambda)^{m-1}$. La base de Jordan est construite lorsqu'on ne peut plus construire de nouveau cycles linéairement indépendant des précédents. La matrice de changement de base P est alors la matrice dont les colonnes sont les coordonnées des vecteurs de notre base. Il est important de vérifier que $J[A] = P^{-1}AP$ (ou si on préfère $PJ[A] = AP$, ce qui permet d'économiser le calcul de P^{-1}).

Exemple 2. On demande de jordaniser la matrice

$$A := \begin{pmatrix} 7 & 0 & 2 \\ 3 & 7 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

Le polynôme caractéristique est $\chi_A(t) = (t-7)^2(t-3)$, en particulier il est scindé et les valeurs propres sont 3 et 7. On a

$$(A-3I_3) = \begin{pmatrix} 4 & 0 & 2 \\ 3 & 4 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad (A-7) = \begin{pmatrix} 0 & 0 & 2 \\ 3 & 0 & 2 \\ 0 & 0 & -4 \end{pmatrix}, \quad (A-7)^2 = \begin{pmatrix} 0 & 0 & -8 \\ 0 & 0 & -2 \\ 0 & 0 & -16 \end{pmatrix}.$$

On voit facilement que $\text{multgeom}_A(7) = 1 < \text{multalg}_A(7) = 2$. Ceci implique que A n'est pas diagonalisable et que le polynôme minimal est donc $\mu_A(t) = (t-7)^2(t-3)$ et la forme normale de Jordan de A est

$$J[A] = J_1(3) \oplus J_2(7) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 1 \\ 0 & 0 & 7 \end{pmatrix}$$

Pour construire une base de Jordan on doit donc trouver un vecteur propre pour la valeur propre 3 et un cycle de longueur 2 pour la valeur propre 7.
Un vecteur propre pour $\lambda = 3$ est $\{X = (4,1,-8)\}$. Pour construire un cycle associé à la valeur propre $\lambda = 7$ on cherche d'abord un vecteur $Y_2 \in \text{Ker}(A-7)^2 \setminus \text{Ker}(A-7)$. On peut choisir $Y_2 = (1,0,0)$. Le deuxième vecteur du cycle est alors $Y_1 = (A-7)Y_2 = (0,3,0)$. La base de Jordan cherchée est

$$\{X, Y_1, Y_2\} = \{(4,1,-8),(0,3,0),(1,0,0)\}\,.$$

Pour finaliser la jordanisation, on pose

$$P = \begin{pmatrix} 4 & 0 & 1 \\ 1 & 3 & 0 \\ -8 & 0 & 0 \end{pmatrix}, \text{ alors } P^{-1} = \begin{pmatrix} 0 & 0 & -1/8 \\ 0 & 1/3 & 1/24 \\ 1 & 0 & 1/2 \end{pmatrix} \quad \text{et} \quad P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 1 \\ 0 & 0 & 7 \end{pmatrix}.$$

Exemple 3. On demande de jordaniser la matrice $B = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$.

Le polynôme caractéristique est $\chi_B(t) = t^4$ et cette matrice est donc nilpotente. On calcule que

$$B^2 = \begin{pmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad B^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

ainsi B est nilpotente d'ordre 3. Le rang de B est 2 et le rang de B^2 est 1. On en déduit que la forme normale de Jordan est

$$J[B] = J_3(0) \oplus J_1(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Cherchons une base de Jordan, elle doit contenir un cycle de longueur 3 et un cycle de longueur 1. Pour construire le cycle de longueur 3, on cherche un vecteur $X_3 \in \text{Ker}(B^3) \setminus \text{Ker}(B^2)$, par exemple $X_3 = (0,0,1)$. On complète le cycle en posant $X_2 = BX_3 = (0,0,0,2)$ et $X_1 = BX_2 = B^2X_3 = (4,0,0)$.
Le cycle de longueur 1 est maintenant donné par un vecteur du noyau de B et qui est linéairement indépendant de X_1. On peut prendre $Y_1 = (0,2,0,-1)$. On a donc construit une base de Jordan pour B :

$$\{X_1, X_2, X_3, Y_1\} = \{(4,0,0,0),\ (0,0,0,2),\ (0,0,1,0),\ (0,2,0,-1)\}\,.$$

On vérifie qu'il s'agit d'une base de Jordan en posant $Q = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & -1 \end{pmatrix}$ et en vérifiant que $Q^{-1}BQ = J[B]$ (ou si on préfère $BQ = J[B]Q$).

Exemple 4. Soit à jordaniser la matrice

$$C = \begin{pmatrix} 2 & 1 & 0 & -1 \\ 1 & 2 & -\frac{1}{2} & 1 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Le polynôme caractéristique est $\chi_C(t) = (t-2)^3(t-3)$, le spectre est $\{2, 3\}$ et on a

$$(C - 3I_4) = \begin{pmatrix} -1 & 1 & 0 & -1 \\ 1 & -1 & -\frac{1}{2} & 1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad (C - 2I_4) = \begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -\frac{1}{2} & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La valeur propre $\lambda = 3$ est de multiplicité algébrique 1 ; l'espace propre $E_3(C) = \operatorname{Ker}(C - 3I_4)$ associé à la valeur propre 3 est de dimension 1 et il est engendré par le vecteur $Y = (1, 0, 0, -1)$. La valeur propre $\lambda = 2$ est de multiplicité algébrique 3 et de multiplicité géométrique égale à $1 = \dim \operatorname{Ker}(C - 2I_4)$. la forme normale de C possède un unique bloc de Jordan pour chaque valeur propre et on peut déjà conclure que

$$J[C] = J_3(2) \oplus J_1(3) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Nous devons construire un cycle de longueur 3 pour la valeur propre $\lambda = 2$. On a on a

$$(C - 2I_4)^2 = \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{et} \qquad (C - 2I_4)^3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

On doit choisir un vecteur $X_3 \in \operatorname{Ker}\left((C - 2I_4)^3\right) \setminus \operatorname{Ker}\left((C - 2I_4)^2\right)$, prenons le vecteur $X_3 = (0, 0, -2, 0)$ (on pourrait prendre $(0, 0, 1, 0)$, mais notre choix va simplifier un peu les calculs). On définit ensuite les vecteurs $X_2 = (C - 2I_4)X_3 = (0, 1, 0, 0)$ et $X_1 = (C - 2I_4)X_2 = (1, 0, 2, 0)$. Nous avons construit notre base de Jordan $\{X_1, X_2, X_3, Y\}$. La matrice de changement de base est donnée par

$$P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & -2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

On vérifie que $CP = PJ[C]$.

9.14 Sur les endomorphismes d'espaces vectoriels réels.

Dans ce paragraphe, nous étudions la structure des endomorphismes d'un espace vectoriel V de dimension finie sur le corps $\mathbb{R}$ des réels. Rappelons que, grâce au théorème fondamental de l'algèbre, le polynôme caractéristique de tout endomorphisme d'un $\mathbb{C}$-espace vectoriel est scindé et donc il admet une base de Jordan.

Dans le cas réel, le polynôme caractéristique n'est pas toujours scindé mais nous rappelons le résultat suivant :

Lemme 9.14.1 (Décomposition d'un polynôme à coefficients réels en facteurs irréductibles)**.** *Tout polynôme $p(t) \in \mathbb{R}[t]$ peut s'écrire*

$$p(t) = \prod_{i=1}^{s}(t-\lambda_i)^{m_i} \cdot \prod_{j=1}^{r} q_j(t)^{n_i},$$

où $\lambda_j \in \mathbb{R}$ pour tout $j = 1, \dots, r$ et $q_i(t) \in \mathbb{R}[t]$ est un polynôme irréductible de degré 2 pour tout $i = 1, \dots, r$.

On dit que les termes $(t-\lambda_i)$ sont les *facteurs linéaires* (ou facteurs du premier degré) et les $q_j(t)$ sont les *facteurs quadratiques* de $p(t)$. Ils sont uniquement déterminés par le polynôme $p(t)$ à permutation des facteurs près. Remarquons que ce lemme implique en particulier que tout polynôme de $\mathbb{R}[t]$ de degré impair admet au moins une racine réelle [6].

Rappelons rapidement la peuve de ce lemme (qui a été vue aux exercices) : Il est clair que $\mathbb{R}[t] \subset \mathbb{C}[t]$, i.e. tout polynôme à coefficients réels est aussi un polynôme à coefficients complexes. En particulier $p(z)$ est un nombre complexe bien défini pour tout $z \in \mathbb{C}$. Mais pour un polynôme à coefficients réels on a $p(\bar{z}) = \overline{p(z)}$. En particulier

$$p(\lambda) = 0 \quad \Rightarrow \quad p(\bar{\lambda}) = 0.$$

En utilisant le théorème fondamental de l'algèbre, on peut maintenant factoriser $p(t)$ comme polynôme à coefficients complexes :

$$p(t) = \prod_{i=1}^{d}(t-\lambda_i)^{m_i},$$

et la remarque précédente nous dit que les racines complexes apparaissent par paires de racines conjuguées. Nous pouvons donc renuméroter les racines de la façon suivante :

$$\lambda_1, \dots, \lambda_s,\ \lambda_{s+1}, \bar{\lambda}_{s+1}, , \dots, \lambda_{s+r}, \bar{\lambda}_{s+r},$$

où $\lambda_i \in \mathbb{R}$ pour $i \leq s$ et $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$ pour $i > s$. On peut alors noter $\lambda_{s+j} = \alpha_j + \sqrt{-1}\beta_j$ (avec $\alpha_j, \beta_j \in \mathbb{R}$ et $\beta_j \neq 0$). On a finalement

$$p(t) = \prod_{i=1}^{s}(t-\lambda_i)^{m_i} \cdot \prod_{j=1}^{r}[(t-\lambda_j)(t-\bar{\lambda}_j)]^{n_j} = \prod_{i=1}^{s}(t-\lambda_i)^{m_i} \cdot \prod_{j=1}^{r} q_j(t)^{n_i},$$

où les facteurs $q_j(t)$ sont des polynômes quadratiques à coefficients réels. Plus précisément

$$\begin{aligned} q_i(t) = (t-\lambda_j)(t-\bar{\lambda}_j) &= t^2 - (\lambda_j + \bar{\lambda}_j)t + \lambda_j \cdot \bar{\lambda}_j \\ &= t^2 - 2\alpha_j t + (\alpha_j^2 + \beta_j^2) \\ &= (t-\alpha_j)^2 + \beta_j^2. \end{aligned}$$

6. Cela se démontre aussi facilement à partir du théorème de la valeur intermédiaire.

$\square$

Considérons maintenant un endomorphisme f d'un espace vectoriel réel V de dimension finie. On note

$$\sigma_{\mathbb{R}}(f) = \text{L'ensemble des racines réelles du polynôme caractéristique } \chi_f(t),$$

et

$$\sigma_{\mathbb{C}}(f) = \text{L'ensemble des racines complexes de } \chi_f(t),$$

On sait que $\sigma_{\mathbb{R}}(f)$ est l'ensemble des valeurs propres de f (i.e. $\lambda \in \sigma_{\mathbb{R}}(f)$ si et seulement si $\lambda \in \mathbb{R}$ et il existe un vecteur non nul v de V tel que $f(v) = \lambda v$), on sait aussi que

$$\lambda \in \sigma_{\mathbb{C}}(f) \quad \Leftrightarrow \quad \bar{\lambda} \in \sigma_{\mathbb{C}}(f).$$

Les éléments de $\sigma_{\mathbb{C}}(f) \setminus \mathbb{R}$ seront appelés les *valeurs propres complexes de f.*

Question : Quelle est la signification réelle des valeurs propres complexes ?

La notion suivante est la clé pour répondre à cette question :

Définition. Soit V un espace vectoriel réel (on ne suppose pas $\dim(V) < \infty$)). On appelle *complexifié* de V, et on note $V_{\mathbb{C}}$ l'espace vectoriel ainsi défini :

(i) Comme groupe abélien $V_{\mathbb{C}} = V \times V$ avec la loi de groupe du produit direct :

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2).$$

(ii) Le multiplication d'un vecteur $(u, v) \in V_{\mathbb{C}}$ par un nombre complexe $\lambda = \alpha + i\beta \in \mathbb{C}$ est définie par

$$(\alpha + i\beta) \cdot (u, v) = (\alpha u - \beta v, \beta u + \alpha v).$$

Proposition 9.14.2. *L'espace $V_{\mathbb{C}}$ est un espace vectoriel sur le corps $\mathbb{C}$ pour les opérations ainsi définies.*

Nous laissons la preuve en exercice (rappelons qu'i s'agit de vérifier 8 axiomes qui ne sont que des règles de cacluls, les 4 premiers axiomes rappellent simplement le fait connu que $V \times V$ est un groupe abélien pour la somme définie selon les composantes).

La structure de cet espace vectoriel est plus intuitive si l'on note un élément $w = (u, v) \in V_{\mathbb{C}}$ sous la forme $w = u + iv = u + \sqrt{-1}v$. Alors la multiplication par un scalaire complexe est

$$\lambda v = (\alpha + i\beta)(u + iv) = (\alpha u - \beta v) + i(\beta u + \alpha v).$$

Avec cette notation la preuve de la proposition est très facile.

Exemples 1.) Le complexifié de l'espace numérique $V = \mathbb{R}^n$ est $V_{\mathbb{C}} = \mathbb{C}^n$.

2.) Le complexifié de l'espace vectoriel $\mathbb{R}[t]$ des polynômes à coefficients réels est l'espace vectoriel $\mathbb{C}[t]$ des polynômes à coefficients complexes.

3.) Le complexifié de l'espace $C^k([a, b], \mathbb{R})$ des fonctions de classe C^k à valeurs réelles sur un intervalle $[a, b]$ est l'espace $C^k([a, b], \mathbb{C})$ des fonctions C^k à valeurs complexes sur $[a, b]$.

Sur l'espace vectoriel complexe $V_{\mathbb{C}}$ on peut définir l'opération de *conjugaison complexe* par $\overline{(u,v)} = (u,-v)$, ou si on préfère :

$$w = u + iv \in V_{\mathbb{C}} \quad \Rightarrow \quad \overline{w} = u - iv.$$

Cette opération est $\mathbb{C}$-*antilinéaire*, c'est-à-dire qu'on a les propriétés :

$$\overline{w_1 + w_2} = \overline{w_1} + \overline{w_2} \quad \text{et} \quad \overline{\lambda w} = \overline{\lambda}\,\overline{w}.$$

Un vecteur w de $V_{\mathbb{C}}$ est réel (i.e. c'est un élément de V) si et seulement si $w = \overline{w}$. On peut alors définir les *parties réelles* et *imaginaires* d'un vecteur $w \in V_{\mathbb{C}}$ du complexifié par

$$\text{Ré}(w) = \frac{w + \overline{w}}{2} \qquad \text{Im}(w) = \frac{w - \overline{w}}{2i}.$$

Observer que si $w = u + iv$, avec $u, v \in V$, alors $\text{Ré}(w) = u$ et $\text{Im}(w) = v$.

A tout endomorphisme $f \in \mathcal{L}(V)$ d'un espace vectoriel réel V on associe un endomorphisme noté $f_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$ de l'espace vectoriel complexifié $V_{\mathbb{C}}$ de V. Cet endomorphisme est simplement défini par

$$f_{\mathbb{C}}(w) = f_{\mathbb{C}}(u + iv) = f(u) + if(v).$$

On vérifie alors facilement la proposition suivante :

Proposition 9.14.3. *Si $\mathcal{B} = \{v_1, \dots, v_n\}$ est une base de l'espace vectoriel réel V, alors $\mathcal{B}$ est aussi une base du complexifié $V_{\mathbb{C}} = V + iV$. De plus si $A = M_{\mathcal{B}}(f)$ est la matrice de f dans la base $\mathcal{B}$, alors A est aussi la matrice de $f_{\mathbb{C}}$ dans cette base :*

$$M_{\mathcal{B}}(f_{\mathbb{C}}) = M_{\mathcal{B}}(f).$$

En particulier $V_{\mathbb{C}}$ a la même dimension (comme espace vectoriel complexe) que V (comme espace vectoriel réel) :

$$\dim_{\mathbb{C}}(V_{\mathbb{C}}) = \dim_{\mathbb{R}}(V).$$

Remarquons cependant que $V_{\mathbb{C}}$ est aussi un espace vectoriel sur le corps $\mathbb{R}$ et qu'on a

$$\dim_{\mathbb{R}}(V_{\mathbb{C}}) = 2\dim_{\mathbb{R}}(V).$$

Revenons à la question de la signification réelle d'une valeur propre complexe. Soit f un endomorphisme d'un espace vectoriel réel V de dimension finie et soit $\lambda = \alpha + i\beta$ une valeur propre complexe (i.e. $\chi_f(\lambda) = 0$). Par la proposition précédente, on déduit que λ est une valeur propre du complexifié $f_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$, il existe donc un vecteur $w = u + iv \in V_{\mathcal{C}}$ tel que $f_{\mathbb{C}}(w) = \lambda w$. En prenant les parties réelles et imaginaires, on a donc

$$f(u) + if(v) = f(u + iv) = (\alpha + i\beta)(u + iv) = (\alpha u - \beta v) + i(\beta u + \alpha v).$$

On a alors le résultat suivant :

Théorème 9.14.4. *Si V est un espace vectoriel de dimension finie et si $\lambda = \alpha + i\beta$ est une valeur propre complexe ($\beta \neq 0$), alors il existe deux vecteurs $u, v \in V$ linéairement indépendants tels que*

$$\begin{cases} f(u) & = \alpha u - \beta v, \\ f(v) & = \beta u + \alpha v. \end{cases}$$

En particulier le sous-espace U de V engendré par u et v est e dimension 2 et il est invariant par f.

Preuve. Il ne reste qu'à démontrer l'indépendance linéaire de u et v. Observons d'abord que $u \neq 0$. En effet, si on avait $u = 0$, alors $\beta v = \alpha u - f(u) = 0$ par la première équation. Or nous supposons que $\beta \neq 0$, donc $v = 0$ ce qui contredit l'hypothèse que $w = u + iv \neq 0$. Pour montrer l'indépendance linéaire de u et v, on suppose maintenant par l'absurde qu'il existe $\gamma \in \mathbb{R}$ tel que $v = \gamma u$, alors

$$(\beta + \alpha\gamma)u = \beta u + \alpha v = f(v) = f(\gamma u) = \gamma f(u) = \gamma(\alpha u - \beta v) = \gamma(\alpha - \beta\gamma)u.$$

Cela implique que $\beta u = -\gamma^2 \beta u$, et puisque $u \neq 0$ et $\beta \neq 0$, on en déduit que $\gamma^2 = -1$. Mais ceci est impossible puisque $\gamma \in \mathbb{R}$.

$\square$

Remarque. Le théorème reste vrai pour un espace vectoriel de dimension infinie à condition que $\lambda \in \mathbb{C}$ soit une valeur propre complexe du complexifié de l'endomorphisme considéré.

Exemple. Le complexifié de $V = C^\infty(\mathbb{R}, \mathbb{R})$ est l'espace vectoriel $V_\mathbb{C} = C^\infty(\mathbb{R}, \mathbb{C})$ des fonctions infiiniment différentiables sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$, et l'opérateur de dérivation $D = \frac{d}{dx}$ s'étend naturellement à $C^\infty(\mathbb{R}, \mathbb{C})$. Tout nombre complexe $\lambda \in \mathbb{C}$ est valeur propre car

$$D(e^{\lambda x}) = \lambda e^{\lambda x}.$$

Si $\lambda = \alpha + i\beta$, avec $\alpha, \beta \in \mathbb{R}$ et $\beta \neq 0$, alors

$$e^{\lambda x} = e^{(\alpha + i\beta)x} = e^{\alpha x} e^{i\beta x} = e^{\alpha x} \left(\cos(\beta x) + i \sin(\beta x)\right).$$

Posons

$$\begin{aligned} \varphi(x) &= \text{Ré}(e^{\lambda x}) = e^{\alpha x} \cos(\beta x) \\ \psi(x) &= \text{Im}(e^{\lambda x}) = e^{\alpha x} \sin(\beta x). \end{aligned}$$

Alors on a bien

$$\begin{aligned} D\left(\varphi(x)\right) = \varphi'(x) = D\left(e^{\alpha x} \cos(\beta x)\right) &= \alpha e^{\alpha x} \cos(\beta x) - \beta e^{\alpha x} \sin(\beta x) \\ &= \alpha\varphi(x) - \beta\psi(x), \end{aligned}$$

et

$$\begin{aligned} D\left(\psi(x)\right) = \psi'(x) = D\left(e^{\alpha x} \sin(\beta x)\right) &= \beta e^{\alpha x} \cos(\beta x) + \alpha e^{\alpha x} \sin(\beta x) \\ &= \alpha\psi(x) + \beta\varphi(x). \end{aligned}$$

Remarque. Si dans le théorème précédent on suppose $\dim(V) = 2$, alors $\{u, v\}$ est une base de V et on vérifie très simplement que la matrice de f dans cette base est la matrice

$$K(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Nous avons donc le résultat suivant

Théorème 9.14.5. *Tout endomorphisme f d'un espace vectoriel réel V de dimension 2 admet une base $\mathcal{B}$ dans laquelle la matrice de f prend l'une des trois formes suivantes :*

$$\mathrm{Diag}(\alpha,\beta) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \qquad J_2(\alpha) = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix} \quad \text{ou} \quad K(\alpha,\beta) = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \tag{9.17}$$

avec $\alpha, \beta \in \mathbb{R}$ (et $\beta \neq 0$ dans le troisième cas).

Remarques. (i) On ne suppose pas que $\beta \neq \alpha$. Lorsque $\alpha = \beta$ le premier et le deuxième cas se distinguent par le polynôme minimal ($\mu_f(t) = (t-\alpha)$ dans le premier cas et $\mu_f(t) = (t-\alpha)^2$ dans le deuxième cas).
(ii) Dans le troisième cas le polynôme caractéristique est $\chi_f(t) = (t-\alpha)^2 + \beta^2$ et il n'a pas de racine réelle car on suppose $\beta \neq 0$.

Preuve. Si le polynôme caractéristique $\chi_f(t)$ est scindé, alors il existe une base de Jordan et nous sommes dans le premier ou le second cas. Si $\chi_f(t)$ n'est pas scindé comme polynôme à coefficients réels, alors il existe une paire de valeurs propres complexes conjuguées et le résultat est démontré dans le théorème précédent.

□

Corollaire 9.14.6. *Toute matrice $A \in M_2(\mathbb{R})$ est semblable à l'une des matrices de type (9.17), i.e. il existe $P \in GL_2(\mathbb{R})$ telle que $P^{-1}AP$ est l'une des matrices de (9.17).*

En dimension 3 nous avons un résultat semblable :

Théorème 9.14.7. *Tout endomorphisme f d'un espace vectoriel réel V de dimension 3 admet une base $\mathcal{B}$ dans laquelle la matrice de f prend l'une des quatre formes suivante :*

$$\mathrm{Diag}(\alpha,\beta,\gamma) = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad J_2(\alpha) \oplus J_1(\gamma) = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad J_3(\alpha) = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix},$$

ou

$$K(\alpha,\beta) \oplus J_1(\gamma) = \begin{pmatrix} \alpha & \beta & 0 \\ -\beta & \alpha & 0 \\ 0 & 0 & \gamma \end{pmatrix},$$

avec $\alpha, \beta, \gamma \in \mathbb{R}$ (et $\beta \neq 0$ dans le dernier cas).

Ce théorème peut naturellement ce reformuler en termes de matrices $A \in M_3(\mathbb{R})$.

Preuve. Si le polynôme caractéristique $\chi_f(t)$ est scindé (comme polynôme à coefficients réel), alors il existe une base de Jordan et la matrice de f dans cette base possède un, deux ou trois blocs de Jordan, ce qui nous donne une des trois premières matrices. Si $\chi_f(t)$ n'est pas scindé sur les réels, alors il admet une racine réelle que nous notons γ et deux racines complexes conjuguées $\alpha \pm i\beta$.
Il existe donc une paire de vecteurs $u, v \in V$ vérifiant les équations du Théorème 9.14.4. Il existe aussi un vecteur propre $z \in V$ pour la valeur propre γ. On vérifie que les vecteurs $u, v, z \in V$ sont nécessairement linéairement indépendants. Ils forment donc une base de V et dans cette base la matrice de f est $K(\alpha,\beta) \oplus J_1(\gamma)$.

□

Ce théorème admet la généralisation suivante :

Théorème 9.14.8. *Soit f un endomorphisme d'un espace vectoriel réel V de dimension finie. Notons γ_i ses valeurs propres réelles ($i = 1, \dots, r$) et $\alpha_j \pm \sqrt{-1}\beta_j$ ses paires de valeurs propres complexes conjuguées, $j = 1, \dots, s$ (on suppose $\beta_j \neq 0$). On suppose que les valeurs propres complexes sont deux-à-deux distinctes. Alors il existe une base de V pour la quelle la matrice de f prend la forme*

$$\widetilde{J} \oplus K(\alpha_1, \beta_1) \oplus \cdots K(\alpha_s, \beta_s),$$

où $\widetilde{J}$ est une matrice générale de Jordan de valeurs propres $\gamma_1, \dots, \gamma_j$.

Généralisation. Mentionnons pour terminer que si les valeurs propres complexes ont des multiplicités ≥ 1, alors on peut encore trouver une base dans laquelle la matrice de f prend une forme standard, que nous expliquons maintenant.

Si V est un espace vectoriel de dimension 4 et $f \in \mathcal{L}(V)$ est un endomorphisme dont le polynôme caractéristique s'écrit

$$\chi_f(t) = (t - \lambda)^2 (t - \overline{\lambda})^2 = \left((t - \alpha)^2 + \beta^2\right)^2,$$

alors il existe une base de V dans laquelle la matrice de f prend l'une des deux formes suivantes :

$$K(\alpha, \beta) \oplus K(\alpha, \beta) = \begin{pmatrix} \alpha & \beta & 0 & 0 \\ -\beta & \alpha & 0 & 0 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{pmatrix}$$

ou

$$K_2(\alpha, \beta) = \begin{pmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{pmatrix}.$$

La première matrice est la matrice diagonale par blocs $K(\alpha, \beta) \oplus K(\alpha, \beta)$ et la seconde matrice possède les même blocs en diagonale et une 2×2 matrice identité I_2 en sur-diagonale. Les deux matrices ont même polynôme caractéristique $\chi_f(t) = \left((t - \alpha)^2 + \beta^2\right)^2$. Le polynôme minimal de la première matrice est $\mu(t) = (t - \alpha)^2 + \beta^2$ et le polynôme minimal de la seconde matrice est $\mu(t) = \left((t - \alpha)^2 + \beta^2\right)^2$.

Cette structure se généralise en toute dimension : tout endomorphisme d'une espace vectoriel réel admet une base dans laquelle sa matrice est est un produit direct de blocs de Jordan et de matrices du type

$$K_{2m}(\alpha, \beta) = \begin{pmatrix} K(\alpha, \beta) & \mathrm{I}_2 & & & \\ 0_2 & K(\alpha, \beta) & & & \\ & & \ddots & & \\ & & & K(\alpha, \beta) & \mathrm{I}_2 \\ & & & 0_2 & K(\alpha, \beta) \end{pmatrix}$$

La matrice $K_{2m}(\alpha, \beta)$ est la matrice réelle de taille $4m \times 4m$ formée de m blocs qui sont des matrices $K(\alpha, \beta)$ et $(m - 1)$ blocs I_2 en surdiagonale. Son polynôme minimal est égale à son polynôme caractéristique :

$$\chi(t) = \mu(t) = \left((t - \alpha)^2 + \beta^2\right)^m.$$

Annexe : Qui est Jordan ?

Le nom "Jordan" en algèbre peut faire référence à l'une des trois personnes suivantes :

• Wilhelm Jordan (1842– 1899), géographe allemand. La méthode d'échelonnage systématique pour les systèmes linéaire lui est attribuée, il voyait cet algorithme comme un raffinement de la méthode d'élimination qu'il attribuait à Gauss.Toutefois il semble que la méthode de Gauss-Jordan a été d'abord découverte par le mathématicien Belge Clasen en 1888.

• Camille Jordan (1838–1922), mathématicien français, professeur à l'École Polytechnique de Paris. Les formes canoniques de Jordan apparaissent dans son livre *Traité des substitutions et des équations algébriques* (1870).

• Pascual Jordan (1902–1980), physicien et mathématicien allemand. On lui doit d'importantes contributions en mécanique quantique et en théorie quantique des champs. Son nom est aussi attaché aux algèbres de Jordan, qui sont une classe d'algèbres non associatives qui sont utilisées dans la formalisation des observables en mécanique quantique.