

Sorbonne Université
Licence de Mathématiques
Année 2019–2020

LU2MA123 - Algèbre linéaire et bilinéaire IIb

RÉDUCTION DES ENDOMORPHISMES

par Alexandru OANCEA

23 avril 2020

Remerciements :

à Vincent Humilière, Patrick Polo, Pierre-Antoine Guihéneuf, Marco Maculan pour m'avoir prêté leurs photocopiés et archives ;

à Adrien Deloro pour sa relecture tout en finesse ;

à mes étudiantes et étudiants, sans qui ce texte n'aurait tout simplement pas existé.

Bibliographie :

X. Gourdon, *Les maths en tête : Algèbre*, Ellipses, 2009.

L. Koelblen, P. Polo, V. Humilière, *Algèbre et géométrie/Algèbre linéaire 2, espaces affines*, photocopiés 2009–2016, Université Pierre et Marie Curie.

<https://webusers.imj-prg.fr/~patrick.polo/LM270/polyLM270-2013.pdf>

<https://webusers.imj-prg.fr/~vincent.humiliere/2M270-2016/Poly2M270-2016.pdf>

M. Audin, O. Debarre, *Algèbre linéaire 2*, photocopié 1998/1999, Université Louis Pasteur, Strasbourg.

<https://www.math.ens.fr/~debarre/DEUG2.pdf>

Page web du cours :

<https://webusers.imj-prg.fr/~alexandru.oancea/2020-L2-LU2MA123/LU2MA123-2020.html>

Alexandru OANCEA

Sorbonne Université

Institut de Mathématiques de Jussieu – Paris Rive Gauche (UMR 7586 du CNRS)

URL : <http://webusers.imj-prg.fr/~alexandru.oancea>

E-mail : alexandru.oancea@imj-prg.fr

TABLE DES MATIÈRES

0. Rappels : valeurs propres, vecteurs propres, diagonalisabilité	1
0.1. Somme directe de sous-espaces	1
0.2. Espaces propres et critères de diagonalisabilité	2
0.3. Polynôme caractéristique d'un endomorphisme	3
1. Polynômes d'endomorphismes	5
1.1. Critère de diagonalisabilité	5
1.2. Sous-espaces stables	6
1.3. Trigonalisation	8
1.4. Polynômes d'endomorphismes	12
1.5. Polynôme minimal	14
1.6. Théorème de Cayley-Hamilton	15
1.7. Espaces caractéristiques	17
1.8. Suite des noyaux et algorithme de trigonalisation – version 2.0	18
A. Appendice (†) : somme directe externe d'espaces vectoriels	20
B. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout	21
C. Appendice (†) : \mathbb{C} est algébriquement clos	22
2. Décompositions de Dunford et Jordan, puissances et exponentielles de matrices	25
2.1. Endomorphismes qui commutent	25
2.2. Endomorphismes nilpotents	26
2.3. Décomposition de Dunford	28
2.4. Forme normale de Jordan	32
2.5. Applications : équations différentielles linéaires et suites récurrentes	37
A. Appendice (†) : démonstration du théorème de Jordan	45
B. Appendice (†) : algorithme pour calculer une forme normale de Jordan	47
C. Appendice (†) : forme normale de Jordan et classes de conjugaison	49
D. Appendice (†) : forme normale de Jordan des matrices 3×3	50
E. Appendice : rappels sur l'exponentielle des matrices	52

CHAPITRE 0

RAPPELS : VALEURS PROPRES, VECTEURS PROPRES, DIAGONALISABILITÉ

Ce chapitre 0 est constitué de **rappels** sur la diagonalisation des endomorphismes, telle que vue dans le cours LU2MA221 “Algèbre linéaire et bilinéaire I” en L2 S3. La référence est la section §4.2 du polycopié de Yves Coudène, qui est disponible à l’adresse

<https://www.lpsm.paris/pageperso/coudene/2MA221-algebre-lineaire-bilineaire-I.html>

Nous allons revisiter ces notions dans le cours. Nous supposons connue la méthode du pivot pour la résolution des systèmes linéaires $AX = Y$.

0.1. Somme directe de sous-espaces

Définition 0.1.1 (Sous-espaces en somme directe). — Soient V un k -espace vectoriel, E_1, \dots, E_n des sous-espaces de V . (Ni V ni les E_i ne sont supposés de dimension finie.)

(1) La *somme des sous-espaces* E_1, \dots, E_n , notée $E_1 + \dots + E_n$ ou $\sum_{i=1}^n E_i$, est le sous-espace de V engendré par $E_1 \cup \dots \cup E_n$. L’on montre que

$$(*) \quad E_1 + \dots + E_n = \{x_1 + \dots + x_n : \forall i, x_i \in E_i\}.$$

(2) On dit que les E_i *sont en somme directe* si pour tous $x_1 \in E_1, \dots, x_n \in E_n$, l’égalité $x_1 + \dots + x_n = 0$ entraîne $x_1 = 0 = \dots = x_n$. Ceci équivaut à dire que tout élément x de $E_1 + \dots + E_n$ s’écrit *de façon unique* $x = x_1 + \dots + x_n$ avec $x_i \in E_i$. Dans ce cas, $E_1 + \dots + E_n$ est noté $E_1 \oplus \dots \oplus E_n$ ou $\bigoplus_{i=1}^n E_i$.

Si les sous-espaces $E_i, i = 1, \dots, n$ sont de dimension finie, alors ils sont en somme directe si et seulement si

$$(*) \quad \dim(E_1 + \dots + E_n) = \dim(E_1) + \dots + \dim(E_n).$$

Terminologie 0.1.2. — Si E_1, \dots, E_n sont en somme directe et si de plus $E_1 \oplus \dots \oplus E_n$ égale V , alors on dit que V est la somme directe des E_i .

Remarques 0.1.3. — (1) Il résulte de la définition que E_1, \dots, E_n sont en somme directe si et seulement si, pour tout $i = 1, \dots, n$, on a : $E_i \cap \sum_{j \neq i} E_j = 0$.

(2) En particulier, si $n = 2$, alors E_1 et E_2 sont en somme directe si et seulement si $E_1 \cap E_2 = (0)$.

(3) **Attention!** Si des sous-espaces sont en somme directe, leur somme n'est pas nécessairement égale à l'espace tout entier : par exemple si E_1, E_2 sont deux droites distinctes dans \mathbb{R}^3 , leur somme est directe, et c'est un plan de \mathbb{R}^3 , et non \mathbb{R}^3 tout entier!

(4) **Attention!** Si $n \geq 3$, la condition $E_i \cap E_j = \{0\}$ pour $i \neq j$ n'entraîne pas que la somme des E_i soit directe : par exemple si E_1, E_2, E_3 sont trois droites distinctes dans \mathbb{R}^2 , elles vérifient $E_i \cap E_j = \{0\}$ pour $i \neq j$, mais leur somme n'est pas directe (car $E_1 + E_2$ égale \mathbb{R}^2 donc contient E_3).

Définition 0.1.4 (Sous-espaces supplémentaires). — Soient V un espace vectoriel, E, F deux sous-espaces de V . On dit que E et F sont des sous-espaces *supplémentaires* si $V = E \oplus F$, c.-à-d., si $E \cap F = (0)$ et $E + F = V$.

Si V est de dimension finie, ceci équivaut à dire que $E \cap F = (0)$ et $\dim(E) + \dim(F) = \dim(V)$.

Proposition 0.1.5. — Soit V un k -espace vectoriel de dimension finie n . Tout sous-espace E de V admet un supplémentaire. \square

Remarque 0.1.6. — Soient V un espace vectoriel et E, F deux sous-espaces de dimension finie. Alors

$$\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F).$$

0.2. Espaces propres et critères de diagonalisabilité

Définition 0.2.1 (valeurs propres, vecteurs propres). — Soit V un k -espace vectoriel et u un endomorphisme de V . Un scalaire $\lambda \in k$ est une *valeur propre de u* s'il existe $x \in V \setminus \{0\}$ tel que $u(x) = \lambda x$, ou encore si

$$\ker(u - \lambda \text{Id}) \neq 0.$$

Un élément $x \in \ker(u - \lambda \text{Id})$ s'appelle *vecteur propre pour la valeur propre λ* . Le sous-espace

$$V_\lambda = \ker(u - \lambda \text{Id}) \subseteq V$$

s'appelle *sous-espace propre pour la valeur propre λ* .

Théorème 0.2.2. — Soient V un k -espace vectoriel (pas nécessairement de dimension finie), u un endomorphisme de V , et $\lambda_1, \dots, \lambda_r$ des valeurs propres de u deux à deux distinctes. Pour $i = 1, \dots, r$, on note

$$E_i = V_{\lambda_i} = \{v \in V \mid u(v) = \lambda_i v\}$$

le sous-espace propre associé. Alors les V_{λ_i} sont en somme directe.

Démonstration. — Montrons par récurrence sur $r \geq 1$ l'assertion suivante : si l'on a une égalité $x_1 + \dots + x_r = 0$ avec $x_i \in V_{\lambda_i}$, alors $x_1 = \dots = x_r = 0$. Il n'y a rien à démontrer pour $r = 1$, donc on peut supposer $r \geq 2$ et l'assertion établie pour $r - 1$. L'on a $u(x_1 + \dots + x_r) - \lambda_r(x_1 + \dots + x_r) = 0$, ou encore

$$(\lambda_1 - \lambda_r)x_1 + \dots + (\lambda_{r-1} - \lambda_r)x_{r-1} = 0.$$

L'hypothèse de récurrence assure que, pour tout $i = 1, \dots, r - 1$, l'on a $(\lambda_i - \lambda_r)x_i = 0$, ou encore $x_i = 0$ puisque $\lambda_i - \lambda_r \neq 0$. Or $x_1 + \dots + x_r = 0$, ce qui entraîne $x_r = 0$ et achève la démonstration. \square

La somme $\bigoplus_{i=1}^r V_{\lambda_i}$ n'est pas nécessairement égale à V ; si V est de dimension finie, c'est le cas si et seulement si il existe une base formée de vecteurs propres de u .

Définition 0.2.3 (Endomorphismes diagonalisables). — Soient V un k -espace vectoriel de dimension finie et $u \in \text{End}_k(V)$. Les conditions suivantes sont équivalentes :

- (i) il existe une base de V dans laquelle la matrice de u est diagonale.
- (ii) V admet une base formée de vecteurs propres de u ;
- (iii) les vecteurs propres de u engendrent V ;
- (iv) la somme des espaces propres de u égale V ;
- (v) V est la somme directe des espaces propres de u .

Si ces conditions sont vérifiées, on dit que u est *diagonalisable*.

Proposition 0.2.4 (Valeurs propres distinctes). — Soit $u \in \text{End}_k(V)$ ($\dim_k(V) = n$). Si u possède n valeurs propres distinctes, alors u est diagonalisable.

Démonstration. — L'endomorphisme u possède n espaces propres distincts V_1, \dots, V_n , qui sont en somme directe d'après le théorème précédent. Alors le sous-espace $E = V_1 \oplus \dots \oplus V_n$ de V est de dimension

$$n \geq \dim E = \sum_{i=1}^n \dim V_i \geq n = \dim V.$$

Ainsi $\dim E = \dim V$ et par conséquent $V = E = V_1 \oplus \dots \oplus V_n$. De plus, chaque V_i est nécessairement de dimension 1. \square

Cette proposition fournit une condition *suffisante* de diagonalisabilité. Bien entendu, cette condition n'est **pas nécessaire** : par exemple la matrice identité I_n (≥ 2) est diagonale et a toutes ses valeurs propres égales à 1.

0.3. Polynôme caractéristique d'un endomorphisme

Soit V un k -espace vectoriel de dimension finie n . L'outil fondamental pour trouver les valeurs propres d'un endomorphisme $u \in \text{End}_k(V)$ est son *polynôme caractéristique*

$$P_u(X) = \det(u - XI_n) \in k[X].$$

Celui-ci est défini de la manière suivante : l'on choisit une base quelconque de V , l'on représente u par une matrice $A \in M_n(k)$ et l'on pose

$$P_u(X) = \det(A - XI_n) \in k[X].$$

La définition ne dépend pas du choix de la base.

Proposition 0.3.1. — Un scalaire $\lambda \in k$ est valeur propre de u si et seulement si il est racine du polynôme caractéristique.

Démonstration. — L'équation $u(x) = \lambda x$, $x \in V$ équivaut à $(u - \lambda \text{Id})x = 0$. L'existence d'une solution non-nulle équivaut à la non-injectivité de l'application linéaire $u - \lambda \text{Id} : V \rightarrow V$. Puisque V est de dimension finie, ceci équivaut à l'annulation du déterminant $\det(u - \lambda I_n) = P_u(\lambda)$. \square

Corollaire 0.3.2. — Soit $u \in \text{End}_k(V)$ ($\dim_k(V) = n$). Si le polynôme caractéristique $P_u(X)$ possède n racines distinctes, alors u est diagonalisable. \square

CHAPITRE 1

POLYNÔMES D'ENDOMORPHISMES

La terme de “réduction des endomorphismes” fait référence au problème suivant.

Problème (réduction des endomorphismes). Étant donné un k -espace vectoriel V (de dimension finie n) et un endomorphisme $u \in \text{End}_k(V)$, trouver une base de V dans laquelle la matrice de u prend la forme “la plus simple possible”.

Les applications les plus simples sont les homothéties λId , $\lambda \in k$. C'est une des raisons pour lesquelles l'on a introduit dans la section précédente les notions de *valeur propre* et *sous-espace propre* de u . Dans ce chapitre, nous allons raffiner cette étude. On y démontre des critères de diagonalisation, un théorème de trigonalisation, un théorème de réduction à une forme diagonale par blocs via l'étude des sous-espaces caractéristiques, ainsi que deux théorèmes importants : le théorème de Cayley-Hamilton et le lemme des noyaux.

Les appendices contiennent des discussions de la notion de somme directe d'espaces vectoriels, du théorème de Bézout et du fait que le corps \mathbb{C} est algébriquement clos.

Sauf mention contraire tous les espaces vectoriels considérés dans ce chapitre sont de dimension finie.

1.1. Critère de diagonalisabilité

Dans la section précédente nous avons vu une condition suffisante de diagonalisabilité. Celle-ci peut être complétée en une condition nécessaire et suffisante (CNS).

Définition et proposition 1.1.1 (Multiplicités algébrique et géométrique d'une valeur propre)

Soient V un \mathbb{C} -espace vectoriel de dimension n , $u \in \text{End}_{\mathbb{C}}(V)$, $\lambda_1, \dots, \lambda_r$ les racines (deux à deux distinctes) du polynôme caractéristique $P_u(X)$ dans \mathbb{C} . D'une part, $P_u(X)$ se factorise

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

où m_i est la multiplicité de λ_i comme racine de $P_u(X)$. D'autre part, d'après la Proposition (0.3.1), $\lambda_1, \dots, \lambda_r$ sont les valeurs propres de u .

On appelle *multiplicité algébrique* (resp. *géométrique*) de la valeur propre λ_i sa multiplicité m_i comme racine de $P_u(X)$ (resp. la dimension n_i de l'espace propre V_{λ_i}).

- (1) On a $\dim V_{\lambda_i} \leq m_i$ pour tout i .
- (2) u est diagonalisable $\iff \dim V_{\lambda_i} = m_i$ pour tout i .

Démonstration. — (1) Pour tout i , soit \mathcal{C}^i une base de V_{λ_i} . Comme les espaces propres sont en somme directe, la famille $\mathcal{C} = \mathcal{C}^1 \cup \dots \cup \mathcal{C}^r$ est une famille libre, donc on peut la compléter en une base \mathcal{B} de V . Alors $A = \text{Mat}_{\mathcal{B}}(u)$ est de la forme suivante :

$$A = \left(\begin{array}{c|c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 & * \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots & * \\ \hline 0 & \ddots & \ddots & 0 & * \\ \hline \vdots & \ddots & 0 & \lambda_r I_{n_r} & * \\ \hline 0 & \cdots & 0 & 0 & B \end{array} \right)$$

où B est une matrice carrée de taille $p = n - (n_1 + \dots + n_r)$. En particulier, A est triangulaire par blocs et l'on en déduit l'égalité

$$P_u(X) = \det(A - XI_n) = P_B(X) \prod_{i=1}^r (\lambda_i - X)^{n_i}.$$

Donc $\prod_{i=1}^r (\lambda_i - X)^{n_i}$ divise $P_u(X)$, d'où $n_i \leq m_i$ pour tout i , ce qui prouve (1).

(2) Si $n_i = m_i$ pour tout i , alors le sous-espace $E = \bigoplus_{i=1}^r V_{\lambda_i}$ est de dimension $\sum_{i=1}^r m_i = n$, donc égale V , donc u est diagonalisable. Réciproquement, si u est diagonalisable, il existe une base \mathcal{B} de V telle que

$$A = \text{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c|c|c} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0 & \cdots & 0 & \lambda_r I_{n_r} \end{array} \right)$$

alors $P_u(X) = \det(A - XI_n) = \prod_{i=1}^r (\lambda_i - X)^{n_i} = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{n_i}$, d'où $n_i = m_i$ pour tout i . \square

Le critère (1.1.1) est utile en pratique : étant connues les racines du polynôme caractéristique avec leurs multiplicités, l'on calcule les sous-espaces propres correspondants. Pour déterminer si l'endomorphisme u est diagonalisable l'on compare la multiplicité algébrique à la multiplicité géométrique.

Les espaces propres sont en somme directe, mais leur somme n'est E que si u est diagonalisable. Par exemple, la seule valeur propre de la matrice carrée

$$U_s = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

est 0 et l'espace propre associé est de dimension 1.

1.2. Sous-espaces stables

La notion clé dans le problème de la réduction des endomorphismes est celle de sous-espace stable.

Définition 1.2.1 (Sous-espace stable). — Soit $u \in \text{End}_k(V)$. On dit qu'un sous-espace E de V est *stable par u* si

$$u(E) \subseteq E.$$

Dans ce cas, la restriction de u à E induit un endomorphisme de E , noté $u|_E$ ou u_E .

Exemple 1.2.2. — Les sous-espaces propres d'un endomorphisme u sont stables par u . La restriction de u à chaque sous-espace propre est une homothétie.

Remarque 1.2.3. — Supposons que E est un sous-espace stable par u , choisissons une base de E et complétons-la en une base de V . La matrice de u dans cette base est triangulaire supérieure par blocs

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

où A représente la matrice de $u|_E$ dans la base choisie.

Supposons que l'on a trouvé une décomposition $V = E_1 \oplus \cdots \oplus E_r$ en somme directe de sous-espaces stables par u , de sorte que $u(E_i) \subseteq E_i$ pour tout i . Définissons une base de V en concaténant des bases de E_1, \dots, E_r . La matrice de u dans cette base est diagonale par blocs

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_r \end{pmatrix},$$

où A_i représente la matrice de $u|_{E_i}$ dans la base choisie.

Lorsque l'endomorphisme u est diagonalisable la situation peut être reformulée en termes de sous-espaces stables de la manière suivante : l'on a une décomposition $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ en somme directe de sous-espaces propres, chaque V_{λ_i} est stable par u et sa dimension est égale à la multiplicité algébrique de λ_i . La restriction $u|_{V_{\lambda_i}}$ est une homothétie de rapport λ_i et la matrice de u dans une base constituée de vecteurs propres est diagonale.

Pour mieux saisir la notion de sous-espace stable, nous donnons maintenant un résultat et un exemple qui en font usage.

Proposition 1.2.4 (Restriction d'un endomorphisme diagonalisable)

Soient V un k -espace vectoriel de dimension finie, $u \in \text{End}_k(V)$ un endomorphisme diagonalisable, E un sous-espace de V stable par u . Alors E admet une base formée de vecteurs propres de u , i.e. la restriction u_E de u à E est diagonalisable.

Démonstration. — La preuve est une variation sur celle du Théorème 0.2.2. D'après 0.2.3, il suffit de montrer que E est engendré par des vecteurs propres de u . Comme u est diagonalisable, tout $x \in E$ s'écrit dans V comme une somme de vecteurs propres :

$$(\dagger) \quad x = x_1 + \cdots + x_r, \quad \text{avec } x_i \in V_{\mu_i} \text{ et } \mu_i \neq \mu_j \text{ si } i \neq j.$$

Montrons par récurrence sur r que pour tout $x \in E$ et toute écriture (\dagger) comme ci-dessus, chaque x_i appartient à E (ce qui prouvera le théorème). C'est OK pour $r = 1$, donc on peut supposer $r \geq 2$ et le résultat démontré pour $r - 1$. Appliquant $u - \mu_r \text{Id}_V$ à (\dagger) on obtient

$$x' = (u - \mu_r \text{Id}_V)(x) = \sum_{i=1}^{r-1} (\mu_i - \mu_r)x_i$$

et $x' \in E$ puisque E est stable par u . Donc par hypothèse de récurrence, chacun des vecteurs propres $(\mu_i - \mu_r)x_i$, $i = 1, \dots, r-1$ appartient à E , donc x_i appartient aussi à E puisque $\mu_i - \mu_r \neq 0$. En reportant ceci dans (†) on obtient $x_r \in E$, ce qui prouve le théorème. \square

Rappels 1.2.5. — Soit k un corps. Si $n \cdot 1_k = 1_k + \dots + 1_k$ (n termes) est $\neq 0$ pour tout entier $n > 0$, on dit que k est de caractéristique 0; c'est le cas par exemple pour $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Sinon, le plus petit entier $p > 0$ tel que $p \cdot 1_k = 0$ est nécessairement un nombre premier (car si $p = rs$ avec $r, s \geq 1$, l'égalité $0 = (r \cdot 1_k)(s \cdot 1_k)$ entraîne que $r \cdot 1_k = 0$ ou $s \cdot 1_k = 0$, disons $r \cdot 1_k = 0$, mais alors la minimalité de p entraîne que $r = p$); dans ce cas on dit que k est de caractéristique p . D'autre part, si V est un k -espace vectoriel et $p \in \text{End}_k(V)$, rappelons qu'on dit que p est un **projecteur** si $p^2 = p \circ p$ est égal à p .

Exemple 1.2.6 (Symétries). — Soient k un corps de caractéristique $\neq 2$ (par exemple, $k = \mathbb{Q}, \mathbb{R}$ ou \mathbb{C}), V un k -espace vectoriel de dimension n , et $s \in \text{End}_k(V)$ tel que $s^2 = \text{Id}_V$. Alors s est diagonalisable; plus précisément, soient

$$p_+ = \frac{\text{Id}_V + s}{2}, \quad p_- = \frac{\text{Id}_V - s}{2}, \quad V_{\pm} = \text{Im}(p_{\pm}).$$

Alors p_+ et p_- sont des projecteurs et l'on a :

$$V = V_+ \oplus V_- \quad \text{et} \quad \forall x \in V_{\pm}, \quad s(x) = \pm x.$$

Donc, si $s \neq \pm \text{Id}_V$, alors V_+ et V_- sont non-nuls et V_{\pm} est l'espace propre associé à la valeur propre ± 1 ; dans ce cas, s est la symétrie par rapport à V_+ parallèlement à V_- .

Pour montrer ces affirmations notons les identités suivantes : $p_+^2 = p_+$, $p_-^2 = p_-$ et $p_- = \text{Id}_V - p_+$ d'où $p_+p_- = 0 = p_-p_+$. Ainsi p_+ et p_- sont des projecteurs et l'on a $V = V_+ \oplus V_-$. De plus, si $x \in V_{\pm}$, on voit aussitôt que $s(x) = \pm x$, ce qui achève la preuve.

Remarque 1.2.6.1. — Attention, si k est de caractéristique 2, c.-à-d., si $2 = 0$ dans k (par exemple, si k est le corps à deux éléments $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$), la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(k)$

vérifie $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$ mais A n'est pas diagonalisable : en effet sa seule valeur propre est 1, donc si A était diagonalisable on aurait $A = I_2$, ce qui n'est pas le cas.

1.3. Trigonalisation

Définition 1.3.1 (Endomorphismes trigonalisables). — Soit $u \in \text{End}_k(V)$. On dit que u est *trigonalisable* s'il existe une base \mathcal{B} de V dans laquelle la matrice $A = \text{Mat}_{\mathcal{B}}(u)$ est triangulaire, disons supérieure. ⁽¹⁾

Dans ce cas, soient $\lambda_1, \dots, \lambda_n$ les coefficients diagonaux ($n = \dim_k V$), et soit X une indéterminée. Alors

$$P_u(X) = \det(A - XI_n) = \prod_{i=1}^n (\lambda_i - X)$$

donc $\lambda_1, \dots, \lambda_n$ sont les n racines (comptées avec multiplicités) du polynôme caractéristique $P_u(X)$. On voit donc qu'une condition *nécessaire* pour que u soit trigonalisable est que $P_u(X)$ ait toutes ses racines dans k . Ceci conduit à la définition suivante :

⁽¹⁾Si la matrice de u dans une base (v_1, \dots, v_n) est triangulaire supérieure, alors la matrice dans la base (v_n, \dots, v_1) est triangulaire inférieure, et vice-versa, donc on pourrait dans la définition remplacer le mot « supérieure » par « inférieure ».

Définition 1.3.2 (Polynômes scindés, corps algébriquement clos)

Soient k un corps et $P \in k[X]$ un polynôme de degré $n \geq 1$.

(1) On dit que P est *scindé* dans $k[X]$ s'il admet n racines $\lambda_1, \dots, \lambda_n$ dans k (comptées avec multiplicités), c.-à-d., si P se factorise dans $k[X]$ en produit de facteurs de degré 1 :

$$P = a(X - \lambda_1) \cdots (X - \lambda_n)$$

(où a est le coefficient dominant de P).

(2) On dit que k est *algébriquement clos* si tout polynôme $P \in k[X]$ de degré ≥ 1 est scindé. Par exemple, on sait que \mathbb{C} est algébriquement clos (une démonstration est donnée dans un appendice à ce chapitre).

Théorème 1.3.3 (Trigonalisation). — *Un endomorphisme $u \in \text{End}_k(V)$ est trigonalisable si et seulement si son polynôme caractéristique est scindé dans k .*

En particulier, lorsque $k = \mathbb{C}$, tout endomorphisme est trigonalisable.

Démonstration. — L'implication directe a été démontrée plus haut, il s'agit de prouver l'implication inverse.

On procède par récurrence sur la dimension n de V . Lorsque $n = 1$ il n'y a rien à démontrer, on suppose donc $n \geq 2$ et l'affirmation démontrée pour $n - 1$. Il existe par hypothèse une racine $\lambda \in k$ du polynôme caractéristique, c'est-à-dire une valeur propre. Soit e un vecteur propre associé, soit $E = \text{Vect}(e)$ et soit E' un supplémentaire de E dans V . Comme discuté dans la Remarque 1.2.3, la matrice de u dans une base de V constituée de e et d'une base \mathcal{C} de E' est "triangulaire supérieure par blocs" du type

$$\begin{pmatrix} \lambda & B \\ 0 & C \end{pmatrix}.$$

Par ailleurs, l'on a $P_u(X) = (\lambda - X)P_C(X)$ et, puisque $P_u(X)$ est scindé sur k , il en est de même pour $P_C(X)$. Soit $p : V \rightarrow E'$ la projection sur E' parallèlement à E . Puisque $\dim E' = n - 1$ nous pouvons appliquer l'hypothèse de récurrence à l'endomorphisme $p \circ u|_{E'} \in \text{End}_k(E')$, dont la matrice dans la base \mathcal{C} est C : dans une base convenable \mathcal{C}' de E' sa matrice est triangulaire supérieure. Si l'on ajoute cette base à e l'on obtient une base de V dans laquelle la matrice de u est triangulaire supérieure. \square

Corollaire 1.3.4 (Déterminant (resp. trace) = produit (resp. somme) des valeurs propres)

Soient $A \in M_n(\mathbb{C})$ et $\lambda_1, \dots, \lambda_n$ les n racines dans \mathbb{C} (comptées avec multiplicité) du polynôme caractéristique $P_A(X)$. Alors

$$\det(A) = \lambda_1 \cdots \lambda_n, \quad \text{Tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

Démonstration. — D'après le théorème 1.3.3, il existe $P \in \text{GL}_n(\mathbb{C})$ telle que $A' = P^{-1}AP$ soit triangulaire; notons $\lambda_1, \dots, \lambda_n$ ses coefficients diagonaux, alors

$$P_{A'}(X) = \prod_{i=1}^n (\lambda_i - X), \quad \det(A') = \lambda_1 \cdots \lambda_n, \quad \text{Tr}(A') = \lambda_1 + \cdots + \lambda_n.$$

D'autre part A' et A ont même polynôme caractéristique, même déterminant et même trace. Alors $\lambda_1, \dots, \lambda_n$ sont les racines de $P_A(X) = P_{A'}(X)$ et l'on a $\det(A) = \det(A') = \lambda_1 \cdots \lambda_n$ et $\text{Tr}(A) = \text{Tr}(A') = \lambda_1 + \cdots + \lambda_n$. \square

Remarque 1.3.5. — Par définition, un endomorphisme $u \in \text{End}_k(V)$ est trigonalisable s'il existe une base $\mathcal{B} = (v_1, \dots, v_n)$ de V telle que la matrice de u dans la base \mathcal{B} soit triangulaire supérieure. Ceci revient à dire que, pour tout $j = 1, \dots, n$, le vecteur $u(v_j)$ est combinaison linéaire des vecteurs v_i avec $i \leq j$, ou encore qu'il existe des scalaires $a_{ij} \in k$ tels que

$$u(v_j) = \sum_{i=1}^j a_{ij} v_i.$$

La démonstration du Théorème 1.3.3 est effective. Elle permet en pratique de trigonaliser un endomorphisme grâce à l'algorithme suivant.

Soit $u \in \text{End}_k(V)$ avec $\dim V = n$.

Test de trigonalisabilité. On calcule d'abord le polynôme caractéristique $P_u(X) \in k[X]$.

– Si $P_u(X)$ n'est pas scindé sur k alors l'endomorphisme u n'est pas trigonalisable et l'algorithme ne fonctionnera pas.

– On suppose donc dorénavant que P_u est scindé sur k . Il s'écrit

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$

avec $\lambda_i \in k$, $m_i \in \mathbb{N}^*$ pour $i = 1, \dots, r$ et $m_1 + \cdots + m_r = n$. L'endomorphisme u est alors trigonalisable et l'algorithme fonctionnera.

L'algorithme reçoit en entrée un sous-espace vectoriel $E \subset V$ et un endomorphisme trigonalisable $f \in \text{End}_k(E)$ dont on connaît les valeurs propres. Il produit à la sortie une base \mathcal{B} de l'espace vectoriel E dans laquelle la matrice de f est triangulaire supérieure.

Pour déterminer une base \mathcal{B} dans laquelle u est trigonalisable, on appelle l'algorithme avec $E = V$ et $f = u$. Les valeurs propres de u sont $\lambda_1, \dots, \lambda_r$.

Algorithme 1 (Algorithme de trigonalisation – version 1.0)

Étape 1. (Calcul d'espaces propres). *On calcule les espaces propres de f . Si la somme de leur dimension vaut $\dim E$ alors f est diagonalisable dans une base de vecteurs propres; on calcule une telle base et l'algorithme la renvoie comme résultat. Sinon, on passe au 2.*

Étape 2. (Calcul d'un supplémentaire). *On calcule un supplémentaire dans E pour la somme directe des espaces propres de f . On le note E' .*

Étape 3. (Trigonalisation de dimension inférieure). *Soit $p : E \rightarrow E'$ la projection linéaire sur E' parallèlement à la somme des espaces propres de f . Nous définissons*

$$f' = p \circ f|_{E'} \in \text{End}_k(E').$$

On appelle l'algorithme de façon récursive avec en entrée le couple constitué de E' et $f' \in \text{End}_k(E')$. On reçoit une base \mathcal{B}' de E' . L'algorithme renvoie une base $\mathcal{B} = \mathcal{C} \cup \mathcal{B}'$, où \mathcal{C} est une base de la somme des espaces propres de f .

Exemple 1.3.6. — Voici comment fonctionne l'algorithme en pratique sur un exemple. Considérons la matrice

$$A = \begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} \in M_3(\mathbb{R}).$$

Son polynôme caractéristique est $P_A(X) = -(X - 1)^3$. Il est scindé sur \mathbb{R} et la matrice est donc trigonalisable. Elle possède une unique valeur propre $\lambda = 1$. On note e_1, e_2, e_3 les vecteurs de la base canonique de \mathbb{R}^3 .

On lance l'algorithme avec $E = \mathbb{R}^3$ et $f = A$.

On calcule une base de l'espace propre $E_1 = \ker(A - I_3)$ par la méthode du pivot de Gauss et l'on trouve $E_1 = \text{Vect}(e_1 + e_3)$. La dimension de E_1 est strictement plus petite que 3 et par conséquent la matrice A n'est pas diagonalisable. On choisit un supplémentaire de E_1 , par exemple $E' = \text{Vect}(e_2, e_3)$. Dans la base $(e_1 + e_3, e_2, e_3)$ l'endomorphisme f est exprimé par la matrice

$$\tilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

En effet, nous avons $A \cdot (e_1 + e_3) = e_1 + e_3$, $Ae_2 = 3(e_1 + e_3) + e_2$, $Ae_3 = 3(e_1 + e_3) + 2e_2 + e_3$. Nous avons eu de la chance et la matrice A' est déjà triangulaire supérieure! ⁽²⁾

Fin

Voyons comment le choix du supplémentaire influe sur le déroulement de l'algorithme. Choisissons comme supplémentaire $E' = \text{Vect}(e_2 + e_3, e_3)$. Dans la base $(e_1 + e_3, e_2 + e_3, e_3)$ l'endomorphisme f est exprimé par la matrice

$$\tilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 6 & 3 \\ 0 & 3 & 2 \\ 0 & -2 & -1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

En effet, nous avons $A \cdot (e_1 + e_3) = e_1 + e_3$, $A \cdot (e_2 + e_3) = 6(e_1 + e_3) + 3(e_2 + e_3) - 2e_3$, $Ae_3 = 3(e_1 + e_3) + 2(e_2 + e_3) - e_3$. Puisque celle-ci n'est pas triangulaire supérieure, il faut itérer l'algorithme.

Appel
récuratif

Considérons l'endomorphisme f' de $E' = \text{Vect}(e_2 + e_3, e_3)$ donné dans cette base par le carré inférieur droit de la matrice \tilde{A} , à savoir

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

(Celui-ci est précisément la composition de la projection sur E' parallèlement à E_1 avec $f|_{E'}$.) Pour calculer $E'_1 = \ker(f' - \text{Id}_{E'})$, nous identifions E' à \mathbb{R}^2 en faisant correspondre à la base $(e_2 + e_3, e_3)$ la base canonique (e'_1, e'_2) de \mathbb{R}^2 . On trouve $\ker(A' - I_2) = \text{Vect}(e'_1 - e'_2)$, ou encore $E'_1 = \ker(f' - \text{Id}_{E'}) = \text{Vect}((e_2 + e_3) - e_3) = \text{Vect}(e_2)$. On choisit comme supplémentaire de E'_1 dans E' le sous-espace $E'' = \text{Vect}(e_3)$. Dans la base $(e_1 + e_3, e_2, e_3)$ la matrice de l'endomorphisme f est triangulaire supérieure

$$\tilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

puisque $A \cdot (e_1 + e_3) = e_1 + e_3$, $Ae_2 = 3(e_1 + e_3) + e_2$, $Ae_3 = 3(e_1 + e_3) + 2e_2 + e_3$ tel que déjà calculé plus haut.

Fin

⁽²⁾En réalité l'algorithme continue à la manière du cas examiné plus bas, mais le résultat final reste inchangé. On aurait pu inclure dans l'algorithme un test de trigonalité pour l'arrêter dès qu'une base de trigonalisation est trouvée, mais nous avons choisi de ne pas le faire pour plus de lisibilité.

1.4. Polynômes d'endomorphismes

Le but de cette section et de la suivante est d'appliquer notre compréhension de l'arithmétique des polynômes à l'algèbre linéaire, via le concept de *polynôme d'endomorphisme*. En particulier le théorème purement arithmétique de Bézout joue un rôle essentiel dans la réduction des endomorphismes.

Définition 1.4.1 (Polynômes d'endomorphismes). — Soient V un k -espace vectoriel et $u \in \text{End}_k(V)$. Étant donné un polynôme

$$Q = a_0 + a_1X + \cdots + a_dX^d$$

à coefficients dans k l'on pose

$$Q(u) = a_0 \text{Id}_V + a_1u + \cdots + a_du^d \in \text{End}_k(V),$$

où u^i désigne $u \circ \cdots \circ u$ (i fois) et $u^0 = \text{Id}_V$. Un élément de $\text{End}_k(V)$ est un *polynôme en u* s'il est de la forme $Q(u)$ pour un certain $Q \in k[X]$.

L'application

$$\phi : k[X] \rightarrow \text{End}_k(V), \quad Q \mapsto Q(u)$$

est un *homomorphisme de k -algèbres*. Par définition, ceci signifie que :

- ϕ est k -linéaire.
- ϕ est unitaire : $\phi(1) = \text{Id}_V$.
- ϕ préserve la multiplication : $\phi(QR) = \phi(Q)\phi(R)$ pour tous $Q, R \in k[X]$.

Cette dernière relation entraîne notamment

$$Q(u) \circ R(u) = (QR)(u) = (RQ)(u) = R(u) \circ Q(u).$$

En particulier, « les polynômes en u commutent entre eux » et « la composition de deux polynômes en u est encore un polynôme en u ».

Définition 1.4.2 (Polynômes de matrices). — Étant donnés $A \in M_n(k)$ et un polynôme $Q = a_0 + a_1X + \cdots + a_dX^d$ à coefficients dans k , l'on note $A^0 = I_n$ et

$$Q(A) = a_0I_n + a_1A + \cdots + a_dA^d \in M_n(k).$$

Une matrice est un *polynôme en A* si elle est de la forme $Q(A)$ pour un certain $Q \in k[X]$.

On obtient comme avant un morphisme de k -algèbres

$$\phi : k[X] \rightarrow M_n(k), \quad Q \mapsto Q(A),$$

« les polynômes en A commutent entre eux », i.e. pour tous $Q, R \in k[X]$ on a

$$Q(A)R(A) = (QR)(A) = (RQ)(A) = R(A)Q(A),$$

et « la multiplication de deux polynômes en A est encore un polynôme en A ».

Exercice 1.4.3. — Soit V un espace vectoriel de dimension n et $u \in \text{End}_k(V)$. Si A est la matrice de u dans une base \mathcal{B} , alors la matrice de $Q(u)$ dans la base \mathcal{B} est $Q(A)$. (Après avoir résolu cet exercice une première fois, rédiger une approche plus abstraite employant le concept d'isomorphisme de k -algèbres.)

Lemme 1.4.4. — Soient $u \in \text{End}_k(V)$ et $x \in V$ un vecteur propre pour une valeur propre $\lambda \in k$. Alors $Q(u)(x) = Q(\lambda)x$ pour tout polynôme $Q \in k[X]$.

En particulier, si $Q(u) = 0$, alors λ est une racine de Q .

Démonstration. — Nous avons

$$u^2(x) = u(u(x)) = u(\lambda x) = \lambda u(x) = \lambda^2 x$$

et l'on montre ainsi, par récurrence sur i , que $u^i(x) = \lambda^i x$ pour tout $i \in \mathbb{N}^*$ (et aussi pour $i = 0$, avec la convention $u^0 = \text{Id}_V$ et $\lambda^0 = 1$). La conclusion du lemme en découle. \square

Nous démontrons maintenant le Lemme des noyaux. C'est l'un des résultats les plus utiles pour la réduction des endomorphismes. Il traduit en algèbre linéaire le théorème de Bézout, démontré en appendice à la fin de ce chapitre.

Théorème 1.4.5 (Théorème de Bézout). — Soient $P_1, \dots, P_r \in k[X]$ des polynômes premiers entre eux. Il existe alors des polynômes $S_1, \dots, S_r \in k[X]$ tels que

$$P_1 S_1 + \dots + P_r S_r = 1.$$

Démonstration. — Voir l'appendice B. \square

Lemme 1.4.6 (Lemme des noyaux). — Soient $P_1, \dots, P_r \in k[X]$ des polynômes premiers entre eux deux à deux. On pose $P = P_1 \cdots P_r$. Soient V un k -espace vectoriel et $u \in \text{End}_k(V)$.

1) On a $\ker P(u) = \ker P_1(u) \oplus \dots \oplus \ker P_r(u)$.

2) Les projections $\ker P(u) \rightarrow \ker P_i(u) \subset \ker P(u)$ relatives à cette décomposition en somme directe sont des polynômes en u .

Dans la démonstration qui suit, on notera le fait que le cœur de l'argument se situe au cas $r = 2$. D'ailleurs, le passage de $r = 1$ à $r = 2$ illustre déjà l'argument de récurrence.

Démonstration. — On procède par récurrence sur $r \geq 2$ (lorsque $r = 1$ il n'y a rien à démontrer).

Soit $r = 2$. Le théorème de Bézout entraîne l'existence de polynômes S_1 et S_2 tels que $1 = S_1 P_1 + S_2 P_2$. Si $x \in \ker P_1(u) \cap \ker P_2(u)$, on a $P_1(u)(x) = P_2(u)(x) = 0$, donc

$$x = S_1(u)P_1(u)(x) + S_2(u)P_2(u)(x) = 0.$$

D'autre part $\ker P_i(u)$ est contenu dans $\ker P(u)$ pour $i = 1, 2$. Si $x \in \ker P(u)$ l'on a

$$x = \underbrace{P_1(u)(S_1(u)(x))}_{\in \ker P_2(u)} + \underbrace{P_2(u)(S_2(u)(x))}_{\in \ker P_1(u)},$$

donc $V = \ker P_1(u) + \ker P_2(u)$. Ceci montre 1). De plus, la projection sur $\ker P_1(u)$ est, sur $\ker P(u)$, égale à $(S_2 P_2)(u)$, et celle sur $\ker P_2(u)$ à $(S_1 P_1)(u)$. Ceci prouve 2).

Supposons le résultat vrai au rang r et montrons-le au rang $r + 1$. On a $P = Q_1 Q_2$ avec $Q_1 = P_1 \cdots P_r$ et $Q_2 = P_{r+1}$. Les polynômes Q_1 et Q_2 sont premiers entre eux puisque P_{r+1} est premier avec chacun des polynômes P_1, \dots, P_r . D'après le cas $r = 2$ nous avons $\ker P(u) = \ker Q_1(u) \oplus \ker Q_2(u)$ et les projections sur $\ker Q_1(u)$ et $\ker Q_2(u)$ sont des polynômes en u . Par hypothèse de récurrence on a $\ker Q_1(u) = \ker P_1(u) \oplus \dots \oplus \ker P_r(u)$ et les projections sur les facteurs $\ker P_i(u)$, $i = 1, \dots, r$ sont des polynômes en u . On déduit

$$\ker P(u) = (\ker P_1(u) \oplus \dots \oplus \ker P_r(u)) \oplus \ker P_{r+1}(u) = \ker P_1(u) \oplus \dots \oplus \ker P_{r+1}(u).$$

Les projections sont des polynômes en u comme composées de polynômes en u . \square

Nous donnons maintenant un critère de diagonalisabilité formulé en termes de polynômes d'endomorphismes. Celui-ci redémontre en particulier le Corollaire 0.3.2.

Proposition 1.4.7 (Polynômes sans racines multiples). — Soit V un k -espace vectoriel de dimension finie. Un endomorphisme $u \in \text{End}_k(V)$ est diagonalisable si et seulement s'il existe $P \in k[X]$ scindé sur k ayant toutes ses racines simples tel que $P(u) = 0$.

Démonstration. — *Condition nécessaire.* Supposons u diagonalisable. Notons $\lambda_1, \dots, \lambda_r$ les valeurs propres (distinctes) de u et $V_{\lambda_1}, \dots, V_{\lambda_r}$ les sous-espaces propres correspondants, de sorte que $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r}$. Posons

$$P = (X - \lambda_1) \dots (X - \lambda_r) \in k[X].$$

Le polynôme P est scindé dans k et il a toutes ses racines simples. Nous affirmons que $P(u) = 0$ puisque $P(u)$ s'annule sur une base constituée de vecteurs propres de u . En effet, la diagonalisabilité de u implique l'existence d'une telle base et, si x_ℓ est un vecteur propre correspondant à la valeur propre λ_ℓ , alors

$$P(u)(x_\ell) = \left(\prod_{i=1}^r (u - \lambda_i \text{Id}) \right) (x_\ell) = \left(\prod_{i \neq \ell} (u - \lambda_i \text{Id}) \right) (u - \lambda_\ell \text{Id})(x_\ell) = 0.$$

Dans la deuxième égalité nous utilisons le fait que les polynômes en u commutent.

Condition suffisante. Soit $P \in k[X]$ scindé sur k avec racines simples tel que $P(u) = 0$. Écrivons $P = c(X - \lambda_1) \dots (X - \lambda_r)$ avec les $\lambda_i \in k$ distincts et $c \neq 0$. Puisque les polynômes $X - \lambda_i$ sont premiers entre eux deux à deux le Lemme des noyaux 1.4.6 s'applique et nous obtenons

$$(\dagger) \quad V = \ker P(u) = \ker(u - \lambda_1 \text{Id}) \oplus \dots \oplus \ker(u - \lambda_r \text{Id}).$$

Chaque $\ker(u - \lambda_i \text{Id})$ qui est non-nul est un sous-espace propre de u . En ne retenant dans (\dagger) que les termes non-nuls on déduit que V est somme directe de sous-espaces propres de u . Ceci équivaut à la diagonalisabilité de u . \square

Corollaire 1.4.8 (Automorphismes d'ordre fini de \mathbb{C}^n). — Soient V un \mathbb{C} -espace vectoriel de dimension n et $u \in \text{End}_{\mathbb{C}}(V)$ tel que $u^d = \text{Id}_V$ pour un entier $d \geq 1$ (dans ce cas, on dit que u est un automorphisme d'ordre fini). Alors u est diagonalisable (et ses valeurs propres sont des racines d -èmes de l'unité). \square

1.5. Polynôme minimal

Définition 1.5.1. — Soit V un k -espace vectoriel de dimension finie et $u \in \text{End}_k(V)$. L'idéal des polynômes qui annulent u est le sous-espace vectoriel

$$I_u = \{Q \in k[X] \mid Q(u) = 0\}.$$

La terminologie est justifiée par le fait que I_u est bien un idéal de $k[X]$ au sens suivant : si $Q \in I_u$ alors $PQ \in I_u$ pour tout $P \in k[X]$. On renvoie à l'Appendice B pour une discussion plus détaillée.

Il est clair que

$$I_u \subsetneq k[X]$$

puisque $1 \notin I$. Par ailleurs

$$I_u \neq \{0\}.$$

En effet, nous pouvons écrire de façon alternative $I_u = \ker \phi$, avec $\phi : k[X] \rightarrow \text{End}_k(V)$, $Q \mapsto Q(u)$ le morphisme de k -algèbres considéré après la Définition 1.4.1. Puisque $k[X]$ est un k -espace vectoriel de dimension infinie et $\text{End}_k(V)$ est un k -espace vectoriel de dimension finie (égale à n^2), il s'ensuit que ϕ ne peut pas être injectif.

Définition et proposition 1.5.2. — Il existe un unique polynôme unitaire μ_u qui engendre l'idéal I_u , au sens où

$$I_u = (\mu_u) = \{P\mu_u \mid P \in k[X]\}.$$

Ce polynôme est appelé polynôme minimal de u .

Démonstration. — Ceci est un cas particulier du Théorème B.3 dans l'Appendice B. \square

Remarque 1.5.3. — Soient V un k -espace vectoriel de dimension n et $u \in \text{End}_k(V)$. Puisque $\dim \text{End}_k(V) = n^2$, les $n^2 + 1$ vecteurs de $\text{End}_k(V)$ donnés par les endomorphismes $\text{Id}_V, u, \dots, u^{n^2}$ forment nécessairement une famille liée sur k . Autrement dit, il existe un polynôme non nul $Q \in k[X]$ de degré $\leq n^2$ qui annule u . Ceci fournit en particulier une borne sur le degré du polynôme minimal, à savoir $\deg \mu_u \leq n^2$. Nous verrons plus bas que le théorème de Cayley-Hamilton améliore cette borne de façon inattendue à $\deg \mu_u \leq n$.

Proposition 1.5.4. — Soit V un k -espace vectoriel de dimension finie. Un endomorphisme $u \in \text{End}_k(V)$ est diagonalisable si et seulement si son polynôme minimal $\mu_u \in k[X]$ est scindé sur k avec racines simples.

Démonstration. — \Leftarrow Si μ_u est scindé sur k avec racines simples, alors u est diagonalisable d'après la proposition 1.4.7 (qui utilise le théorème de Bézout).

\Rightarrow Si u est diagonalisable nous avons vu qu'il existe un polynôme scindé sur k avec racines simples tel que $P(u) = 0$. Donc $P \in I_u$ et μ_u divise P_u . Ceci entraîne que μ_u est scindé sur k avec racines simples. \square

Remarque 1.5.5. — Il est en général délicat de calculer le polynôme minimal d'un endomorphisme. Néanmoins, son existence peut souvent être très utile.

1.6. Théorème de Cayley-Hamilton

Dans cette section nous définissons la notion d'espace caractéristique et nous en déduisons une autre sorte de réduction d'un endomorphisme u , à savoir sous forme de matrice diagonale par blocs dont la taille est égale à la multiplicité des valeurs propres. Ceci s'appuie sur un théorème important dit « de Cayley-Hamilton », ⁽³⁾ et sur le Lemme des noyaux 1.4.6.

Théorème 1.6.1 (Théorème de Cayley-Hamilton). — Soit V un \mathbb{C} -espace vectoriel de dimension n , et soient $u \in \text{End}_{\mathbb{C}}(V)$ et $P_u(X)$ son polynôme caractéristique. Alors l'endomorphisme $P_u(u)$ est nul, c.-à-d. : « u est annulé par son polynôme caractéristique ».

Démonstration. — On a $P_u(X) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$, où $\lambda_1, \dots, \lambda_n$ sont les n racines (comptées avec multiplicité) de $P_u(X)$ dans \mathbb{C} . D'après le théorème 1.3.3, il existe une base $\mathcal{B} = (f_1, \dots, f_n)$ de V dans laquelle la matrice de u est triangulaire supérieure, avec $\lambda_1, \dots, \lambda_n$ sur la diagonale. Ceci équivaut à dire que, pour tout $i = 1, \dots, n$, le sous-espace F_i de V engendré par f_1, \dots, f_i est stable par u et, plus précisément, que l'on a, pour $i = 1, \dots, n$:

$$(u - \lambda_i \text{Id}_V)(F_i) \subseteq F_{i-1},$$

⁽³⁾Le théorème a été observé et démontré pour certaines matrices 4×4 par William Rowan Hamilton (1805–1865) en 1853 et pour des matrices 3×3 par Arthur Cayley (1821–1895) en 1858. Le théorème a été démontré en toute généralité par Ferdinand Georg Frobenius (1849–1917) en 1878. Cf. https://en.wikipedia.org/wiki/Cayley-Hamilton_theorem.

avec la convention $F_0 = \{0\}$. Comme $F_n = V$, on déduit des inclusions ci-dessus que $(u - \lambda_n \text{Id}_V)(V) \subseteq F_{n-1}$, puis que $(u - \lambda_{n-1} \text{Id}_V)(u - \lambda_n \text{Id}_V)(V) \subseteq F_{n-2}$, etc., d'où finalement :

$$(u - \lambda_1 \text{Id}_V) \cdots (u - \lambda_n \text{Id}_V)(V) \subseteq F_0 = \{0\}.$$

Ceci montre que $(-1)^n P_u(u) = 0$, ou encore $P_u(u) = 0$. \square

Corollaire 1.6.2 (Cayley-Hamilton pour $k \subseteq \mathbb{C}$). — Soient k un sous-corps de \mathbb{C} (par exemple $k = \mathbb{R}$) et V un k -espace vectoriel de dimension n . Soient $u \in \text{End}_k(V)$ et $P_u(X)$ son polynôme caractéristique. Alors $P_u(u) = 0$.

Démonstration. — Soient \mathcal{B} une base de V et $A = \text{Mat}_{\mathcal{B}}(u) \in M_n(k)$. D'une part, $P_u(X) = P_A(X)$; notons P ce polynôme. D'autre part, comme $k \subseteq \mathbb{C}$, on peut considérer A comme élément de $M_n(\mathbb{C})$, donc d'après le théorème de Cayley-Hamilton on a $P(A) = 0$. Or $P(A)$ est la matrice dans la base \mathcal{B} de $P(u)$, d'où $P(u) = 0$. \square

Remarque 1.6.3. — Tout corps k peut être réalisé comme sous-corps d'un corps \bar{k} « algébriquement clos », i.e., tel que tout polynôme de à coefficients dans \bar{k} soit scindé dans \bar{k} (on dit aussi que \bar{k} est une « clôture algébrique de k »). Les preuves du Théorème 1.6.1 et du Corollaire 1.6.2 peuvent être directement adaptées pour montrer que le théorème de Cayley-Hamilton reste valable pour des k -espaces vectoriels de dimension finie avec k un corps quelconque.

Remarque 1.6.4. — Il existe des dizaines (!) de démonstrations différentes du théorème de Cayley-Hamilton, ce qui indique son statut central en algèbre linéaire. C'est un phénomène incontournable, digne d'une étude approfondie.

Dans la suite nous allons nous restreindre pour plus de commodité à des \mathbb{C} -espaces vectoriels, mais la remarque précédente montre que la plupart des résultats restent valables pour des coefficients dans un corps k quelconque.

Corollaire 1.6.5. — *Ê* Soit $u \in \text{End}_k(V)$. Le polynôme minimal μ_u divise le polynôme caractéristique P_u . Leurs racines sur k sont les mêmes, à savoir les valeurs propres de u .

Démonstration. — La première affirmation est conséquence de la définition de μ_u et du fait que P_u annule u (théorème de Cayley-Hamilton).

Si λ est valeur propre de u alors $\mu_u(\lambda) = 0$ par le Lemme 1.4.4. Réciproquement, si λ est racine de μ_u alors $P_u(\lambda) = 0$ puisque μ_u divise P_u , donc λ est valeur propre de u . \square

Remarque 1.6.6. — Le polynôme minimal joue un rôle essentiel dans la théorie des extensions finies de corps, avec des applications notamment en cryptographie. On parle par exemple de *polynôme minimal d'un nombre algébrique sur \mathbb{Q}* : c'est un cas particulier de la notion que nous étudions ici.⁽⁴⁾

⁽⁴⁾Voir A. Chambert-Loir, *Algèbre corporelle*, Éditions de l'École Polytechnique, 2005. <http://www.cmls.polytechnique.fr/perso/chambert/teach/algebre.pdf> (consultée le 10 février 2020), ou B. Martin, *Codage, cryptologie et applications*, Presses Polytechniques et Universitaires Romandes (PPUR), 2004, chapitre 4 (« Codes cycliques »).

1.7. Espaces caractéristiques

Définition 1.7.1 (Espaces caractéristiques). — Soient V un \mathbb{C} -espace vectoriel de dimension finie, $u \in \text{End}_{\mathbb{C}}(V)$ et $\lambda \in \mathbb{C}$ une valeur propre de u dont la multiplicité algébrique (i.e., sa multiplicité comme racine de $P_u(X)$) vaut m . On pose

$$V_{(\lambda)} = \ker(u - \lambda \text{Id}_V)^m$$

et on l'appelle *l'espace caractéristique associé à λ* .

Remarque 1.7.2. — 1. Tout espace caractéristique $V_{(\lambda)}$ est stable par u . En effet, $V_{(\lambda)}$ est stable par $u - \lambda \text{Id}_V$ donc aussi par $u = (u - \lambda \text{Id}_V) + \lambda \text{Id}_V$.

2. L'espace caractéristique $V_{(\lambda)}$ contient l'espace propre $V_{\lambda} = \ker(u - \lambda \text{Id}_V)$. L'inclusion est stricte en général.

Exemple 1.7.3. — Si u est l'endomorphisme de \mathbb{C}^2 de matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (i.e., $u(e_1) = 0$ et $u(e_2) = e_1$), alors $P_u(X) = X^2$ a 0 comme unique racine. Par un calcul direct (ou encore d'après le théorème de Cayley-Hamilton) on a $u^2 = 0$, donc $V_{(0)} = \mathbb{C}^2$ tandis que $V_0 = \ker(u) = \mathbb{C}e_1$.

Théorème 1.7.4 (Décomposition en espaces caractéristiques)

Soient V un \mathbb{C} -espace vectoriel de dimension n et $u \in \text{End}_{\mathbb{C}}(V)$. Écrivons $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{m_i}$, où $\lambda_1, \dots, \lambda_r$ sont les valeurs propres, deux à deux distinctes, de u et m_i est la multiplicité algébrique de λ_i . Alors :

- (1) V est la somme directe des espaces caractéristiques $V_{(\lambda_1)}, \dots, V_{(\lambda_r)}$ et les projections $V \rightarrow V_{(\lambda_i)}$ relatives à cette décomposition sont des polynômes en u .
- (2) On a $\dim V_{(\lambda_i)} = m_i$ pour tout i .

Démonstration. — Pour tout i , posons

$$P_i = (X - \lambda_i)^{m_i},$$

de sorte que

$$V_{(\lambda_i)} = \ker P_i(u).$$

Les polynômes P_i sont clairement premiers deux à deux et le Lemme des noyaux 1.4.6 s'applique. On en déduit que

$$\ker(P_1 \dots P_r)(u) = V_{(\lambda_1)} \oplus \dots \oplus V_{(\lambda_r)}.$$

Or $P_1 \dots P_r = (-1)^n P_u$ et donc, par le théorème de Cayley-Hamilton, on a

$$\ker(P_1 \dots P_r)(u) = \ker P_u(u) = V.$$

Par ailleurs, il découle aussi du lemme des noyaux que les projections $V \rightarrow V_{(\lambda_i)}$ sont des polynômes en u . Ceci démontre (1).

Pour tout $i = 1, \dots, r$ posons $d_i = \dim V_{(\lambda_i)}$, choisissons une base \mathcal{B}_i de $V_{(\lambda_i)}$ et notons $u_i = u|_{V_{(\lambda_i)}}$. Alors $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ est une base de V et, notant $A_i = \text{Mat}_{\mathcal{B}_i}(u_i)$, on a :

$$(\dagger) \quad \text{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c|c|c} A_1 & 0 & \cdots & 0 \\ \hline 0 & A_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & A_r \end{array} \right), \quad \text{d'où} \quad P_u(X) = \prod_{i=1}^r P_{u_i}(X).$$

Montrons que λ_i est la seule valeur propre de u_i . À cet effet, soit μ une valeur propre quelconque de u_i et $x \in V_{(\lambda_i)}$ un vecteur propre associé. Alors $(u_i - \lambda_i \text{Id}_V)(x) = (\mu - \lambda_i)x$, d'où $(u_i - \lambda_i \text{Id}_V)^p(x) = (\mu - \lambda_i)^p x$ pour tout $p \in \mathbb{N}^*$. Or $(u_i - \lambda_i \text{Id}_V)^{m_i}(x) = 0$ par définition de $V_{(\lambda_i)}$, ce qui entraîne $(\mu - \lambda_i)^{m_i} x = 0$ et donc $\mu = \lambda_i$ puisque $x \neq 0$.

On en déduit que $P_{u_i}(X) = (-1)^{d_i}(X - \lambda_i)^{d_i}$ et, d'après (†), on obtient $P_u(X) = (-1)^n \prod_{i=1}^r (X - \lambda_i)^{d_i}$. Ceci entraîne $d_i = m_i$ pour tout i et démontre (2). \square

Remarque 1.7.5. — Le Théorème de décomposition en espaces caractéristiques 1.7.4 est effectif, puisque les sous-espaces caractéristiques peuvent être calculés explicitement pourvu que les valeurs propres soient connues. ⁽⁵⁾ Il fournit une décomposition de V en une somme directe de sous-espaces stables de dimensions égales à la multiplicité algébrique des valeurs propres. En termes de matrices, il fournit une réduction de toute matrice $A \in M_n(\mathbb{C})$ à une matrice diagonale par blocs comme dans (†) dont la taille est égale à la multiplicité algébrique de chaque valeur propre. Cette décomposition est revisitée dans la section suivante en lien avec la trigonalisation, et dans le chapitre suivant en lien avec la décomposition de Dunford.

1.8. Suite des noyaux et algorithme de trigonalisation – version 2.0

Nous nous appuyons sur la décomposition en espaces caractéristiques pour donner un algorithme de trigonalisation plus efficace que celui de la section 1.3.

Définition et proposition 1.8.1 (Suite des noyaux et indice)

Soit $u \in \text{End}_k(V)$ avec $\dim V$ finie. Il existe un unique entier $s \in \mathbb{N}$ tel que

$$\{0\} = \ker u^0 \subsetneq \ker u \subsetneq \ker u^2 \subsetneq \dots \subsetneq \ker u^s = \ker u^{s+1} = \ker u^{s+2} = \dots$$

L'entier s s'appelle indice de u . C'est le plus petit entier naturel p tel que $\ker u^p = \ker u^{p+1}$.

Démonstration. — On a $\ker u^p \subset \ker u^{p+1}$ pour $p \in \mathbb{N}$. La suite croissante $(\dim \ker u^p)_{p \in \mathbb{N}}$ prend ses valeurs dans l'ensemble fini $\{0, \dots, \dim V\}$, elle est donc stationnaire à partir d'un certain rang. Soit $s = \min\{p \in \mathbb{N} \mid \dim \ker u^p = \dim \ker u^{p+1}\} \in \mathbb{N}$. Par définition l'on a $\ker u^p \subsetneq \ker u^{p+1}$ pour tout $p < s$ et $\ker u^s = \ker u^{s+1}$. Montrons l'égalité $\ker u^p = \ker u^{p+1}$ pour tout $p > s$. L'inclusion $\ker u^p \subset \ker u^{p+1}$ a déjà été discutée. Pour prouver l'inclusion inverse $\ker u^{p+1} \subset \ker u^p$, on considère un vecteur $x \in \ker u^{p+1}$ et on montre que $u^p(x) = 0$. Or l'égalité

$$u^{p+1}(x) = u^{s+1}(u^{p-s}(x)) = 0$$

entraîne bien

$$u^s(u^{p-s}(x)) = u^p(x) = 0$$

puisque $\ker u^s = \ker u^{s+1}$. \square

Corollaire 1.8.2. — (i) Soient V un \mathbb{C} -espace vectoriel de dimension finie, $u \in \text{End}_{\mathbb{C}}(V)$ et $\lambda \in \mathbb{C}$ une valeur propre de u de multiplicité algébrique m . La suite des noyaux de $(u -$

⁽⁵⁾Étant donné un polynôme général de degré ≥ 5 , il n'existe pas de formule « universelle » pour trouver ses racines qui soit analogue de la bien connue « $x_{1,2} = (-b \pm \sqrt{\Delta})/2a$ avec $\Delta = b^2 - 4ac$ » (Le lecteur intéressé par le sujet devra suivre un cours de théorie de Galois, ou bien consulter un livre tel A. Chambert-Loir, *loc. cit.* De telles formules existent par ailleurs en degrés 3 et 4.) Dans cette généralité, la discussion faite ici reste purement théorique. En pratique, on utilise des méthodes d'algèbre linéaire pour approximer les racines des polynômes, et en particulier pour calculer des valeurs propres de matrices.

$\lambda \text{Id}_V)^p$, $p \in \mathbb{N}$, est strictement croissante avant d'être stationnaire ; elle a comme premier terme non-nul l'espace propre V_λ et comme terme stationnaire l'espace caractéristique $V_{(\lambda)}$:

$$\{0\} \subsetneq V_\lambda = \ker(u - \lambda \text{Id}_V) \subsetneq \cdots \subsetneq V_{(\lambda)} = \ker(u - \lambda \text{Id}_V)^s = \cdots = \ker(u - \lambda \text{Id}_V)^m.$$

(ii) Soit \mathcal{B} une base de $V_{(\lambda)} = \ker(u - \lambda \text{Id}_V)^m$ obtenue en considérant d'abord une base de $V_\lambda = \ker(u - \lambda \text{Id}_V)$, que l'on complète ensuite en une base de $\ker(u - \lambda \text{Id}_V)^2$, que l'on complète ensuite en une base de $\ker(u - \lambda \text{Id}_V)^3$ etc. La matrice de $u|_{V_{(\lambda)}}$ dans la base \mathcal{B} est triangulaire supérieure avec tous les coefficients diagonaux égaux à λ .

Démonstration. — (i) La conclusion est équivalente à montrer que l'indice s de $u - \lambda \text{Id}_V$ est au plus égal à m . Soient $\lambda_1, \dots, \lambda_r$ les valeurs propres de u distinctes deux à deux, avec $\lambda_1 = \lambda$. Considérons la décomposition en espaces caractéristiques $V = V_{(\lambda_1=\lambda)} \oplus V_{(\lambda_2)} \oplus \cdots \oplus V_{(\lambda_r)}$. Chaque espace caractéristique est stable par u et donc par $u - \lambda \text{Id}_V$. Puisque $u - \lambda \text{Id}_V$ est inversible sur chaque $V_{(\lambda_i)}$ avec $i \geq 2$, on déduit $\ker(u - \lambda \text{Id}_V)^s \subset V_{(\lambda)} = \ker(u - \lambda \text{Id}_V)^m$, donc $s \leq m$ par définition de l'indice s .

(ii) Notons $E_i = \ker(u - \lambda \text{Id}_V)^i$ pour $i = 0, \dots, s$, de sorte que $\{0\} = E_0 \subsetneq E_1 \subsetneq \cdots \subsetneq E_s$. La conclusion découle alors directement de ce que

$$(u - \lambda \text{Id}_V)E_i \subseteq E_{i-1} \quad \text{pour tout } i \geq 1.$$

En effet, si $x \in E_i$ alors $(u - \lambda \text{Id}_V)^{i-1}(u - \lambda \text{Id}_V)(x) = (u - \lambda \text{Id}_V)^i(x) = 0$. □

Ce résultat fournit l'algorithme de trigonalisation suivant. L'algorithme reçoit en entrée un espace vectoriel V et un endomorphisme $u \in \text{End}_{\mathbb{C}}(V)$ dont on connaît les valeurs propres $\lambda_1, \dots, \lambda_r$ et leurs multiplicités respectives m_1, \dots, m_r . Il fournit en sortie une base \mathcal{B} de V dans laquelle la matrice de u est triangulaire supérieure.

Algorithme 2 (Algorithme de trigonalisation – version 2.0)

Pour chaque $i = 1, \dots, r$:

On calcule une base \mathcal{B}_i de $V_{(\lambda_i)} = \ker(u - \lambda_i \text{Id}_V)^{m_i}$ en calculant d'abord une base de $V_{\lambda_i} = \ker(u - \lambda_i \text{Id}_V)$, en la complétant en une base de $\ker(u - \lambda_i \text{Id}_V)^2$, en complétant celle-ci en une base de $\ker(u - \lambda_i \text{Id}_V)^3$ etc., jusqu'à arriver à une base de $\ker(u - \lambda_i \text{Id}_V)^{m_i}$.

La base \mathcal{B} recherchée est $\mathcal{B} = \mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_r$.

Exemple 1.8.3. — Voici comment fonctionne l'algorithme sur la matrice considérée dans l'exemple 1.3.6, à savoir

$$A = \begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} \in M_3(\mathbb{R}).$$

Le polynôme caractéristique est $P_A(X) = -(X - 1)^3$, il est scindé sur \mathbb{R} et la matrice est trigonalisable. Elle possède une unique valeur propre $\lambda = 1$ de multiplicité algébrique 3. On a

$$A - I_3 = \begin{pmatrix} -3 & 3 & 3 \\ -2 & 0 & 2 \\ -3 & 3 & 3 \end{pmatrix}.$$

Nous calculons une base de l'espace propre $V_1 = \ker(A - I_3)$ par l'algorithme du pivot de Gauss sur colonnes, en rajoutant la première colonne à la deuxième et à la troisième. La matrice devient échelonnée et nous trouvons $V_1 = \text{Vect}(e_1 + e_3)$.

On a

$$(A - I_3)^2 = \begin{pmatrix} -6 & 0 & 6 \\ 0 & 0 & 0 \\ -6 & 0 & 6 \end{pmatrix}$$

et nous calculons une base de $\ker(A - I_3)^2$ par l'algorithme du pivot de Gauss sur colonnes en rajoutant la première colonne à la troisième. La matrice devient échelonnée et nous trouvons $\ker(A - I_3)^2 = \text{Vect}(e_1 + e_3, e_2)$.

Enfin, nous savons que $V_{(1)} = \mathbb{R}^3 = \ker(A - I_3)^3$, ou encore $(A - I_3)^3 = 0$. (Ceci peut bien sûr être vérifié directement.) Nous complétons la base $(e_1 + e_3, e_2)$ de $\ker(A - I_3)^2$ en une base de \mathbb{R}^3 en lui adjoignant par exemple le vecteur e_3 (mais nous pourrions aussi lui adjoindre e_1 , ou $e_2 + e_3$). En conclusion, nous trouvons que la matrice A devient triangulaire supérieure dans la base $(e_1 + e_3, e_2, e_3)$ de \mathbb{R}^3 .

Vérifions la réponse : la matrice

$$T = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

est bien triangulaire supérieure.

A. Appendice (†) : somme directe externe d'espaces vectoriels

Définition A.1 (Sommes directes). — Soient V_1, \dots, V_n des k -espaces vectoriels. L'ensemble produit

$$V_1 \times \cdots \times V_n = \{(v_1, \dots, v_n) \mid v_i \in V_i\}$$

est muni d'une structure d'espace vectoriel définie « composante par composante », c.-à-d., $t \cdot (v_1, \dots, v_n) = (t \cdot v_1, \dots, t \cdot v_n)$ et $(v_1, \dots, v_n) + (v'_1, \dots, v'_n) = (v_1 + v'_1, \dots, v_n + v'_n)$.

On l'appelle la *somme directe* (externe) des V_i et on le note

$$V_1 \oplus \cdots \oplus V_n \quad \text{ou} \quad \bigoplus_{i=1}^n V_i.$$

De même, un n -uplet (v_1, \dots, v_n) (avec $v_i \in V_i$) est aussi noté $v_1 + \cdots + v_n$ ou $\sum_{i=1}^n v_i$, c.-à-d., on identifie l'élément $v_i \in V_i$ au n -uplet $(0, \dots, 0, v_i, 0, \dots, 0)$ (où v_i est à la i -ème place).

Une base de $V_1 \oplus \cdots \oplus V_n$ est obtenue en adjoignant des bases de V_1, \dots, V_n , de sorte que

$$\dim \bigoplus_{i=1}^n V_i = \sum_{i=1}^n \dim V_i.$$

Remarque A.2. — Supposons maintenant que E_1, \dots, E_n soient des sous-espaces d'un k -espace vectoriel V . D'une part, on note $E_1 + \cdots + E_n$ le sous-espace de V engendré par $E_1 \cup \cdots \cup E_n$, défini comme étant l'ensemble de toutes les sommes

$$x_1 + \cdots + x_n, \quad \text{avec } x_i \in E_i.$$

D'autre part, on peut former, la somme directe externe $S = E_1 \oplus \cdots \oplus E_n$ des E_i ; ce n'est pas un sous-espace de V , mais on a une application linéaire naturelle

$$\sigma : E_1 \oplus \cdots \oplus E_n \rightarrow V, \quad (x_1, \dots, x_n) \mapsto x_1 + \cdots + x_n$$

dont l'image est le sous-espace $E_1 + \cdots + E_n$ de V , et le noyau est le sous-espace de S formé des n -uplets (x_1, \dots, x_n) tels que $x_1 + \cdots + x_n = 0$.

On voit donc que $\ker \sigma = \{0\}$ si et seulement si les sous-espaces E_1, \dots, E_n sont en somme directe dans V , et dans ce cas σ est un isomorphisme de la somme directe externe S sur le sous-espace de V noté $E_1 \oplus \dots \oplus E_n$ en 0.1.1. Ceci justifie l'usage de la notation $E_1 \oplus \dots \oplus E_n$ dans les deux cas. Pour des espaces vectoriels arbitraires E_1, \dots, E_n , la somme directe « externe » $E_1 \oplus \dots \oplus E_n$ sera appelée simplement « somme directe ».

B. Appendice (†) : division euclidienne dans $\mathbb{C}[X]$ et théorème de Bézout

Théorème B.1 (Division euclidienne dans $k[X]$). — Soit k un corps et soit $P \in k[X]$ un polynôme de degré $d \geq 1$. Pour tout $F \in k[X]$, il existe un unique couple (Q, R) d'éléments de $k[X]$ tel que

$$F = PQ + R, \quad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

On appelle Q (resp. R) le quotient (resp. le reste) de la division euclidienne de F par P .

Démonstration. — Montrons l'existence de Q, R en procédant par récurrence sur $\deg(F)$. Si $F = 0$ ou si $\deg(F) < d = \deg(P)$, on prend $Q = 0$ et $R = F$. Soit $n \geq d$ et supposons l'existence établie pour les degrés $< n$. Soit F de degré n . Notons a le coefficient dominant de F et c celui de P . Alors $ac^{-1}X^{n-d}P$ est de degré n et de coefficient dominant a , donc $F - ac^{-1}X^{n-d}P$ est de degré $< n$. Par hypothèse de récurrence, il existe $Q, R \in k[X]$ tels que

$$F - ac^{-1}X^{n-d}P = PQ + R, \quad \text{et} \quad R = 0 \text{ ou bien } \deg(R) < \deg(P).$$

Alors $F = P(Q + ac^{-1}X^{n-d}) + R$, ce qui prouve le résultat d'existence.

Montrons l'unicité : si Q_1, R_1 vérifient les mêmes conditions, les égalités $PQ + R = F = PQ_1 + R_1$ donnent

$$P(Q - Q_1) = R_1 - R.$$

Si $Q - Q_1$ était $\neq 0$ alors $P(Q - Q_1)$ serait de degré $d + \deg(Q - Q_1) \geq d$. Or, $R_1 - R$ est nul ou de degré $< d$. Donc nécessairement $Q - Q_1 = 0$, d'où $R_1 - R = 0$, d'où $Q = Q_1$ et $R = R_1$. Ceci prouve l'unicité. \square

Définitions B.2. — 1) Soit I un sous-ensemble de $k[X]$. On dit que I est un *idéal* de $k[X]$ si c'est un sous-espace vectoriel et si, pour tout $P \in I$ et $S \in k[X]$, on a $SP \in I$.

2) L'intersection de deux idéaux est un idéal. Étant donnés des polynômes $P_1, \dots, P_r \in k[X]$, l'idéal

$$(P_1, \dots, P_r) = \bigcap_{P_1, \dots, P_r \in I \text{ idéal}} I$$

est appelé *l'idéal engendré par P_1, \dots, P_r* . C'est l'idéal le plus petit, pour l'ordre partiel donné par l'inclusion, qui contient P_1, \dots, P_r . L'on démontre l'égalité

$$(P_1, \dots, P_r) = \{S_1P_1 + \dots + S_rP_r \mid S_1, \dots, S_r \in k[X]\}.$$

3) On dit qu'un idéal I de $k[X]$ est *principal* s'il peut être engendré par un seul élément, c.-à-d., s'il existe $P \in I$ tel que $I = \{SP \mid S \in k[X]\} = (P)$.

Théorème B.3. — Soit k un corps. Tout idéal I de $k[X]$ est principal. Plus précisément, si I est un idéal non nul de $k[X]$, il existe un unique polynôme unitaire $P \in I$ tel que $I = (P)$.

Démonstration. — Si I est l'idéal nul $\{0\}$, il est engendré par le polynôme nul 0. Donc on peut supposer $I \neq \{0\}$. Dans ce cas, l'ensemble $\{\deg(Q) \mid Q \in I - \{0\}\}$ est un sous-ensemble non-vide de \mathbb{N} , donc admet un plus petit élément d . Soit $P \in I - \{0\}$ tel que $\deg(P) = d$; quitte à remplacer P par $a^{-1}P$, où a est le coefficient dominant de P , on peut supposer P unitaire.

Soit F un élément arbitraire de I , d'après le théorème B.1, on peut écrire $F = PQ + R$, avec $R = 0$ ou bien $\deg(R) < \deg(P) = d$. Comme I est un idéal, alors $PQ \in I$ et donc $R = F - PQ$ appartient à I . Si on avait $R \neq 0$, ce serait un élément non nul de I de degré $< d$, contredisant la minimalité de d . Donc $R = 0$ et donc $F = PQ$. Il en résulte que $I = \{PQ \mid Q \in k[X]\} = (P)$, i.e. I est principal, engendré par le polynôme unitaire P . De plus, P est unique. En effet, si P_1 est un second polynôme unitaire tel que $I = (P_1)$, alors il existe $Q, Q_1 \in k[X]$ tels que $P_1 = PQ$ et $P = P_1Q_1$. Il en résulte que Q et Q_1 sont de degré zéro, donc des éléments de k , et comme P et P_1 sont unitaires, l'égalité $P_1 = PQ$ entraîne $Q = 1$ d'où $P_1 = P$. \square

Théorème B.4 (Théorème de Bézout sur \mathbb{C}). — Soient $P_1, \dots, P_r \in \mathbb{C}[X]$ des polynômes non nuls, sans racine commune. Alors il existe $S_1, \dots, S_r \in \mathbb{C}[X]$ tels que $S_1P_1 + \dots + S_rP_r = 1$.

Démonstration. — Soit $I = (P_1, \dots, P_r) \subset \mathbb{C}[X]$ l'idéal engendré par P_1, \dots, P_r et $D \in \mathbb{C}[X]$ son unique générateur unitaire. Puisque les polynômes P_1, \dots, P_r n'ont pas de racine commune et \mathbb{C} est algébriquement clos, il en découle que $\deg D = 0$, ou encore $D = 1$. En effet, dans le cas contraire on aurait $\deg D > 0$, donc D aurait au moins une racine dans \mathbb{C} , qui serait aussi racine commune pour P_1, \dots, P_r puisque D divise chacun des P_i .

Or l'idéal (P_1, \dots, P_r) est l'ensemble des sommes $S_1P_1 + \dots + S_rP_r$, avec $S_1, \dots, S_r \in \mathbb{C}[X]$. En particulier $D = 1$ peut être exprimé comme une telle somme, ce qui finit la preuve. \square

Le théorème de Bézout se généralise au cas des polynômes à coefficients dans un corps quelconque (en particulier $k = \mathbb{R}$) de la manière suivante.

Définition B.5. — Les polynômes $P_1, \dots, P_r \in k[X]$ sont dits *premiers entre eux* si tout polynôme D qui les divise simultanément est nécessairement de degré 0, ou encore s'il n'existe pas de polynôme de degré ≥ 1 qui les divise simultanément.

La démonstration du Théorème B.4 s'adapte directement pour démontrer le

Théorème B.6 (Théorème de Bézout). — Soient $P_1, \dots, P_r \in k[X]$ des polynômes non nuls premiers entre eux. Il existe $S_1, \dots, S_r \in k[X]$ tels que $S_1P_1 + \dots + S_rP_r = 1$. \square

Remarque B.7. — Soit k un corps algébriquement clos (par exemple $k = \mathbb{C}$). Les polynômes $P_1, \dots, P_r \in k[X]$ sont premiers entre eux si et seulement s'ils n'ont pas de racine commune. De façon équivalente, il existe $D \in k[X]$ avec $\deg D \geq 1$ qui divise simultanément les polynômes P_1, \dots, P_r si et seulement si les P_i ont une racine commune dans k . L'implication directe découle du fait qu'un tel polynôme D a nécessairement une racine dans k , et celle-ci est aussi une racine commune des P_i . L'implication inverse découle du fait que, si $\alpha \in k$ est racine commune des P_i , alors $X - \alpha$ les divise simultanément.

C. Appendice (†) : \mathbb{C} est algébriquement clos

Théorème C.1. — \mathbb{C} est algébriquement clos, c.-à-d., tout polynôme $P \in \mathbb{C}[X]$ non constant admet une racine dans \mathbb{C} .

Démonstration. — Soit $P \in \mathbb{C}[X]$ un polynôme de degré $n \geq 1$. Sans perte de généralité, on peut supposer P unitaire, i.e. de coefficient dominant égal à 1. Écrivons

$$P = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$$

avec $a_n \neq 0$. Raisonnons par l'absurde et supposons que P ne s'annule pas sur \mathbb{C} . Alors, en particulier, $a_0 \neq 0$. Notons $|\cdot|$ la norme usuelle sur \mathbb{C} , c.-à-d., $|z| = \sqrt{z\bar{z}}$.

Montrons d'abord que la fonction continue $f : \mathbb{C} \rightarrow \mathbb{R}_+$, $z \mapsto |P(z)|$ atteint son minimum $r_0 > 0$ sur \mathbb{C} . Rappelons que $\lim_{|z| \rightarrow +\infty} |P(z)| = +\infty$, de sorte qu'il existe $R > 0$ tel que

$$(1) \quad |z| \geq R \implies |P(z)| \geq |a_0|.$$

Comme le disque fermé D de centre 0 et rayon R est compact, la fonction f y atteint son minimum r_0 , et $r_0 > 0$ puisqu'on a supposé que P ne s'annule pas. Puisque pour tout $z \notin D$ on a $f(z) = |P(z)| \geq |a_0| = |P(0)| \geq r_0$, alors r_0 est le minimum de f sur \mathbb{C} .

Soit $z_0 \in D$ tel que $f(z_0) = r_0$. En remplaçant P par le polynôme de même degré $Q(X) = P(z_0)^{-1}P(X + z_0)$, on se ramène au cas où $z_0 = 0$ et où $P(0) = 1$ est le minimum de f sur \mathbb{C} .

Notons k l'ordre d'annulation en 0 de $P - 1$. On peut alors écrire

$$P(X) = 1 + a_k X^k + \dots + a_n X^n$$

avec $1 \leq k < n$ et $a_k, a_n \neq 0$. (Le cas où $k = n$ est exclu puisque le polynôme $1 + a_n X^n$ possède des racines complexes.) Écrivons $a_k = r e^{i\theta}$ avec $r > 0$ et $\theta \in [0, 2\pi[$. Posons $z_\varepsilon = \varepsilon e^{i(\pi-\theta)/k}$ pour tout $\varepsilon \in \mathbb{R}_+^*$. Comme $z_\varepsilon^k = \varepsilon^k e^{i(\pi-\theta)} = -\varepsilon^k e^{-i\theta}$, alors

$$P(z_\varepsilon) = 1 - r\varepsilon^k + \varepsilon^k h(\varepsilon), \quad \text{où} \quad h(\varepsilon) = \sum_{j=1}^{n-k} a_{k+j} z_\varepsilon^j.$$

Comme $\lim_{\varepsilon \rightarrow 0} \varepsilon^k = 0$ et $\lim_{\varepsilon \rightarrow 0} h(\varepsilon) = 0$, il existe $\varepsilon_0 \in]0, 1[$ tel que

$$\forall \varepsilon \leq \varepsilon_0, \quad r\varepsilon^k < 1 \quad \text{et} \quad |h(\varepsilon)| \leq \frac{r}{2}.$$

On a alors

$$|P(z_{\varepsilon_0})| = |1 - r\varepsilon_0^k + \varepsilon_0^k h(\varepsilon_0)| \leq |1 - r\varepsilon_0^k| + \frac{r}{2}\varepsilon_0^k = 1 - r\varepsilon_0^k + \frac{r}{2}\varepsilon_0^k = 1 - \frac{r}{2}\varepsilon_0^k < 1.$$

Ceci contredit l'hypothèse que $1 = P(0)$ était le minimum de $f = |P|$ sur \mathbb{C} . Cette contradiction montre que l'hypothèse que P ne s'annule pas sur \mathbb{C} est impossible. Ceci achève la démonstration du théorème C.1. \square

Corollaire C.2. — *Tout polynôme $P \in \mathbb{C}[X]$ de degré $n \geq 1$ se factorise en produit de facteurs de degré 1, i.e.*

$$P = a(X - \lambda_1) \dots (X - \lambda_n),$$

où $a \in \mathbb{C}^*$ est le coefficient dominant de P .

Démonstration. — Nous procédons par récurrence sur $n \geq 1$. L'affirmation est évidente pour $n = 1$, nous pouvons donc supposer $n \geq 2$ et le résultat établi pour $n - 1$. Soit P de degré n et de coefficient dominant a . D'après le Théorème C.1 le polynôme P admet au moins une racine λ_1 dans \mathbb{C} . Faisant la division euclidienne de P par $X - \lambda_1$, on peut écrire

$$P = (X - \lambda_1)P_1 + R, \quad \text{avec} \quad R = 0 \text{ ou bien } \deg(R) < 1.$$

Donc $R = 0$ ou bien R est une constante $c \neq 0$. Or, évaluant l'égalité ci-dessus en $X = \lambda_1$, on trouve $R(\lambda_1) = P(\lambda_1) = 0$, donc nécessairement $R = 0$. Ainsi $P = (X - \lambda_1)P_1$, avec P_1 non nul, de degré $n - 1$ et de coefficient dominant a . Par hypothèse de récurrence, P_1 se factorise en $P_1 = a(X - \lambda_2) \dots (X - \lambda_n)$, et donc $P = (X - \lambda_1)P_1$ égale $a(X - \lambda_1) \dots (X - \lambda_n)$. \square

CHAPITRE 2

DÉCOMPOSITIONS DE DUNFORD ET JORDAN, PUISSANCES ET EXPONENTIELLES DE MATRICES

Résumé : Dans le §2.3 nous discutons la décomposition de Dunford, qui est un raffinement du concept de trigonalisation. Dans le §2.4 nous introduisons la forme normale de Jordan, qui est une réduction à une matrice triangulaire supérieure ayant une forme canonique. À la fin du chapitre on se place sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , on introduit les exponentielles de matrices et nous les utilisons pour résoudre des équations différentielles linéaires à coefficients constants.

La question pratique qui sous-tend ce chapitre est la suivante :

Question. *Étant donné une matrice carrée A , comment calculer de manière efficace des puissances A^N pour $N \gg 1$?*

Ici, « l'efficacité » d'un algorithme de calcul de puissances de matrices est entendue en termes de nombre de multiplications effectuées avec les entrées/coefficients, dans le corps de base. On parle aussi de « complexité d'un algorithme ».

Nous avons déjà fait un pas dans la direction d'une réponse à cette question en discutant la trigonalisabilité des matrices. Néanmoins, une forme triangulaire supérieure quelconque ne permet pas un tel calcul efficace hors de la diagonale. C'est l'une des raisons d'être des formes triangulaires raffinées que sont les décompositions de Dunford et de Jordan.

2.1. Endomorphismes qui commutent

Nous nous intéressons dans cette section aux propriétés communes à deux endomorphismes qui commutent. Cette discussion préliminaire est utilisée aussi bien dans la preuve du théorème de décomposition de Dunford, que dans ses applications aux calculs de puissances et exponentielles de matrices.

Définition 2.1.1. — Soient V un k -espace vectoriel et $u, v \in \text{End}_k(V)$. On dit que u et v commutent si $uv = vu$. On dit que u et v sont simultanément diagonalisables s'il existe une base \mathcal{B} de V les diagonalisant en même temps, i.e. telle que les matrices $M_{\mathcal{B}}(u)$ et $M_{\mathcal{B}}(v)$ soient toutes les deux diagonales.

Lemme 2.1.2. — Soient V un k -espace vectoriel et $u, v \in \text{End}_k(V)$ deux endomorphismes qui commutent. Alors v laisse stable tout espace propre de u , tout espace caractéristique de u , et l'image de u .

Démonstration. — Clairement v laisse stable $\ker u$: si $x \in \ker u$ alors $u(v(x)) = v(u(x)) = v(0) = 0$. L'énoncé plus général concernant les espaces propres et les espaces caractéristiques découle de ce que, si v commute avec u , alors v commute avec $(u - \lambda \text{Id}_V)^m$ pour tout $\lambda \in k$ et tout entier $m \geq 1$.

Montrons que v laisse stable $\text{im } u$: si $y = u(x)$, alors $v(y) = v(u(x)) = u(v(x)) \in \text{im } u$. □

Proposition 2.1.3 (Diagonalisation simultanée). — Soient V un k -espace vectoriel de dimension finie et $u, v \in \text{End}_k(V)$ deux endomorphismes diagonalisables. Alors u et v commutent si et seulement s'ils sont simultanément diagonalisables.

Démonstration. — Si u et v sont simultanément diagonalisables, alors u et v commutent clairement : deux matrices diagonales de même taille commutent toujours.

Supposons maintenant que u et v commutent. Par le Lemme 2.1.2, tout espace propre de u est stable par v . Par la Proposition 1.2.4 on déduit que la restriction de v à chaque espace propre de u est diagonalisable, ce qui entraîne la conclusion. □

Exemple 2.1.4. — Les matrices $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ sont diagonalisables mais ne commutent pas (vérification directe). Elles ne sont donc pas diagonalisables dans une même base.

Remarques 2.1.5. — 1. Si u et v sont diagonalisables et commutent, alors $u + v$ et uv sont diagonalisables. En effet, dans une base qui diagonalise simultanément u et v , les endomorphismes $u + v$ et uv sont diagonaux.

La condition de commutation est nécessaire : les matrices A, B de l'exemple 2.1.4 sont diagonalisables, ne commutent pas, et les matrices $A+B$ et AB ne sont pas diagonalisables.

2. Si u et v commutent, alors

$$(1) \quad (u + v)^p = \sum_{i=0}^p \binom{p}{i} u^{p-i} v^i$$

pour tout $p \in \mathbb{N}$. Cette formule est une instance de la formule binomiale de Newton et se démontre par récurrence. La formule n'est pas vraie si u et v ne commutent pas.

2.2. Endomorphismes nilpotents

Dans le contexte de ce cours, les endomorphismes nilpotents sont importants en vue de la décomposition de Dunford discutée dans le §2.3 : le « défaut de diagonalisabilité » d'un endomorphisme quelconque est « mesuré » par un endomorphisme nilpotent. En d'autres termes, les nilpotents forment, avec les diagonalisables, les autres briques fondamentales de la nature des endomorphismes.

Définition 2.2.1. — Soit V un k -espace vectoriel et $u \in \text{End}_k(V)$. On dit que u est *nilpotent* s'il existe un entier $\ell \geq 1$ tel que $u^\ell = 0$. Le plus petit entier ℓ ayant cette propriété est appelé *indice de nilpotence* de u .

Par extension, on parle de *matrice nilpotente* et d'*indice de nilpotence* d'une matrice $A \in M_n(k)$ en la regardant comme endomorphisme de l'espace vectoriel k^n .

Exemples 2.2.2. — 1. La matrice nulle est nilpotente. C'est l'unique matrice dont l'indice de nilpotence vaut 1.

2. La matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est nilpotente et son indice de nilpotence vaut 2.

3. La matrice

$$U_n = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \in M_n(k)$$

est nilpotente et son indice de nilpotence vaut n .

4. Toute matrice $A \in M_n(k)$ triangulaire supérieure ayant des zéros sur la diagonale⁽¹⁾ est nilpotente avec indice de nilpotence au plus égal à n . Ceci peut être vu de deux manières : en remarquant que le polynôme caractéristique est $P_A(X) = (-1)^n X^n$ et en invoquant le théorème de Cayley-Hamilton, ou bien directement en remarquant le fait que $A(\text{Vect}(e_1, \dots, e_i)) \subseteq \text{Vect}(e_1, \dots, e_{i-1})$ pour tout $i = 1, \dots, n$, où (e_1, \dots, e_n) est la base canonique de k^n et l'on utilise la convention $e_0 = 0$.

Lemme 2.2.3. — Soit $u \in \text{End}_k(V)$ un endomorphisme nilpotent. Alors u possède une unique valeur propre $\lambda = 0$. En particulier, un endomorphisme nilpotent est diagonalisable si et seulement s'il est nul.

Démonstration. — Si u est nilpotent alors u n'est pas injectif, donc 0 est une valeur propre. Par ailleurs, toute valeur propre λ est nécessairement nulle : si $x \neq 0$ est tel que $u(x) = \lambda x$ et $\ell \geq 1$ est l'indice de nilpotence de u , alors $0 = u^\ell(x) = \lambda^\ell x$, donc $\lambda = 0$.

Si u est diagonalisable alors V est somme directe des sous-espaces propres de u . Si u est nilpotent son unique sous-espace propre est $\ker u$. Ainsi $V = \ker u$, ou encore $u = 0$. \square

Le Lemme 2.2.3 montre qu'il y a bien disjonction des classes

$$\{\text{diagonalisables}\} \text{ et } \{\text{nilpotents}\}.$$

Ceci vient soutenir le point de vue heuristique exprimé au début de ce paragraphe.

Exercice 2.2.4. — Soit V un espace vectoriel de dimension finie égale à n . L'indice de nilpotence de tout endomorphisme nilpotent est $\leq n$.

Exercice 2.2.5. — Soit $u \in \text{End}_k(V)$ un endomorphisme dont le polynôme caractéristique est scindé sur k . Soient $\lambda \in k$ une valeur propre de u et $V_{(\lambda)}$ le sous-espace caractéristique associé. La restriction de $u - \lambda \text{Id}_V$ à $V_{(\lambda)}$ est un endomorphisme nilpotent de $V_{(\lambda)}$, dont l'indice de nilpotence est égal à l'indice de $u - \lambda \text{Id}_V$ au sens de la Définition 1.8.1.

⁽¹⁾Une telle matrice est appelée parfois « matrice triangulaire (supérieure) stricte ».

2.3. Décomposition de Dunford

Dans cette section nous décomposons un endomorphisme en une partie « facile⁽²⁾ », car diagonalisable, et une partie nilpotente. Ceci permettra de calculer de manière efficace de grandes puissances de matrices. Cette décomposition, dite « de Dunford », est en fait due à Jordan et Chevalley.⁽³⁾

Proposition 2.3.1 (Décomposition de Dunford). — Soient V un k -espace vectoriel de dimension finie et $u \in \text{End}_k(V)$ un endomorphisme dont le polynôme caractéristique est scindé sur k . Il existe une unique paire (d, n) d'endomorphismes de V tels que

$$u = d + n$$

et

- (1) d est diagonalisable.
- (2) n est nilpotent.
- (3) $dn = nd$.

De plus, d et n sont des polynômes en u .

Démonstration. — Soient $\lambda_1, \dots, \lambda_r \in k$ les valeurs propres deux à deux distinctes de u . On rappelle que l'on a une décomposition en somme directe d'espaces caractéristiques $V = \bigoplus_i V_{(\lambda_i)}$. Pour tout $i = 1, \dots, r$ on note $\pi_{\lambda_i} : V \rightarrow V_{(\lambda_i)}$ la projection relativement à cette décomposition, qui est un polynôme en u . Puisque

$$\sum_i \pi_{\lambda_i} = \text{Id}_V$$

on peut écrire

$$u = \sum_i u \pi_{\lambda_i} = \underbrace{\sum_i \lambda_i \pi_{\lambda_i}}_d + \underbrace{\sum_i (u - \lambda_i \text{Id}_V) \pi_{\lambda_i}}_n.$$

Nous avons alors toutes les propriétés recherchées. Le fait que d et n commutent vient de ce que toutes les projections π_{λ_i} sont des polynômes en u , qui commutent tous entre eux. La nilpotence de n vient de ce que la restriction de $u - \lambda_i \text{Id}_V$ à $V_{(\lambda_i)}$ est un endomorphisme nilpotent (Exercice 2.2.5).

Pour montrer l'unicité nous nous appuyons sur la Proposition 2.1.3. Si $u = d' + n'$ est une autre décomposition vérifiant les propriétés (1), (2) et (3), les endomorphismes d' et n' commutent avec u , donc avec tout polynôme en u , donc avec d et n construits plus haut. Par la Proposition 2.1.3 les endomorphismes d et d' se diagonalisent dans une même base et donc l'endomorphisme $d - d' = n' - n$ est diagonalisable. Par ailleurs n et n' sont nilpotents et commutent ; donc, par la formule (1), l'endomorphisme $n' - n$ est nilpotent. On déduit par le Lemme 2.2.3 que l'endomorphisme $d - d' = n' - n$, qui est à la fois diagonalisable et nilpotent, est nécessairement nul. \square

⁽²⁾Le mot est à utiliser toujours avec précaution. Les lecteurs et lectrices pourront considérer la question suivante : en dimension 7, est-il plus facile de calculer la puissance 10^e pour un endomorphisme diagonalisable (mais non diagonal), ou pour un endomorphisme nilpotent ?

⁽³⁾La forme normale de Jordan du §2.4 apparaît déjà dans les travaux de Camille Jordan (1838-1922). La décomposition en partie diagonalisable et partie nilpotente est étudiée de façon approfondie par Claude Chevalley (1909-1984) dans son ouvrage classique *Théorie des groupes de Lie. Tome II. Groupes Algébriques*, Hermann, Paris, 1951. Nelson Dunford (1906-1986) généralise le concept aux espaces d'opérateurs linéaires et continus sur des espaces fonctionnels dans son ouvrage, classique aussi, co-écrit avec Jacob Schwartz (1930-2009), *Linear Operators. Parts I-III*, New York, John Wiley and Sons Inc., 1958-1971.

Exemples 2.3.2. — 1. Soient $\lambda \in k$ et $T_\lambda \in M_n(k)$ une matrice triangulaire supérieure dont tous les coefficients sur la diagonale sont égaux à λ . Alors la décomposition de Dunford de T_λ est

$$T_\lambda = D + N, \quad \text{avec } D = \lambda I_n \quad \text{et} \quad N = T_\lambda - \lambda I_n,$$

ou encore

$$\underbrace{\begin{pmatrix} \lambda & * & \cdots & \cdots & * \\ 0 & \lambda & * & & \vdots \\ \vdots & 0 & \ddots & * & \vdots \\ \vdots & & 0 & \lambda & * \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}}_{T_\lambda} = \underbrace{\begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 0 & & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & & 0 & \lambda & 0 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}}_D + \underbrace{\begin{pmatrix} 0 & * & \cdots & \cdots & * \\ 0 & 0 & * & & \vdots \\ \vdots & 0 & \ddots & * & \vdots \\ \vdots & & 0 & 0 & * \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}}_N.$$

Vérifions les conditions (1), (2) et (3) de la Proposition 2.3.1 : la matrice D est diagonale, la matrice N est nilpotente comme vu dans l'exemple 2.2.2.4, et enfin la matrice D commute avec N puisque D est un multiple de la matrice identité.

2. Soit T une matrice triangulaire par blocs T_{λ_i} comme ci-dessus :

$$(2) \quad T = \begin{pmatrix} T_{\lambda_1} & 0 & \cdots & 0 \\ 0 & T_{\lambda_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & T_{\lambda_r} \end{pmatrix}, \quad \text{avec } T_{\lambda_i} \in M_{m_i}(k).$$

Alors la décomposition de Dunford de T est

$$T = D + N, \quad \text{avec } D = \begin{pmatrix} \lambda_1 I_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{m_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r I_{m_r} \end{pmatrix} \quad \text{et } N = T - D.$$

La vérification est similaire à celle du point 1.

3. Soit $A \in M_n(k)$ une matrice dont le polynôme caractéristique est scindé. L'algorithme de trigonalisation 2.0 fournit une matrice de passage $P \in \text{GL}_n(k)$ telle que $T = P^{-1}AP$ est triangulaire supérieure par blocs T_{λ_i} comme au point 2. Si

$$T = D + N$$

est la décomposition de Dunford de T comme ci-dessus, alors on vérifie directement que

$$A = PTP^{-1} = (PDP^{-1}) + (PNP^{-1})$$

est la décomposition de Dunford de A .

Remarque 2.3.3. — Le point de vue adopté dans le point 3. ci-dessus permet de donner une démonstration alternative de l'existence de la décomposition de Dunford.

Exemple 2.3.4 (L'algorithme 2.0 appliqué à la décomposition de Dunford)

Voici comment fonctionne en pratique l'algorithme 2.0 pour calculer la décomposition de Dunford. Considérons la matrice

$$A = \begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} \in M_3(\mathbb{R})$$

vue dans l'exemple 1.8.3. Son polynôme caractéristique est $P_A(X) = -(X-1)^3$, scindé sur \mathbb{R} , la matrice est donc trigonalisable sur \mathbb{R} . Nous avons calculé une forme triangulaire

$$T = P^{-1}AP = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{avec } P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

La décomposition de Dunford de T est

$$T = D + N, \quad \text{avec } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } N = \begin{pmatrix} 0 & 3 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

La décomposition de Dunford de $A = PTP^{-1}$ est alors

$$\begin{pmatrix} -2 & 3 & 3 \\ -2 & 1 & 2 \\ -3 & 3 & 4 \end{pmatrix} = A = (PDP^{-1}) + (PNP^{-1}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -3 & 3 & 3 \\ -2 & 0 & 2 \\ -3 & 3 & 3 \end{pmatrix}.$$

Les lecteurs et lectrices pourront trouver utile de vérifier directement que la matrice

$$PNP^{-1} = \begin{pmatrix} -3 & 3 & 3 \\ -2 & 0 & 2 \\ -3 & 3 & 3 \end{pmatrix}$$

est bien nilpotente, d'indice de nilpotence égal à 3.

Exemple 2.3.5 (Décomposition de Dunford et trigonalisations)

Dans 2.3.2 la décomposition de Dunford se lit directement sur les matrices triangulaires de la forme (2), et d'ailleurs toute matrice trigonalisable peut être mise sous cette forme en vue du 2.3.4. Néanmoins, pour une matrice triangulaire supérieure A quelconque, il n'est *pas* vrai que sa décomposition de Dunford est $D + N$, où D est la matrice des coefficients diagonaux et $N = A - D$. Le point qui fait défaut est la commutation de D et N .

Les lectrices et lecteurs trouveront utile de détailler le calcul suivant. Considérons la matrice triangulaire supérieure

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}.$$

En posant naïvement

$$\tilde{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{et} \quad \tilde{N} = \begin{pmatrix} 0 & 3 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix},$$

l'on voit que $A = \tilde{D} + \tilde{N}$, \tilde{D} est diagonale, \tilde{N} est nilpotente, mais

$$\tilde{D}\tilde{N} = \begin{pmatrix} 0 & 3 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{pmatrix} = \tilde{N}\tilde{D}.$$

Cette décomposition n'est donc *pas* la décomposition de Dunford.

Pour déterminer la décomposition de Dunford nous appliquons l'algorithme 2.0. On calcule $\ker(A - I_3) = \text{Vect}(e_1)$, $\ker(A - I_3)^2 = \text{Vect}(e_1, e_2)$, $\ker(A - 2I_3) = \text{Vect}(14e_1 + 4e_2 +$

e_3), et nous obtenons comme forme triangulaire supérieure dans la base $\mathcal{B} = (e_1, e_2, 14e_1 + 4e_2 + e_3)$ la matrice

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

Celle-ci est de la forme (2) et sa décomposition de Dunford se lit directement

$$T = D + N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On a $T = P^{-1}AP$ avec

$$P = \begin{pmatrix} 1 & 0 & 14 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} 1 & 0 & -14 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}.$$

La décomposition de Dunford de $A = PTP^{-1}$ est alors

$$A = PDP^{-1} + PNP^{-1} = \begin{pmatrix} 1 & 0 & 14 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 3 & -12 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

L'on vérifie que l'on a bien la relation de commutation attendue

$$\begin{pmatrix} 1 & 0 & 14 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 & -12 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 & -12 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 & -12 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 14 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}$$

Remarque 2.3.6. — La décomposition de Dunford est utile pour calculer des puissances de matrices et, comme nous le verrons plus tard, des exponentielles de matrices. Si

$$A = (PDP^{-1}) + (PNP^{-1})$$

est la décomposition de Dunford d'une matrice $A \in M_n(k)$, avec P inversible, D diagonale, N nilpotente d'indice de nilpotence $\ell \leq n$ et $DN = ND$, alors la formule (1) donne

$$A^p = \sum_{i=0}^{\ell-1} \binom{p}{i} PD^{p-i}N^iP^{-1}$$

pour tout $p \geq \ell$. Cette formule est remarquable lorsque p est grand.

Remarque 2.3.7. — Il existe un algorithme effectif, issu des travaux de 1951 de Chevalley, pour calculer la décomposition de Dunford d'une matrice quelconque *sans passer* par le calcul des valeurs propres. Voir à ce sujet l'article de D. Couty, J. Esterle et R. Zarouf, « Décomposition effective de Jordan-Chevalley », *Gazette des Mathématiciens*, 129 (2011), 29–49. Les lecteurs et lectrices intéressés pourront le consulter à l'adresse http://smf4.emath.fr/Publications/Gazette/2011/129/smf_gazette_129_29-49.pdf.

2.4. Forme normale de Jordan

Soit V un k -espace vectoriel de dimension finie et $u \in \text{End}_k(V)$ un endomorphisme dont le polynôme caractéristique est scindé sur k . Sur chaque espace caractéristique $V_{(\lambda)}$ l'endomorphisme u est somme de l'homothétie de rapport λ et d'un endomorphisme nilpotent. La représentation matricielle d'une homothétie ne dépend pas du choix de base ; mais celle d'un nilpotent, oui. Il est donc souhaitable de savoir présenter les endomorphismes nilpotents de manière particulièrement simple.

Nous cherchons donc à fournir une écriture normalisée d'un endomorphisme u , où le choix des blocs standard est conventionnel et non intrinsèque. Nous allons montrer que, dans une base convenable, la matrice d'un endomorphisme nilpotent est diagonale par blocs du type U_n comme dans l'exemple 2.2.2. Par conséquent, on aura trouvé une base dans laquelle la matrice de u est diagonale par blocs du type $\lambda I_n + U_n$. Noter le fait que le même λ peut intervenir plusieurs fois, comme dans la matrice

$$(3) \quad \begin{pmatrix} 2 & 1 & 0 & & & \\ 0 & 2 & 1 & & & \\ 0 & 0 & 2 & & & \\ & & & 2 & 1 & \\ & & & 0 & 2 & \end{pmatrix}.$$

Définition 2.4.1. — Le bloc de Jordan de taille n et valeur propre λ est la matrice

$$J_n(\lambda) = \lambda I_n + U_n = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \lambda & 1 \\ 0 & \dots & \dots & \dots & \dots & \lambda \end{pmatrix}.$$

Nous avons en particulier $U_n = J_n(0)$.

Théorème 2.4.2 (Forme normale de Jordan). — Soient V un k -espace vectoriel de dimension finie et u un endomorphisme dont le polynôme caractéristique est scindé sur k .

– (EXISTENCE) Il existe une base de V dans laquelle la matrice de u est diagonale par blocs de Jordan.

– (UNICITÉ) Deux formes matricielles de u qui sont diagonales par blocs de Jordan coïncident à permutation des blocs de Jordan près. \square

Nous renvoyons à l'Appendice A pour la preuve. Une telle forme matricielle de u est dite « forme normale de Jordan ». Une base dans laquelle la matrice de u prend cette forme est dite « base de Jordan ». Il conviendra de noter qu'il n'y a pas unicité des bases de Jordan.

Exemple 2.4.3. — La matrice du (3) s'écrit sous forme diagonale par blocs de Jordan

$$\begin{pmatrix} J_3(2) & \\ & J_2(2) \end{pmatrix}.$$

Corollaire 2.4.4 (Blocs de Jordan et espaces propres). — Le nombre de blocs de Jordan de valeur propre λ est égal à la dimension de l'espace propre $V_\lambda = \ker(u - \lambda \text{Id}_V)$.

Démonstration. — Ceci découle directement de ce que la dimension de l'espace propre $\ker(J_n(\lambda) - \lambda I_n)$ est égale à 1 pour chaque bloc de Jordan. \square

Corollaire 2.4.5 (Blocs de Jordan et espaces caractéristiques)

(i) La somme des tailles des blocs de Jordan de valeur propre λ est égale à la dimension de l'espace caractéristique $V_{(\lambda)}$ (et donc à la multiplicité algébrique de λ).

(ii) La taille du plus grand bloc de Jordan de valeur propre λ est égale à l'indice de nilpotence de la restriction de $u - \lambda \text{Id}_V$ à l'espace caractéristique $V_{(\lambda)}$.

Démonstration. — Le point (i) découle du fait que l'affirmation est clairement vraie pour une forme de Jordan. Le point (ii) découle de ce que l'indice de nilpotence de la matrice $J_n(\lambda) - \lambda I_n$ est égal à n . \square

Corollaire 2.4.6 (Forme normale de Jordan et polynôme minimal)

Soit $u \in \text{End}_k(V)$ un endomorphisme dont le polynôme caractéristique est scindé sur k , avec racines deux à deux distinctes $\lambda_1, \dots, \lambda_r$. Le polynôme minimal de u est

$$(4) \quad \mu_u(X) = \prod_{i=1}^r (X - \lambda_i)^{d_i},$$

avec d_i égal à la taille maximale d'un bloc de Jordan de valeur propre λ_i .

Démonstration. — Le polynôme minimal a les mêmes racines que le polynôme caractéristique (Corollaire 1.6.5), il est donc de la forme (4) avec des exposants $d_i \geq 1$ à déterminer. Le polynôme minimal de u est le même que celui d'une forme normale de Jordan. Pour cette dernière, on voit directement qu'un polynôme de la forme (4) est annulateur si chaque d_i est au moins égal à la taille maximale d'un bloc de Jordan de valeur propre λ_i , et qu'il n'est pas annulateur dès que l'un des d_i est strictement plus petit que cette valeur. \square

Le lemme suivant contient le principe d'engendrement de la forme normale de Jordan.

Lemme 2.4.7. — Soient V un k -espace vectoriel de dimension finie, v un endomorphisme nilpotent et $x_0 \in V$ un vecteur non-nul. Soit s le plus petit entier positif tel que $v^s(x_0) = 0$. Posons

$$x_1 = v(x_0), x_2 = v(x_1), \dots, x_{s-1} = v(x_{s-2}).$$

Alors :

- Les vecteurs x_0, x_1, \dots, x_{s-1} forment une famille libre.
- Le sous-espace $E = \text{Vect}(x_0, x_1, \dots, x_{s-1})$ est stable par v .
- La matrice de $v|_E$ dans la base $\mathcal{B} = (x_{s-1}, x_{s-2}, \dots, x_0)$ est U_s .

Démonstration. — Les deux dernières affirmations sont claires (les lecteurs et lectrices devront néanmoins les vérifier!). Montrons la première affirmation. Soit $a_0 x_0 + a_1 x_1 + \dots + a_{s-1} x_{s-1} = 0$ une relation de liaison. Appliquant v^{s-1} on trouve $a_0 v^{s-1}(x_0) = 0$, donc $a_0 = 0$. Revenant à la première équation, et lui appliquant v^{s-2} , on a maintenant $a_1 = 0$. De proche en proche on établit la liberté. (Le lecteur soucieux de s'exercer peut formaliser la récurrence.) \square

Remarque 2.4.8. — La base \mathcal{B} du Lemme 2.4.7 est constituée des vecteurs x_0, x_1, \dots, x_{s-1} écrits dans l'ordre *inverse*. Si nous avons pris le parti de choisir pour blocs fondamentaux de Jordan des matrices triangulaires inférieures, nous n'aurions pas dû changer l'ordre de ces vecteurs. Ceci est une question de convention, pas de contenu mathématique.

L'appendice A contient la preuve du théorème 2.4.2. Nous invitons les lectrices et lecteurs à lire la Remarque A.1, qui en donne une esquisse. L'appendice B contient la description d'un algorithme général pour calculer la forme normale de Jordan. Pour des matrices de petite taille il n'est pourtant pas nécessaire de faire appel à la forme générale de l'algorithme : le Lemme 2.4.7 suffit, comme montré par les exemples 2.4.9 et D.1.

Exemple 2.4.9 (Matrices 2×2). — Soit $A \in M_2(k)$ une matrice dont le polynôme caractéristique est scindé sur k , avec racines λ_1, λ_2 . On note $V = k^2$.

• Si $\lambda_1 \neq \lambda_2$, la matrice est diagonalisable (Proposition 0.2.4). Une forme normale de Jordan est la même chose qu'une forme diagonale, ici composée de deux blocs de taille 1 :

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}.$$

• Si $\lambda_1 = \lambda_2 = \lambda$, deux cas se présentent.

— Si $\dim \ker(A - \lambda I_2) = 2$, la matrice est diagonalisable (Proposition 1.1.1). Une forme normale de Jordan est la même chose qu'une forme diagonale, ici composée de deux blocs de taille 1 :

$$\begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}.$$

— Si $\dim \ker(A - \lambda I_2) = 1$, la matrice n'est pas diagonalisable, mais elle est trigonalisable et possède une forme normale de Jordan avec un bloc de Jordan de taille 2 :

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Pour déterminer une base de Jordan, il suffit de déterminer un vecteur x dans $\ker(A - \lambda I_2)^2 \setminus \ker(A - \lambda I_2) = V \setminus \ker(A - \lambda I_2)$. Une base de Jordan est alors $\mathcal{B} = ((A - \lambda I_2)x, x)$. Dans la pratique, on peut déterminer un vecteur propre f et choisir un vecteur x qui complète f en une base.

Voici un calcul explicite. Considérons la matrice

$$A = \begin{pmatrix} -11 & 25 \\ -4 & 9 \end{pmatrix} \in M_2(\mathbb{R}).$$

Son polynôme caractéristique est $P_A(X) = (X + 1)^2$, avec une racine réelle double $\lambda = -1$. La matrice $A + I_2 = \begin{pmatrix} -10 & 25 \\ -4 & 10 \end{pmatrix}$ est de rang 1, l'espace propre $V_{-1} = \ker(A + I_2)$ est de dimension 1, engendré par $f = 5e_1 + 2e_2$. La matrice n'est donc pas diagonalisable et sa forme de Jordan est constituée du seul bloc $J_2(-1)$ de taille 2. Pour déterminer une base de Jordan nous complétons le vecteur f en une base (f, x) de \mathbb{R}^2 , ce qui équivaut à choisir $x \in \ker(A + I_2)^2 \setminus \ker(A + I_2) = \mathbb{R}^2 \setminus \ker(A + I_2)$. (Le lecteur attentif pourra trouver satisfaction en vérifiant que $(A + I_2)^2 = 0$.)

Prenons par exemple $x = e_2$. Une base de Jordan est alors

$$\mathcal{B} = ((A + I_2)x, x) = \left(\begin{pmatrix} 25 \\ 10 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).$$

Vérifions-le : en posant

$$P = \begin{pmatrix} 25 & 0 \\ 10 & 1 \end{pmatrix}, \quad \text{avec} \quad P^{-1} = \begin{pmatrix} \frac{1}{25} & 0 \\ -\frac{2}{5} & 1 \end{pmatrix},$$

nous avons bien

$$P^{-1}AP = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

2.4.10. Forme normale de Jordan et puissances de matrices. — La forme normale de Jordan est particulièrement adaptée au calcul des puissances de matrices tout simplement parce que les puissances des blocs nilpotents U_n sont très faciles à calculer :

$$U_n = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}, U_n^2 = \begin{pmatrix} 0 & 0 & 1 & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}, \dots, U_n^{n-1} = \begin{pmatrix} 0 & 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix},$$

et bien sûr $U_n^n = 0$.

Étant donné une matrice A dont le polynôme caractéristique est scindé, soit

$$T_{\text{Jordan}} = P^{-1}AP$$

une forme normale de Jordan pour A . Pour tout $p \geq 0$ nous avons alors

$$A^p = PT_{\text{Jordan}}^p P^{-1}.$$

Pour calculer la puissance p -ème de la forme normale de Jordan, il suffit de le faire pour chaque bloc de Jordan puisque les blocs de Jordan n'interagissent pas lors des multiplications successives. Puisque la décomposition de Dunford d'un bloc de Jordan se lit directement sur la matrice, i.e.

$$J_n(\lambda) = \lambda I_n + U_n,$$

nous obtenons avec la formule du binôme de Newton (pour $p \geq n - 1$)

$$J_n(\lambda)^p = \sum_{i=0}^{n-1} \binom{p}{i} \lambda^{p-i} U_n^i.$$

De façon explicite (pour $p \geq n - 1$)

$$(5) \quad J_n(\lambda)^p = \begin{pmatrix} \lambda^p & p\lambda^{p-1} & \frac{p(p-1)}{2}\lambda^{p-2} & \dots & \dots & \binom{p}{n-1}\lambda^{p-n+1} \\ 0 & \lambda^p & p\lambda^{p-1} & \frac{p(p-1)}{2}\lambda^{p-2} & \dots & \binom{p}{n-2}\lambda^{p-n+2} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \frac{p(p-1)}{2}\lambda^{p-2} \\ 0 & \dots & \dots & \dots & \lambda^p & p\lambda^{p-1} \\ 0 & \dots & \dots & \dots & \dots & \lambda^p \end{pmatrix}.$$

Exemple 2.4.11. — Considérons la matrice

$$A = \begin{pmatrix} -11 & 25 \\ -4 & 9 \end{pmatrix}$$

de l'exemple 2.4.9. Sa forme de Jordan est $J_2(-1) = -I_2 + U_2$. Avec P une matrice de changement de base telle que $J_2(-1) = P^{-1}AP$, on obtient pour tout $p \geq 1$:

$$A^p = P((-1)^p I_2 + p(-1)^{p-1} U_2) P^{-1} = P \begin{pmatrix} (-1)^p & p(-1)^{p-1} \\ 0 & (-1)^p \end{pmatrix} P^{-1}.$$

Nous renvoyons à l'exemple B.2 pour le calcul des puissances d'une matrice 4×4 .

2.4.12. Forme normale de Jordan et exponentielles de matrices. — La forme normale de Jordan est aussi bien adaptée au calcul d'exponentielles de matrices. Dans cette section le corps de base est \mathbb{R} ou \mathbb{C} . Nous le notons \mathbb{K} pour le mettre en évidence. Nous avons rassemblé dans l'Appendice E des rappels sur l'exponentielle des matrices.

Vu les propriétés (SCALAIRE), (COMMUTATION), (NILPOTENCE) et (DIAGONALE) de l'exponentielle (*loc. cit.*), pour un bloc de Jordan $J_n(\lambda) = \lambda I_n + U_n$ nous avons

$$\exp(J_n(\lambda)) = \exp(\lambda I_n) \exp(U_n) = e^\lambda \left(I_n + U_n + \frac{U_n^2}{2} + \cdots + \frac{U_n^{n-1}}{(n-1)!} \right),$$

ou encore

$$\exp(J_n(\lambda)) = e^\lambda \begin{pmatrix} 1 & 1 & \frac{1}{2} & \cdots & \cdots & \frac{1}{(n-1)!} \\ 0 & 1 & 1 & \frac{1}{2} & \cdots & \frac{1}{(n-2)!} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \frac{1}{2} \\ 0 & \cdots & \cdots & \cdots & 1 & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

Pour calculer l'exponentielle d'une matrice mise sous forme de Jordan l'on pourra combiner le calcul ci-dessus avec les propriétés (CONJUGAISON) et (DIAGONALE).

Exercice 2.4.13. — Pour $t \in \mathbb{K}$ quelconque montrer l'égalité

$$(6) \quad \exp(tJ_n(\lambda)) = e^{t\lambda} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \cdots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \frac{t^2}{2} & \cdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \frac{t^2}{2} \\ 0 & \cdots & \cdots & \cdots & 1 & t \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

Exemple 2.4.14. — Considérons la matrice

$$A = \begin{pmatrix} -11 & 25 \\ -4 & 9 \end{pmatrix}$$

de l'exemple 2.4.9, dont la forme de Jordan est $J_2(-1) = -I_2 + U_2$. Avec $P = \begin{pmatrix} 25 & 0 \\ 10 & 1 \end{pmatrix}$ la matrice de changement de base telle que $J_2(-1) = P^{-1}AP$, on obtient

$$\begin{aligned} \exp(tA) &= P \exp(tJ_2(-1)) P^{-1} \\ &= e^{-t} \begin{pmatrix} 25 & 0 \\ 10 & 1 \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{25} & 0 \\ -\frac{2}{5} & 1 \end{pmatrix} \\ &= e^{-t} \begin{pmatrix} 1 - 10t & 25t \\ -4t & 1 + 10t \end{pmatrix}. \end{aligned}$$

Dans l'exemple B.3 nous donnons un calcul d'exponentielle pour une matrice 4×4 .

2.5. Applications : équations différentielles linéaires et suites récurrentes

Nous donnons dans cette section deux applications importantes de la réduction des endomorphismes : résolution d'équations différentielles linéaires d'un côté (§2.5.1, §2.5.6) et résolution de récurrences linéaires de l'autre (§2.5.4, §2.5.13). Les deux contextes s'éclairent réciproquement et constituent deux facettes d'un même problème : la modélisation mathématique des processus d'évolution, « continus » dans le premier cas, « discrets » dans le deuxième. Les lecteurs pourront apprécier la manière dont la théorie de la réduction des endomorphismes éclaire à la fois l'oscillateur harmonique et les nombres de Fibonacci.

2.5.1. Équations différentielles linéaires. — Le corps de base est ici $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Dimension 1. En guise de motivation, commençons par le problème suivant :

Étant donné $\lambda \in \mathbb{K}$, déterminer les fonctions dérivables $f : \mathbb{R} \rightarrow \mathbb{K}$ qui vérifient l'équation

$$(7) \quad f' = \lambda f.$$

La réponse est la suivante : si f est une telle fonction, alors $t \mapsto e^{-\lambda t} f(t)$ est dérivable et $\frac{d}{dt} e^{-\lambda t} f(t) = -\lambda e^{-\lambda t} f(t) + e^{-\lambda t} f'(t) = 0$. Il existe donc une constante $c \in \mathbb{K}$ telle que $e^{-\lambda t} f(t) = c$, ou encore $f(t) = ce^{\lambda t}$. Réciproquement, toute fonction de cette forme est bien une solution de (7). En conclusion :

L'ensemble \mathcal{S} des solutions de l'équation (7) est un \mathbb{K} -espace vectoriel de dimension 1, engendré par la fonction $t \mapsto e^{\lambda t}$.

On dit que (7) est une *équation différentielle linéaire d'ordre 1* : « équation différentielle d'ordre 1 » puisqu'elle porte sur une fonction en une variable et sa dérivée d'ordre 1, « linéaire » puisque l'ensemble de ses solutions est un espace vectoriel par rapport aux opérations de multiplication par des scalaires et d'addition des fonctions.

Dimension $n \geq 1$. Le problème analogue en dimension supérieure est le suivant :

Étant donné une matrice $A \in M_n(\mathbb{K})$, déterminer les fonctions dérivables $x : \mathbb{R} \rightarrow \mathbb{K}^n$ qui vérifient l'équation différentielle linéaire d'ordre 1

$$(8) \quad x' = Ax.$$

Vérifions que l'adjectif « linéaire » est justifié : l'ensemble \mathcal{S} des solutions de (8) est bien un \mathbb{K} -espace vectoriel : si x et y sont solutions et $\lambda \in \mathbb{K}$, alors l'application $\lambda x + y : \mathbb{R} \rightarrow \mathbb{K}^n$ définie par $(\lambda x + y)(t) = \lambda x(t) + y(t)$ est encore une solution.

Remarque 2.5.2 (Terminologie et interprétation physique)

On conçoit la variable t comme modèle mathématique du temps et l'on conçoit x comme modèle mathématique de l'évolution temporelle d'un système physique à n paramètres réels ($\mathbb{K} = \mathbb{R}$) ou complexes ($\mathbb{K} = \mathbb{C}$). L'équation (8) exprime la vitesse d'une courbe tracée dans \mathbb{K}^n comme fonction de la position. Les fonctions $x : \mathbb{R} \rightarrow \mathbb{K}^n$ qui vérifient (8) sont aussi appelées *solutions globales* puisque leur domaine de définition est \mathbb{R} .

Théorème 2.5.3. — *L'espace vectoriel \mathcal{S} des solutions de l'équation différentielle (8) est constitué des fonctions $\mathbb{R} \rightarrow \mathbb{K}^n$ de la forme*

$$t \mapsto \exp(tA)x_0,$$

avec $x_0 \in \mathbb{K}^n$ quelconque.

L'application $\mathcal{S} \rightarrow \mathbb{K}^n$, $x \mapsto x(0)$ est un isomorphisme d'espace vectoriels. L'espace \mathcal{S} est de dimension n , engendré par les fonctions $t \mapsto \exp(tA)e_i$, avec e_1, \dots, e_n base de \mathbb{K}^n .

Un corollaire direct (v. Lemme E.1) est que toutes les solutions de (8) sont de classe C^∞ .

Démonstration. — La preuve suit de près l'argument donné en dimension 1, en faisant usage de l'exponentielle matricielle. Si x est une solution de l'équation $x' = Ax$, alors $t \mapsto \exp(-tA)x(t)$ est dérivable (v. Lemme E.1) et

$$\begin{aligned} \frac{d}{dt} \exp(-tA)x(t) &= \exp(-tA)' x(t) + \exp(-tA) x'(t) \\ &= \exp(-tA) (-A)x(t) + \exp(-tA) Ax(t) \\ &= 0. \end{aligned}$$

Il existe donc une constante $x_0 \in \mathbb{K}^n$ telle que $\exp(-tA)x(t) = x_0$, ou encore $x(t) = \exp(tA)x_0$. Réciproquement, toute fonction $x : \mathbb{R} \rightarrow \mathbb{K}^n$ de la forme $x(t) = \exp(tA)x_0$ vérifie $x'(t) = A \exp(tA)x_0 = Ax(t)$. C'est donc une solution de (8).

Remarquons le fait que $x(0) = x_0$. Cela entraîne l'injectivité de l'application $\mathcal{S} \rightarrow \mathbb{K}^n$, $x \mapsto x(0)$. La surjectivité découle de ce qui précède, et la linéarité est évidente. \square

Résolution d'équations différentielles linéaires via la forme normale de Jordan ou via la décomposition de Dunford. Soit $A \in M_n(\mathbb{K})$ une matrice dont le polynôme caractéristique est scindé sur \mathbb{K} . Pour résoudre $x' = Ax$ on pourra trouver une forme de Jordan $T = P^{-1}AP$, ou bien simplement une matrice triangulaire de Dunford $T = P^{-1}AP$ diagonale par blocs qui possèdent une unique valeur propre (v. Exemple 2.3.2). L'unique solution de l'équation $x' = Ax$ telle que $x(0) = x_0$ est alors $x(t) = \exp(tA)x_0$, ou encore

$$x(t) = P \exp(tT) P^{-1} x_0.$$

Les propriétés calculatoires remarquables de telles matrices T trouvent ici bon usage.

2.5.4. Récurrences linéaires d'ordre 1. — Nous travaillons ici sur un corps de base k quelconque. Le cadre considéré dans cette section est l'analogie « discret » du cadre « continu » considéré dans la section précédente.

Dimension 1. Considérons le problème suivant :

Étant donné $\lambda \in k$, déterminer les suites $(x_p)_{p \geq 0}$ d'éléments de k qui vérifient la relation

$$(9) \quad x_{p+1} = \lambda x_p \quad \text{pour tout } p \geq 0.$$

On montre par récurrence que $x_p = \lambda^p x_0$ avec $x_0 \in k$ quelconque et l'on conclut :

L'ensemble \mathcal{S} des solutions de l'équation (9) est un k -espace vectoriel de dimension 1, engendré par la suite $(1, \lambda, \lambda^2, \dots, \lambda^p, \dots)$.

On dit que (9) est une *relation de récurrence linéaire d'ordre 1* : « relation de récurrence d'ordre 1 » puisqu'elle porte sur tous les termes de la suite et leurs successeurs d'ordre 1, « linéaire » puisque l'ensemble de ses solutions est un espace vectoriel par rapport aux opérations de multiplication par des scalaires et d'addition des suites. Précisons ces dernières : l'addition des suites et la multiplication par des scalaires s'effectue terme à terme.

Dimension $n \geq 1$. Le problème analogue en dimension supérieure est le suivant :

Étant donné une matrice $A \in M_n(\mathbb{K})$, déterminer les suites $(x_p)_{p \geq 0}$ d'éléments de k^n qui vérifient la relation de récurrence linéaire d'ordre 1

$$(10) \quad x_{p+1} = Ax_p \quad \text{pour tout } p \geq 0.$$

L'adjectif « linéaire » est justifié : l'ensemble des solutions de (10) est un k -espace vectoriel par rapport aux opérations des multiplication par des scalaires et d'addition des suites.

L'analogie discret du Théorème 2.5.3 est le résultat suivant. La démonstration ne présente pas de difficulté et est laissée à la charge du lecteur.

Proposition 2.5.5. — *L'espace vectoriel \mathcal{S} des solutions de la relation de récurrence (8) est constitué des suites de la forme*

$$x_p = A^p x_0 \quad \text{pour tout } p \geq 0,$$

avec $x_0 \in \mathbb{K}^n$ quelconque.

L'application $\mathcal{S} \rightarrow \mathbb{K}^n$, $(x_p)_{p \geq 0} \mapsto x_0$ est un isomorphisme d'espaces vectoriels. L'espace \mathcal{S} est de dimension n , engendré par les suites $(A^p e_i)_{p \geq 0}$, avec e_1, \dots, e_n base de \mathbb{K}^n . \square

Calculs explicites via la forme normale de Jordan ou via la décomposition de Dunford. Il devrait maintenant être clair que les techniques de calcul de puissances de matrices développées dans les sections précédentes s'appliquent pour calculer les suites définies par des relations de récurrence linéaires d'ordre 1. Si le polynôme caractéristique de A est scindé sur k on pourra considérer une forme de Jordan $T = P^{-1}AP$ ou une matrice triangulaire de Dunford $T = P^{-1}AP$ diagonale par blocs qui possèdent une unique valeur propre. Nous aurons alors pour tout $p \geq 0$:

$$x_p = A^p x_0 = PT^p P^{-1} x_0.$$

2.5.6. Équations différentielles linéaires d'ordre supérieur (†). — ⁽⁴⁾ Nous revenons dans cette section au corps de base $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

L'équation différentielle linéaire (7) a été généralisée au §2.5.1 en augmentant la dimension de l'espace dans lequel les solutions prennent leurs valeurs. Nous donnons ici une autre généralisation importante, où l'on continue à regarder des fonctions à valeurs scalaires mais on augmente le nombre de dérivées. Le problème qui en résulte est le suivant :

Étant donné $n \geq 1$ et $a_0, a_1, \dots, a_{n-1} \in \mathbb{K}$, déterminer les fonctions $f : \mathbb{R} \rightarrow \mathbb{K}$ qui sont n fois dérivables et vérifient l'équation

$$(11) \quad f^{(n)} + a_{n-1}f^{(n-1)} + a_{n-2}f^{(n-2)} + \dots + a_1 f' + a_0 f = 0.$$

On dit que (11) est une « équation différentielle linéaire d'ordre n » : « linéaire » puisque l'ensemble de ses solutions est un \mathbb{K} -espace vectoriel (vérification directe), « d'ordre n » puisque l'équation implique les dérivées de f jusqu'à l'ordre n .

Proposition 2.5.7 (Réduction à une équation d'ordre 1)

Soit \mathcal{S} l'espace des solutions de (11). L'application linéaire

$$\mathcal{S} \rightarrow \mathbb{K}^n, \quad f \mapsto \begin{pmatrix} f(0) \\ f'(0) \\ \vdots \\ f^{(n-1)}(0) \end{pmatrix}$$

⁽⁴⁾Cette section n'a pas été traitée en cours et n'est par conséquent pas exigible.

est un isomorphisme. En particulier, \mathcal{S} est un \mathbb{K} -espace vectoriel de dimension n .

Toutes les solutions de (11) sont de classe C^∞ .

Démonstration. — L'équation (11) équivaut à l'équation différentielle linéaire d'ordre 1

$$(12) \quad y' = Ay, \quad A = \begin{pmatrix} 0 & 1 & & \dots & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

en posant

$$y(t) = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix}, \text{ et réciproquement } f = \text{ première composante de } y.$$

La conclusion découle alors du Théorème 2.5.3 appliqué à l'équation (12). \square

Remarque 2.5.8 (Espace des phases). — La proposition précédente est une manifestation d'un principe puissant : on peut toujours reformuler un problème différentiel d'ordre $n \geq 1$ comme un problème différentiel d'ordre 1 en considérant les dérivées d'ordre $< n$ comme des variables supplémentaires. Lorsque $n = 2$, on se ramène à travailler dans l'« espace des phases », où les coordonnées sont la position et la quantité de mouvement. Celui-ci est sous-jacent à la formulation hamiltonienne des lois de la mécanique classique.⁽⁵⁾ Lorsque $n > 2$, on travaille dans ce qui s'appelle l'« espace des jets d'ordre $n - 1$ ».

La Proposition 2.5.7 permet de résoudre (11) en résolvant l'équation d'ordre 1 associée et en considérant la première composante des solutions trouvées. Le résultat est le Théorème 2.5.9 ci-dessous. C'est un exercice intéressant que les lecteurs pourront envisager. La démonstration du Théorème 2.5.9 que nous présentons plus bas est différente.⁽⁶⁾ Elle s'appuie sur une application directe du Lemme des noyaux 1.4.6.

Théorème 2.5.9. — *Supposons que le polynôme*

$$P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

est scindé sur \mathbb{K} . Soient $\lambda_1, \dots, \lambda_r$ ses racines, de multiplicité m_1, \dots, m_r , de sorte que

$$P(X) = \prod_{i=1}^r (X - \lambda_i)^{m_i}.$$

L'espace \mathcal{S} des solutions de l'équation différentielle (11) est constitué des fonctions

$$f(t) = \sum_{i=1}^r e^{\lambda_i t} Q_i(t),$$

où les $Q_i \in \mathbb{K}[X]$ sont des polynômes de degré $\leq m_i - 1$. En particulier, une base de \mathcal{S} est constituée des fonctions

$$\{t \mapsto t^j e^{\lambda_i t} : i = 1, \dots, r, 0 \leq j \leq m_i - 1\}.$$

⁽⁵⁾Voici deux références exceptionnelles sur le sujet : V. I. Arnol'd, *Mathematical Methods of Classical Mechanics*, GTM 60, Springer, 1997, et R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics*, vol. 1, Basic Books, 2011.

⁽⁶⁾Celle-ci est tirée de X. Gourdon, *Les maths en tête : Analyse*, Ellipses, 2008, §6.2.2.

La preuve du théorème repose sur la réinterprétation suivante – surprenante ! – de l'espace \mathcal{S} . Considérons l'opérateur de différentiation

$$D : C^\infty(\mathbb{R}, \mathbb{K}) \rightarrow C^\infty(\mathbb{R}, \mathbb{K}), \quad f \mapsto f'.$$

Ceci est un endomorphisme linéaire du \mathbb{K} -espace vectoriel $C^\infty(\mathbb{R}, \mathbb{K})$. Avec $P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$ nous avons alors

$$P(D)f = f^{(n)} + a_{n-1}f^{(n-1)} + a_{n-2}f^{(n-2)} + \dots + a_1f' + a_0f,$$

et donc

$$\mathcal{S} = \ker P(D).$$

(Ici nous utilisons le fait que toute solution est de classe C^∞ , tel que donné par la Proposition 2.5.7.)

Démonstration du Théorème 2.5.9. — Par le Lemme des noyaux 1.4.6 nous avons⁽⁷⁾

$$\ker P(D) = \bigoplus_{i=1}^r \ker(D - \lambda_i \text{Id})^{m_i}.$$

L'on montre ensuite par récurrence l'affirmation suivante : pour tout $\lambda \in \mathbb{K}$ et $m \geq 1$,

$$\ker(D - \lambda \text{Id})^m = \{e^{\lambda t}Q(t) : Q \in \mathbb{K}[X] \text{ de degré } \leq m-1\}.$$

L'initialisation de la récurrence n'est rien d'autre que la résolution de l'équation (7). L'hérédité ne pose pas de problème et nous la laissons à la charge des lectrices intéressées. \square

Remarque 2.5.10. — La matrice A de l'équation (12) s'appelle *la matrice compagnon du polynôme* $P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$. Son polynôme caractéristique est égal à $(-1)^n P(X)$. Les matrices compagnons jouent un rôle important dans l'étude de réductions raffinées des endomorphismes linéaires. Par conséquent, le polynôme P est appelé *polynôme caractéristique de l'équation différentielle* (11).

Exercice 2.5.11. — Écrire l'espace des solutions de l'équation différentielle (8) comme noyau d'un opérateur linéaire approprié.

Exemple 2.5.12 (L'équation $f'' + \lambda f = 0$ avec $\lambda \in \mathbb{R}$). — Les solutions de l'équation $f'' - \lambda f = 0$ avec $\lambda > 0$ sont

$$f(t) = c_1 e^{\sqrt{\lambda}t} + c_2 e^{-\sqrt{\lambda}t} \quad \text{avec } c_1, c_2 \in \mathbb{R}.$$

Les solutions de l'équation $f'' = 0$ sont

$$f(t) = c_1 + c_2 t \quad \text{avec } c_1, c_2 \in \mathbb{R}.$$

Les solutions à valeurs *complexes* de l'équation $f'' + \lambda f = 0$ avec $\lambda > 0$ sont

$$f(t) = c_1 e^{i\sqrt{\lambda}t} + c_2 e^{-i\sqrt{\lambda}t} \quad \text{avec } c_1, c_2 \in \mathbb{C}.$$

Les solutions à valeurs *réelles* s'obtiennent en prenant la partie réelle de cette expression :

$$g(t) = a_1 \cos \sqrt{\lambda}t + a_2 \sin \sqrt{\lambda}t \quad \text{avec } a_1, a_2 \in \mathbb{R}.$$

⁽⁷⁾Le Lemme des noyaux est bien valable dans des espaces vectoriels de dimension infinie !

2.5.13. Récurrences linéaires d'ordre supérieur (†). — ⁽⁸⁾ Nous revenons dans cette section à un corps de base quelconque k .

La récurrence linéaire à valeurs scalaires (9) admet aussi comme généralisation l'étude des récurrences d'ordre supérieur. Le problème qui en résulte est le suivant :

Étant donnés $n \geq 1$ et $a_0, a_1, \dots, a_{n-1} \in \mathbb{K}$, déterminer les suites $(x_p)_{p \geq 0}$ avec $x_p \in k$ qui vérifient la relation de récurrence

$$(13) \quad x_{p+n} + a_{n-1}x_{p+n-1} + a_{n-2}x_{p+n-2} + \dots + a_1x_{p+1} + a_0x_0 = 0.$$

On dit que (13) est une « récurrence linéaire d'ordre n » : « linéaire » puisque l'espace \mathcal{S} des solutions est un \mathbb{K} -espace vectoriel (vérification directe), « d'ordre n » puisqu'elle porte sur tous les termes de la suite ainsi que leurs successeurs jusqu'à l'ordre n . Nous avons par ailleurs clairement $\dim \mathcal{S} = n$ puisque toute suite qui vérifie (13) est uniquement déterminée par ses n premiers termes.

Exercice 2.5.14. — Montrer que la récurrence (13) équivaut à la récurrence linéaire d'ordre 1

$$y_{p+1} = Ay_p, \quad \text{avec } A = \begin{pmatrix} 0 & 1 & & \dots & 0 \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}, \quad y_p \in k^n, \quad p \geq 0,$$

en posant

$$y_p = \begin{pmatrix} x_p \\ x_{p+1} \\ \vdots \\ x_{p+n-1} \end{pmatrix}, \quad \text{et réciproquement } x_p = \text{première composante de } y_p.$$

□

Soit S l'espace des suites à valeurs dans k . En analogie avec la section précédente, pour étudier l'espace \mathcal{S} des solutions de (13) nous introduisons l'endomorphisme de décalage

$$D : S \rightarrow S, \quad \mathbf{x} = (x_0, x_1, x_2, \dots) \mapsto D\mathbf{x} = (x_1, x_2, \dots),$$

ou encore

$$(D\mathbf{x})_p = x_{p+1} \quad \text{pour } \mathbf{x} = (x_p)_{p \geq 0}.$$

Alors $(D^i\mathbf{x})_p = x_{p+i}$ pour tout $i \geq 1$. En posant

$$P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

nous obtenons

$$(P(D)\mathbf{x})_p = x_{p+n} + a_{n-1}x_{p+n-1} + a_{n-2}x_{p+n-2} + \dots + a_1x_{p+1} + a_0x_0$$

et donc

$$\mathcal{S} = \ker P(D).$$

⁽⁸⁾Cette section n'a pas été traitée en cours et n'est par conséquent pas exigible.

Théorème 2.5.15. — Supposons que le polynôme

$$P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$$

est scindé sur \mathbb{K} . Soient $\lambda_1, \dots, \lambda_r$ ses racines, de multiplicité m_1, \dots, m_r , de sorte que

$$P(X) = \prod_{i=1}^r (X - \lambda_i)^{m_i}.$$

L'espace \mathcal{S} des suites qui vérifient (13) admet une base formée des suites suivantes :

$$\left\{ (\lambda_i^p)_{p \geq 0}, (p\lambda_i^p)_{p \geq 0}, \dots, \left(\binom{p}{m_i - 1} \lambda_i^{p-m_i+1} \right)_{p \geq 0} : i = 1, \dots, r \right\}.$$

(Par convention $\binom{p}{m_i - 1} = 0$ pour $p < m_i - 1$.)

Démonstration. — Par le Lemme des noyaux 1.4.6 nous avons

$$\ker P(D) = \bigoplus_{i=1}^r \ker (D - \lambda_i \text{Id})^{m_i}.$$

L'on montre ensuite par récurrence l'affirmation suivante : pour tout $\lambda \in \mathbb{K}$ et $m \geq 1$,

$$\ker (D - \lambda \text{Id})^m = \text{Vect} \left\{ (\lambda_i^p)_{p \geq 0}, (p\lambda_i^p)_{p \geq 0}, \dots, \left(\binom{p}{m-1} \lambda_i^{p-m+1} \right)_{p \geq 0} \right\}.$$

L'initialisation de la récurrence n'est rien d'autre que la résolution de la récurrence linéaire (9). L'hérédité ne pose pas de problème particulier et nous la laissons en exercice. \square

Le polynôme $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ est appelé *polynôme caractéristique* de la relation de récurrence (13).

Exercice 2.5.16. — Écrire l'espace des solutions de la récurrence linéaire (10) comme noyau d'un opérateur linéaire approprié.

Exemple 2.5.17 (Suite de Fibonacci). — Considérons la *suite de Fibonacci* $(x_p)_{p \geq 0}$ définie par la récurrence linéaire d'ordre 2 suivante :

$$x_{p+2} = x_{p+1} + x_p, \quad x_0 = 0, \quad x_1 = 1.$$

Le polynôme caractéristique est ici

$$P(X) = X^2 - X - 1,$$

et ses racines sont $(1 \pm \sqrt{5})/2$. On obtient alors

$$x_p = c_1 \left(\frac{1 + \sqrt{5}}{2} \right)^p + c_2 \left(\frac{1 - \sqrt{5}}{2} \right)^p$$

pour tout $p \geq 0$, avec $c_1, c_2 \in \mathbb{R}$ des constantes uniquement déterminées par les deux premiers termes de la suite, ici $x_0 = 0$ et $x_1 = 1$. La relation $x_0 = 0$ implique $c_2 = -c_1$ et la relation $x_1 = 1$ implique $1 = c_1(1 + \sqrt{5})/2 - c_1(1 - \sqrt{5})/2 = c_1\sqrt{5}$, ou encore $c_1 = 1/\sqrt{5}$. Par conséquent

$$x_p = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^p - \left(\frac{1 - \sqrt{5}}{2} \right)^p \right)$$

pour tout $p \geq 0$. Il s'agit là d'une formule fameuse pour les nombres de Fibonacci. (Pouvez-vous montrer directement que les x_p définis par cette formule sont des entiers ?)

Remarque 2.5.18. — Les lectrices et lecteurs attentifs auront remarqué le fait que la base fournie par le Théorème 2.5.15 pour les récurrences linéaires d'ordre n est constituée de suites dont les termes se lisent sur la première ligne des puissances successives des blocs de Jordan de taille m_i , $i = 1, \dots, r$, v. (5). Ceci est en lien avec le fait que la forme normale de Jordan de la matrice compagnon du polynôme $\prod_{i=1}^r (X - \lambda_i)^{m_i}$ possède un unique bloc de Jordan pour chaque valeur propre λ_i , de taille maximale m_i (exercice!).

Dans la même veine, la base fournie par le Théorème 2.5.9 pour les équations différentielles linéaires d'ordre n est formée des fonctions qui constituent la première ligne de l'exponentielle des blocs de Jordan de taille m_i , $i = 1, \dots, r$, v. (6).

————— o —————

Ceci est la fin du cours. Quelle est la suite? Bien évidemment, l'auteur de ce texte vous encourage à suivre un cours d'équations différentielles. Ceci vous permettra d'approfondir et apprécier le rôle joué par les valeurs propres, qui gouvernent l'allure globale des solutions pour des équations linéaires, ou encore leur comportement local pour des équations non-linéaires. D'ici là, on pourra bien se permettre une petite dose d'émerveillement. Nous aurons vu dans ce cours l'algèbre des polynômes à l'œuvre pour comprendre la structure interne des endomorphismes, la réduction des endomorphismes appliquée à des problèmes d'évolution, ou encore un parallèle (surprenant, peut-être?) entre aspects « discrets » et aspects « continus ». Les mathématiques jouissent d'une remarquable unité interne.



A. Appendice (†) : démonstration du théorème de Jordan

[Les appendices non discutés en cours ne seront pas exigibles. Leur lecture devrait néanmoins faire passer un temps agréable aux lectrices et lecteurs.]

Remarque A.1 (Survol de la preuve du théorème 2.4.2)

Pour comprendre la preuve de l'existence et de l'unicité de la forme normale de Jordan qui suit, il est utile d'assimiler la vue d'ensemble suivante.

L'existence sera démontrée en modifiant légèrement l'algorithme 2.0 de trigonalisation : dans celui-ci, nous construisons une base de trigonalisation en choisissant successivement des supplémentaires quelconques de $\ker(u - \lambda \text{Id}_V)^{i-1}$ dans $\ker(u - \lambda \text{Id}_V)^i$ lorsque i croissait de 1 à p , l'indice de nilpotence de $(u - \lambda \text{Id}_V)|_{V_{(\lambda)}}$; pour obtenir la forme normale de Jordan, nous allons choisir des supplémentaires *adaptés* de $\ker(u - \lambda \text{Id}_V)^{i-1}$ dans $\ker(u - \lambda \text{Id}_V)^i$ lorsque i décroît de p à 1.

L'unicité découlera de ce que, pour un endomorphisme nilpotent u , le nombre de blocs de Jordan de taille donnée se lit sur la suite des noyaux de u , qui, elle, est intrinsèque et indépendante de tout choix de base.

Démonstration du théorème 2.4.2. — Nous démontrons d'abord l'existence de la forme normale de Jordan, et ensuite son unicité.

EXISTENCE. En travaillant avec la restriction de $u - \lambda \text{Id}_V$ à chaque espace caractéristique $V_{(\lambda)}$, nous pouvons supposer sans perte de généralité que u est nilpotent.

Soit p l'indice de nilpotence de u et

$$\{0\} = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_p = V$$

la suite des noyaux de u , où $N_i = \ker u^i$ pour $i = 0, \dots, p$. Nous construisons par récurrence *descendante* sur $i = p, \dots, 0$ des supplémentaires M_i de N_{i-1} dans N_i tels que

$$u(M_i) \subset M_{i-1}.$$

Initialisation de la récurrence pour $i = p$: on prend pour M_p un supplémentaire quelconque de N_{p-1} dans $N_p = V$.

Hérédité : si M_i est construit pour $i \geq 2$, nous prenons pour M_{i-1} un supplémentaire quelconque de N_{i-2} dans N_{i-1} qui contient $u(M_i)$. Noter que cette condition a un sens puisque d'une part $u(M_i)$ est contenu dans N_{i-1} car $u(M_i) \subset u(N_i) \subset N_{i-1}$, et d'autre part $u(M_i) \cap N_{i-2} = \{0\}$ puisque $M_i \cap N_{i-1} = \{0\}$ et $u^{-1}(N_{i-2}) = N_{i-1}$. Remarquons par ailleurs que la restriction de u à M_i est injective pour $i \geq 2$ puisque $M_i \cap N_1 = \{0\}$.

Avec ces choix, l'endomorphisme u agira de la manière suivante :

$$\begin{array}{ccccc} N_i & = & N_{i-1} & \oplus & M_i \\ u \downarrow & & u \downarrow & & u \downarrow \\ N_{i-1} & = & N_{i-2} & \oplus & M_{i-1}. \end{array}$$

On construit alors une base de Jordan de V de la manière suivante. On renvoie au tableau ci-dessous pour une représentation imagée de la construction. Posons $m_i = \dim M_i$. On choisit une base $(e_{p,1}, \dots, e_{p,m_p})$ de M_p et on applique des puissances successives de u aux éléments de cette base, à la manière du Lemme 2.4.7. On choisit ensuite une base $(e_{p-1,m_p+1}, \dots, e_{p-1,m_{p-1}})$ d'un supplémentaire de $u(M_p) = \text{Vect}(u(e_{p,1}), \dots, u(e_{p,m_p}))$ dans

M_{p-1} , et on applique des puissances successives de u aux éléments de cette base. On continue de la même manière, en choisissant une base $(e_{i,m_{i+1}+1}, \dots, e_{i,m_i})$ d'un supplémentaire de $u(M_{i+1})$ dans M_i et en appliquant des puissances successives de u à ses éléments.

$$\begin{array}{ccccccc}
N_1 & & N_2 & & \dots & & N_{p-1} & & N_p \\
\cup & & \cup & & & & \cup & & \cup \\
M_1 & \xleftarrow{u} & M_2 & \xleftarrow{u} & \dots & \xleftarrow{u} & M_{p-1} & \xleftarrow{u} & M_p \\
\\
u^{p-1}(e_{p,1}) & \leftarrow \dots \leftarrow & u^{p-2}(e_{p,1}) & \leftarrow \dots \leftarrow & \dots & \leftarrow \dots \leftarrow & u(e_{p,1}) & \leftarrow \dots \leftarrow & e_{p,1} \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
u^{p-1}(e_{p,m_p}) & \leftarrow \dots \leftarrow & u^{p-2}(e_{p,m_p}) & \leftarrow \dots \leftarrow & \dots & \leftarrow \dots \leftarrow & u(e_{p,m_p}) & \leftarrow \dots \leftarrow & e_{p,m_p} \\
\\
u^{p-2}(e_{p-1,m_{p+1}}) & \leftarrow \dots \leftarrow & u^{p-3}(e_{p-1,m_{p+1}}) & \leftarrow \dots \leftarrow & \dots & \leftarrow \dots \leftarrow & e_{p-1,m_{p+1}} & & \\
\vdots & & \vdots & & & & & & \\
u^{p-2}(e_{p-1,m_{p-1}}) & \leftarrow \dots \leftarrow & u^{p-3}(e_{p-1,m_{p-1}}) & \leftarrow \dots \leftarrow & \dots & \leftarrow \dots \leftarrow & e_{p-1,m_{p-1}} & & \\
\vdots & & \vdots & & & & & & \\
e_{1,m_{2+1}} & & & & & & & & \\
\vdots & & & & & & & & \\
\vdots & & & & & & & & \\
e_{1,m_1} & & & & & & & &
\end{array}$$

Par le Lemme 2.4.7, les lignes de ce tableau sont constituées de vecteurs linéairement indépendants. On peut montrer sans difficulté que les vecteurs qui constituent tout le tableau forment une base. On ordonne cette base en lisant successivement les lignes de haut en bas et de gauche à droite. Toujours en vertu du Lemme 2.4.7, la matrice de u dans cette base est diagonale par blocs de type U_n , avec m_p blocs de type U_p , puis $m_{p-1} - m_p$ blocs de type U_{p-1} , $m_{p-2} - m_{p-1}$ blocs de type U_{p-2} et ainsi de suite jusqu'à $m_1 - m_2$ blocs de type U_1 (c'est-à-dire la matrice carrée d'ordre 1 nulle).

Unicité. On se restreint sans perte de généralité au cas d'un endomorphisme u nilpotent. Considérons une forme de Jordan pour u et notons m_s le nombre de blocs de Jordan $U_s = J_s(0)$ pour $1 \leq s \leq p$, avec p l'indice de nilpotence de u . Il s'agit de montrer que les entiers m_s sont uniquement déterminés par u . À cet effet, on considère la suite des noyaux de u comme ci-dessus, avec $N_i = \ker u^i$ pour $i = 1, \dots, p$. L'on voit alors directement que

$$\dim N_i = \sum_{s=1}^i s m_s + i \sum_{s=i+1}^p m_s, \quad i = 1, \dots, p.$$

Interprétons ceci comme un système linéaire de p équations en p indéterminées m_1, \dots, m_p :

$$\begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ 1 & 2 & \dots & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & & & \ddots & \\ 1 & 2 & 3 & \dots & p \end{pmatrix} \cdot \begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_p \end{pmatrix} = \begin{pmatrix} \dim N_1 \\ \dim N_2 \\ \dim N_3 \\ \vdots \\ \dim N_p \end{pmatrix}.$$

La matrice du système est inversible, de déterminant 1. En effet, nous pouvons la mettre sous forme triangulaire supérieure avec des 1 sur la diagonale en soustrayant d'abord la 1^{re} ligne des lignes 2 à p , en soustrayant ensuite la 2^e ligne des lignes 3 à p , et en continuant ainsi jusqu'à soustraire la ligne $p - 1$ de la ligne p . Par conséquent, le système admet une unique solution. \square

B. Appendice (†) : algorithme pour calculer une forme normale de Jordan

La démonstration du théorème 2.4.2 est effective si l'on connaît les valeurs propres de l'endomorphisme $u \in \text{End}_k(V)$. L'algorithme qui en résulte est un raffinement de l'algorithme 2.0. Soient λ une valeur propre de u et $N_j = \ker(u - \lambda \text{Id}_V)^j$, $j \geq 1$ la suite des noyaux. Pour décrire l'algorithme nous introduisons les notations suivantes.

1. Soit \mathcal{B} une famille libre de vecteurs dans $V_{(\lambda)}$. On note \mathcal{B}^j les vecteurs de \mathcal{B} qui sont contenus dans N_j mais pas dans N_{j-1} .

2. Soit $\mathcal{C} = (x_0, y_0, \dots)$ une famille libre de vecteurs dans $N_j \setminus N_{j-1}$. On note $u^*\mathcal{C}$ la famille $(u^{j-1}(x_0), u^{j-2}(x_0), \dots, u(x_0), x_0, u^{j-1}(y_0), u^{j-2}(y_0), \dots, u(y_0), y_0, \dots)$. Autrement dit : on « pousse » chacun des vecteurs de \mathcal{C} à travers la suite des noyaux de $u - \lambda \text{Id}_V$ à la manière du Lemme 2.4.7.

L'algorithme prend en entrée un endomorphisme $u \in \text{End}_k(V)$ dont le polynôme caractéristique est scindé avec racines λ_i , $i = 1, \dots, r$ connues de multiplicité m_i . Il fournit à la sortie une base de Jordan \mathcal{B} pour u .

Algorithme 3 (Algorithme 2.1 pour calculer la forme normale de Jordan)

Pour chaque $i = 1, \dots, r$, on construit une base \mathcal{B}_i de $V_{(\lambda_i)} = \ker(u - \lambda_i \text{Id}_V)^{m_i}$:

Soit $\mathcal{B}_i = \emptyset$.

Pour $j = m_i, \dots, 1$:

on choisit un base \mathcal{C} d'un supplémentaire de $N_{j-1} \oplus \text{Vect}(\mathcal{B}_i^j)$ dans N_j .

on pose $\mathcal{B}_i := \mathcal{B}_i \sqcup (u - \lambda_i \text{Id}_V)^*\mathcal{C}$.

La base \mathcal{B} recherchée est $\mathcal{B} = \mathcal{B}_1 \sqcup \dots \sqcup \mathcal{B}_r$.

Exemple B.1. — Voici comment fonctionne l'algorithme sur un exemple. Considérons la matrice

$$A = \begin{pmatrix} 2 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} \in M_4(\mathbb{R}).$$

Son polynôme caractéristique est $P_A(X) = (X - 2)^4$, il possède une unique racine $\lambda = 2$ de multiplicité 4.

Nous avons

$$A - 2I_4 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

et l'on calcule $\ker(A - 2I_4) = \text{Vect}(e_2, e_1 + e_4)$. Nous aurons donc une forme de Jordan avec 2 blocs (Corollaire 2.4.4). On calcule ensuite

$$(A - 2I_4)^2 = 0.$$

La taille maximale d'un bloc de Jordan est donc égale à 2 (Corollaire 2.4.5), de sorte que la forme de Jordan est nécessairement

$$\begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & 1 \\ & & 0 & 2 \end{pmatrix}.$$

On considère une base d'un supplémentaire de $\ker(A - 2I_4)$ dans $\ker(A - 2I_4)^2 = \mathbb{R}^4$, par exemple constituée de e_3 et e_4 . On applique $A - 2I_4$ à e_3 et e_4 pour déterminer une base de Jordan : $(A - 2I_4)e_3 = e_1 + e_4$, $(A - 2I_4)e_4 = e_2$. Une base de Jordan est donc $\mathcal{B} = ((A - 2I_4)e_3, e_3, (A - 2I_4)e_4, e_4) = (e_1 + e_4, e_3, e_2, e_4)$. On vérifie qu'avec les matrices de passage

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

nous avons bien

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & 1 \\ & & 0 & 2 \end{pmatrix}.$$

Exemple B.2. — Voici comment se déroule le calcul des puissances de la matrice A de l'exemple précédent, v. §2.4.10. Sa forme de Jordan est $T_{\text{Jordan}} = 2I_4 + \text{Diag}(U_2, U_2)$. Avec P la matrice de changement de base telle que $T_{\text{Jordan}} = P^{-1}AP$, on obtient pour tout $p \geq 1$:

$$A^p = P (2^p I_4 + p2^{p-1} \text{Diag}(U_2, U_2)) P^{-1} = P \begin{pmatrix} 2^p & p2^{p-1} & & \\ 0 & 2^p & & \\ & & 2^p & p2^{p-1} \\ & & 0 & 2^p \end{pmatrix} P^{-1}.$$

Exemple B.3. — Voici comme se déroule le calcul de l'exponentielle $\exp(tA)$ pour la matrice A ci-dessus, v. §2.4.12. On note à nouveau $T_{\text{Jordan}} = 2I_4 + \text{Diag}(U_2, U_2)$ sa forme de Jordan et P la matrice de changement de base telle que $T_{\text{Jordan}} = P^{-1}AP$. Alors :

$$\begin{aligned} \exp(tA) &= P \exp(tT_{\text{Jordan}}) P^{-1} \\ &= P e^{2t} \text{Diag}(\exp(tU_2), \exp(tU_2)) P^{-1} \\ &= e^{2t} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & t & & \\ 0 & 1 & & \\ & & 1 & t \\ & & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \\ &= e^{2t} \begin{pmatrix} 1 & 0 & t & 0 \\ -t & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & t & 1 \end{pmatrix}. \end{aligned}$$

C. Appendice (†) : forme normale de Jordan et classes de conjugaison

Nous donnons ici une application de la forme normale de Jordan aux classes de conjugaison de matrices.

Soient $A, B \in M_n(k)$ deux matrices dont le polynôme caractéristique est scindé sur k . On dit que A et B ont la *même forme normale de Jordan à permutation des blocs* près si elles ont les mêmes valeurs propres et le même nombre de blocs de Jordan pour chaque valeur propre et pour chaque taille de bloc de Jordan.

On dit que deux matrices $A, B \in M_n(k)$ sont *conjuguées* s'il existe une matrice inversible $P \in \text{GL}_n(k)$ telle que $B = P^{-1}AP$.

Exercice C.1. — La relation de conjugaison est une relation d'équivalence sur $M_n(k)$.

Le résultat suivant est une conséquence directe de l'unicité dans le théorème 2.4.2.

Corollaire C.2. — Soient $A, B \in M_n(k)$ deux matrices dont le polynôme caractéristique est scindé sur k . Les affirmations suivantes sont équivalentes :

- (1) A, B sont conjuguées ;
- (2) A, B ont la même forme normale de Jordan à permutation des blocs de Jordan près.

Démonstration. — Supposons que A et B sont conjuguées. Il existe alors un endomorphisme $u \in \text{End}_k(k^n)$ et des bases \mathcal{B} et \mathcal{C} de k^n telles que $M_{\mathcal{B}}(u) = A$ et $M_{\mathcal{C}}(u) = B$. Les formes normales de Jordan de A et B sont des formes normales de Jordan pour u . Par unicité dans le théorème 2.4.2, elles coïncident à permutation des blocs de Jordan près.

Supposons que A et B ont la même forme normale de Jordan à permutation des blocs de Jordan près. Deux telles formes normales de Jordan sont conjuguées puisqu'elles s'obtiennent l'une de l'autre en permutant les éléments de la base canonique. Puisqu'elles sont par définition conjuguées à A , respectivement B , on déduit que A et B sont conjuguées. \square

Exemple C.3 (Classes de conjugaison de matrices nilpotentes)

Dans $M_n(k)$ toute matrice nilpotente possède une forme normale de Jordan puisque son polynôme caractéristique est scindé sur k puisqu'égal à $(-1)^n X^n$. Ceci rend possible l'énumération des classes de conjugaison de matrices nilpotentes.

- $n = 2$. Dans $M_2(k)$ il y a 2 classes de conjugaison de matrices nilpotentes. Elles correspondent aux formes normales de Jordan suivantes : 2 blocs de Jordan de taille 1, respectivement 1 bloc de Jordan de taille 2. De façon explicite :

$$\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}.$$

- $n = 3$. Dans $M_3(k)$ il y a 3 classes de conjugaison de matrices nilpotentes. Elles correspondent aux formes normales de Jordan suivantes : 3 blocs de Jordan de taille 1 ; 1 bloc de Jordan de taille 1 et 1 bloc de Jordan de taille 2 ; 1 bloc de Jordan de taille 3.

$$\begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & & \\ & 0 & 1 \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

- $n = 4$. Dans $M_4(k)$ il y a 5 classes de conjugaison de matrices nilpotentes (lesquelles ?)

D. Appendice (†) : forme normale de Jordan des matrices 3×3

Nous discutons ici les formes normales de Jordan des matrices 3×3 et la manière de les obtenir en utilisant le Lemme 2.4.7, sans passer de manière explicite par l'algorithme 2.1.

Exemple D.1 (Matrices 3×3). — Soit $A \in M_3(k)$ une matrice dont le polynôme caractéristique est scindé sur k . Notons $\lambda_1, \lambda_2, \lambda_3$ ses racines. On note $V = k^3$.

• Si $\lambda_1, \lambda_2, \lambda_3$ sont des racines simples, la matrice est diagonalisable. Une forme normale de Jordan est la même chose qu'une forme diagonale, composée de trois blocs de taille 1 :

$$\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}.$$

• Si λ_1 est une racine simple et $\lambda_2 = \lambda_3 = \lambda$, nous distinguons deux cas.

— Si $\dim \ker(A - \lambda I_3) = 2$, la matrice est diagonalisable. Une forme normale de Jordan est la même chose qu'une forme diagonale, ici composée de trois blocs de taille 1 :

$$\begin{pmatrix} \lambda_1 & & \\ & \lambda & \\ & & \lambda \end{pmatrix}.$$

— Si $\dim \ker(A - \lambda I_3) = 1$, la matrice n'est pas diagonalisable. Sa forme de Jordan comprend un bloc de taille 1 (valeur propre λ_1) et un bloc de taille 2 (valeur propre λ) :

$$\begin{pmatrix} \lambda_1 & & \\ & \lambda & 1 \\ & 0 & \lambda \end{pmatrix}.$$

Pour déterminer une base de Jordan on procède comme dans l'exemple 2.4.9.

• Si $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, nous distinguons trois cas.

— Si $\dim \ker(A - \lambda I_3) = 3$, la matrice est diagonalisable. Une forme normale de Jordan est la même chose qu'une forme diagonale, ici composée de trois blocs de taille 1 :

$$\begin{pmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{pmatrix}.$$

— Si $\dim \ker(A - \lambda I_3) = 2$, la matrice n'est pas diagonalisable. Une forme normale de Jordan est constituée de deux blocs de Jordan (Corollaire 2.4.4). Nous trouvons nécessairement un bloc de taille 1 et un autre de taille 2, les deux de valeur propre λ :

$$\begin{pmatrix} \lambda & & \\ & \lambda & 1 \\ & 0 & \lambda \end{pmatrix}.$$

Pour déterminer une base de Jordan on utilise à nouveau le Lemme 2.4.7, en gardant à l'esprit que l'indice de nilpotence de $A - \lambda I_3$ est égal à celui de sa forme de Jordan, donc égal à 2 : nous choisissons un vecteur $x \in \ker(A - \lambda I_3)^2 \setminus \ker(A - \lambda I_3) = V \setminus \ker(A - \lambda I_3)$; nous lui appliquons $A - \lambda I_3$ pour obtenir la famille libre $((A - \lambda I_3)x, x)$; le vecteur $f = (A - \lambda I_3)x$ est propre pour λ et nous choisissons un deuxième vecteur propre f' linéairement indépendant avec f . La famille $\mathcal{B} = (f', f, x)$ est alors une base de Jordan.

— Si $\dim \ker(A - \lambda I_3) = 1$, la matrice n'est pas diagonalisable. La forme normale de Jordan est constituée d'un seul bloc (Corollaire 2.4.4), nécessairement de taille 3 :

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

L'indice de nilpotence de $A - \lambda I_3$ est égal à celui de sa forme de Jordan, donc égal à 3. Pour déterminer une base de Jordan nous utilisons directement le Lemme 2.4.7 (ce cas est en fait plus simple que le précédent) : nous choisissons un vecteur $x \in \ker(A - \lambda I_3)^3 \setminus \ker(A - \lambda I_3)^2 = V \setminus \ker(A - \lambda I_3)^2$ et nous lui appliquons successivement $A - \lambda I_3$. La base $\mathcal{B} = ((A - \lambda I_3)^2 x, (A - \lambda I_3)x, x)$ est alors une base de Jordan.

E. Appendice : rappels sur l'exponentielle des matrices

Nous faisons ici des rappels sur l'exponentielle des matrices et renvoyons au polycopié du cours 2MA122 pour les détails.⁽⁹⁾ Dans cette section le corps de base est $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Étant donné $A \in M_n(\mathbb{K})$, son exponentielle $\exp(A) \in \text{GL}_n(\mathbb{K})$ est définie comme la somme de la série absolument convergente suivante :

$$\exp(A) = I_n + A + \frac{A^2}{2} + \cdots + \frac{A^p}{p!} + \dots$$

Nous utiliserons librement les propriétés suivantes.

0. (SCALAIRE) Pour $n = 1$ l'exponentielle des matrices coïncide avec la fonction exponentielle définie sur \mathbb{K} , ou encore $\exp(a) = e^a$ pour tout $a \in \mathbb{K}$.

1. (CONJUGAISON) Pour toute matrice $P \in \text{GL}_n(\mathbb{K})$,

$$\exp(P^{-1}AP) = P^{-1}\exp(A)P.$$

2. (COMMUTATION) Si A et B commutent, alors

$$\exp(A + B) = \exp(A)\exp(B) = \exp(B)\exp(A).$$

3. (INVERSIBLE) On a $I_n = \exp(0_n) = \exp(A)\exp(-A)$. Ainsi $\exp(A)$ est bien inversible (la définition comme somme de série ne fournit *a priori* que $\exp(A) \in M_n(\mathbb{K})$) et

$$\exp(A)^{-1} = \exp(-A).$$

4. (NILPOTENCE) Si A est nilpotente d'indice ℓ , alors

$$\exp(A) = I_n + A + \frac{A^2}{2} + \cdots + \frac{A^{\ell-1}}{\ell!}.$$

En effet, la série qui définit l'exponentielle se réduit dans ce cas à une somme finie.

5. (DIAGONALE) Si $A = \text{Diag}(A_1, \dots, A_r)$ est diagonale par blocs, alors

$$\exp(A) = \text{Diag}(\exp(A_1), \dots, \exp(A_r)).$$

En effet, lors du calcul des puissances de A les différents blocs n'interagissent pas.

Lemme E.1. — *L'application $\mathbb{R} \rightarrow M_n(\mathbb{K})$, $t \mapsto \exp(tA)$ est dérivable avec*

$$\frac{d}{dt} \exp(tA) = A \exp(tA) = \exp(tA)A.$$

(Il s'ensuit par un argument de récurrence que cette application est de classe C^∞ .)

Démonstration. — Notons $\|A\| = \sup\{\|Ax\|/\|x\| : x \neq 0\}$ la norme opératorielle sur $M_n(\mathbb{K})$, qui jouit de la propriété remarquable $\|AB\| \leq \|A\| \cdot \|B\|$. En considérant les sommes partielles de la série qui définit l'exponentielle on montre que

$$\|\exp(A) - I_n - tA\| \leq |t|^2 e^{\|A\|}$$

pour tout $t \in \mathbb{K}$. Ceci entraîne que $t \mapsto \exp(tA)$ est dérivable en $t = 0$ avec dérivée égale à A . La preuve de la dérivabilité et le calcul de la dérivée en un point t_0 quelconque sont similaires, en utilisant l'égalité $\exp(tA) = \exp((t - t_0)A)\exp(t_0A)$. \square

⁽⁹⁾Ces résultats sont tout à fait classiques et les lectrices et lecteurs les trouveront démontrés dans de nombreux livres. Une référence possible, que nous avons déjà utilisée, est X. Gourdon, *Les maths en tête : Algèbre*, Ellipses, 2009, §4.3.3.