Réduction des endomorphismes

Carl Tipler

Université de Brest
E-mail addresses: `carl.tipler@univ-brest.fr`

Résumé. Ces notes de cours sont issues d'un cours de L2 donné en 2021 à l'université de Brest. Elles contiennent une introduction à la réduction des endomorphismes, et sont largement inspirées de [**MB**].

Réduire un endomorphisme u, c'est déterminer une base $\mathcal{B}$ dans laquelle la représentation matricielle $[u]_{\mathcal{B}}$ de u soit la plus simple possible afin de faciliter les calculs. Idéalement on cherche une forme diagonale, telle que la matrice suivante :

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

ou, le cas échéant, triangulaire supérieure, telle que la matrice :

$$C = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

La recherche d'une telle forme est motivée par les applications. Ainsi, il est beaucoup plus simple de résoudre le système linéaire $AX = B$ si A est une matrice diagonale ou triangulaire supérieure. Cependant une telle forme n'existe pas toujours, et on verra dans ce cours (Chapitre II) des critères nécessaires et suffisant pour pouvoir diagonaliser (resp. trigonaliser) un endomorphisme (ou de manière équivalente une matrice qui le représente). Les premiers critères porteront sur des calculs effectifs de *valeurs propres* (les $\lambda_1, \lambda_2, \lambda_3$ des matrices D et C) et d'espaces propres associés. On introduira au passage la notion fondamentale de *polynôme caractéristique* χ_u d'un endomorphisme u dont les racines sont les valeurs propres de u (pour D ou C, le polynôme $(T - \lambda_1)(T - \lambda_2)(T - \lambda_3)$).

Dans un second temps, motivés par le théorème de Cayley–Hamilton qui stipule que $\chi_u(u) = 0$, on étudiera de manière plus approfondie l'algèbre des polynômes en u (Chapitre III). On verra en particulier qu'il existe un unique *polynôme minimal* qui divise tout polynôme qui annule u, et dont la décomposition en facteurs premiers permet de conclure si u est diagonalisable. On verra par ailleurs que ce polynôme minimal induit une décomposition en somme directe de l'espace vectoriel ambiant en termes de sous-espaces caractéristiques via le très utile lemme des noyaux. Ces notions seront utilisées dans le Chapitre IV pour réduire de manière efficace les endomorphismes. On y présentera en particulier les formes de Jordan et la décomposition de Dunford.

Enfin, on abordera dans un dernier Chapitre V les applications classiques de la réduction des endomorphismes, telles que les calculs de puissance ou de l'exponentielle, la résolution de systèmes différentiels linéaires ou encore l'étude de suites définies par une relation de récurrence linéaire.

Les références données en fin de texte [**LFA**, **MB**, **MM**, **RW2**] permettront au lecteur intéressé de trouver de nombreux exemples et exercices pour s'entraîner à manipuler les notions abordées. On recommandera aux plus enthousiastes qui souhaitent aller plus loin dans le sujet la lecture du livre [**MM**] qui traite des invariants de similitude, des tableaux de Young ou encore de la localisation des valeurs propres.

Table des matières

Chapitre I

Rappels d'algèbre linéaire

Dans ce chapitre, on rappelle les notions d'algèbre linéaire qui seront utilisées dans ce cours. Les preuves des résultats énoncés peuvent être trouvées dans [**RW1**], ou tout autre livre d'algèbre de première année de licence ou de classe préparatoire.

1. Quelques structures algébriques usuelles

Nous allons rencontrer dans ce cours des espaces vectoriels sur des *corps commutatifs*, des *groupes* de matrices inversibles, ainsi que des *anneaux* de polynômes ou d'applications linéaires. Il convient de définir ces notions.

1.1. Lois de composition interne et externe. Soit E un ensemble. Une *loi de composition interne* sur E, notée $\cdot$, $+$, $*$, $\times$ ou $\circ$ selon les circonstances, est une application

$$E \times E \quad \to \quad E.$$

Étant donné un autre ensemble F, une *loi de composition externe (à gauche)* de F sur E est une application

$$F \times E \quad \to \quad E.$$

Ces lois de composition permettent de définir entre autres les notions de groupes, anneaux, corps et algèbres.

Exemple I.1.1. L'addition est une loi de composition interne sur $\mathbb{N}$, sur $\mathbb{Z}$, sur $\mathbb{R}$, etc. La multiplication est une loi de composition interne sur $\mathbb{Z}$, sur $\mathbb{R}^*$, etc.

1.2. Groupe. Un *groupe* est un couple $(G, \circ)$ où G est un ensemble non vide muni d'une loi de composition interne $\circ$ vérifiant les axiomes suivant :

(i) Associativité : pour tout triplet $(a, b, c) \in G^3$, on a

$$a \circ (b \circ c) = (a \circ b) \circ c$$

(ii) Existence d'un élément neutre : il existe $e \in G$ tel que pour tout $a \in G$,

$$e \circ a = a \circ e = a$$

(iii) Existence d'un inverse (ou symétrique) : pour tout $a \in G$, il existe un inverse a^{-1} vérifiant

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

Dans ce cas, l'élément neutre est unique et tout élément de G admet un unique inverse. Si de plus $(G, \circ)$ vérifie pour tout couple $(a, b) \in G$

$$a \circ b = b \circ a,$$

on dit que $(G, \circ)$ est *commutatif* ou *abélien.*

Exemple I.1.2. L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de l'addition, est un groupe commutatif. L'élément neutre est 0 et l'inverse de $n \in \mathbb{Z}$ est dans ce cas l'opposé $-n$. On utilise ici la *notation additive*. De même, $(\mathbb{R}, +)$ et $(\mathbb{C}, +)$ sont des groupes abéliens.

Contre-exemple I.1.3. L'ensemble $\mathbb{N}$ des entiers naturels muni de l'addition n'est pas un groupe. Il manque les inverses de tous les entiers non nuls.

Exemple I.1.4. Les ensembles $(\mathbb{R}^*, \times)$ et $(\mathbb{C}^*, \times)$ sont des groupes abéliens, où $\times$ désigne la multiplication usuelle. Ici l'élément neutre est 1 et l'inverse de x est x^{-1}. On utilise ici la *notation multiplicative.*

Exemple I.1.5. L'ensemble des permutations d'un ensemble fini E, noté $\mathfrak{S}(E)$, est un groupe pour la loi $\circ$ de composition des applications définies de la façon suivante pour $\sigma_1, \sigma_2 \in \mathfrak{S}(E)$:

$$\sigma_1 \circ \sigma_2 : \begin{array}{ccc} E & \to & E \\ x & \mapsto & \sigma_1(\sigma_2(x)). \end{array}$$

Cet exemple est non commutatif dès que E est de cardinal supérieur à 3.

L'exemple important dans ce cours est le groupe des applications linéaires inversibles d'un espace vectoriel de dimension finie V, noté $GL(V)$. On reviendra plus tard sur cet exemple.

1.3. Anneaux et corps. Un *anneau* (unitaire[1]) est un triplet $(\mathbb{A}, +, \times)$ où $\mathbb{A}$ est un ensemble muni de deux lois de composition interne $+$ et $\times$ qui vérifient les axiomes suivants :

(i) $(\mathbb{A}, +)$ est un groupe abélien,

(ii) $\times$ admet un élément neutre,

(iii) $\times$ est associative, i.e. pour tout $(x, y, z) \in \mathbb{A}$,

$$x \times (y \times z) = (x \times y) \times z,$$

(iv) $\times$ est distributive par rapport à $+$: pour tout $(x, y, z) \in \mathbb{A}^3$,

$$x \times (y + z) = x \times y + x \times z$$

et

$$(y + z) \times x = y \times x + z \times x.$$

Si de plus pour tout $(x, y) \in \mathbb{A}^2$ on a

$$x \times y = y \times x,$$

on dira que $(\mathbb{A}, +, \times)$ est un *anneau commutatif.* L'élément neutre de $(\mathbb{A}, +)$ sera appelé l'élément nul, souvent noté 0, et le neutre de $(\mathbb{A}, \times)$ sera appelé l'unité, souvent noté 1.

Exemple I.1.6. Le prototype d'anneau commutatif est l'ensemble des entiers relatifs $\mathbb{Z}$ muni de son addition et de sa multiplication usuelles. Les ensembles des entiers rationnels $\mathbb{Q}$, des réels $\mathbb{R}$ ou des complexes $\mathbb{C}$ sont également des anneaux commutatifs.

Exemple I.1.7. On verra bientôt un autre exemple important : celui des matrices carrées à coefficients réels (resp. complexes) $\mathrm{M}_n(\mathbb{R})$ (resp. $\mathrm{M}_n(\mathbb{C})$). C'est un anneau, non commutatif si $n \geqslant 2$, pour l'addition et la multiplication des matrices.

Exemple I.1.8. L'ensemble $\mathbb{R}[X]$ (resp. $\mathbb{C}[X]$) des polynômes à coefficients dans $\mathbb{R}$ (resp. dans $\mathbb{C}$) est un anneau pour l'addition et la multiplication des polynômes.

---

1. L'axiome (*ii*) d'existence d'un élément neutre pour la loi $\times$ n'est pas toujours requis dans la définition d'anneau. Dans ce cours, tous les anneaux seront unitaires, et on inclura cet axiome dans la définition.

Un *corps commutatif* est un triplet $(\mathbb{K}, +, \times)$ où $(\mathbb{K}, +, \times)$ est un anneau commutatif non nul (i.e. non réduit à $\{0\}$) tel que tout élément non nul admette un inverse pour la loi $\times$. Ceci implique en particulier que $(\mathbb{K}\backslash\{0\}, \times)$ est un groupe commutatif.

Exemple I.1.9. Les réels $(\mathbb{R}, +, \times)$ et les complexes $(\mathbb{C}, +, \times)$ forment des exemples de corps commutatifs. Même si dans ce cours la plupart des exemples seront traités sur $\mathbb{R}$ ou $\mathbb{C}$, il faut garder à l'esprit qu'il existe de nombreux autres corps qui seront étudiés en licence (les corps finis, le corps des rationnels, les corps de nombres, les corps de fractions).

Contre-exemple I.1.10. L'ensemble des entiers relatifs $(\mathbb{Z}, +, \times)$ n'est pas un corps.

2. Espaces vectoriels

Soient $n \in \mathbb{N}^*$ un entier positif et $\mathbb{K}$ un corps commutatif (que l'on pourra supposer être égal à $\mathbb{R}$ ou $\mathbb{C}$ lors d'une première lecture).

2.1. Définitions, exemples. La notion d'espace vectoriel a été introduite dans les années 1843–1845 indépendamment par Cayley et Grassmann. Elle a permis de généraliser des formules algébriques obtenues dans le plan et l'espace et de systématiser l'utilisation des transformations linéaires déjà utilisées à l'époque comme changements de coordonnées.

Définition I.2.1. Un $\mathbb{K}$-espace vectoriel (ou espace vectoriel sur $\mathbb{K}$) est un triplet $(E, +, \cdot)$ où E est un ensemble muni d'une loi de composition interne $+$ et d'une loi de composition externe (à gauche) $\cdot$ vérifiant les axiomes suivants (pour $\lambda \in \mathbb{K}$ et $x \in E$, on notera $\lambda \cdot x$ par λx pour alléger le texte) :

(i) $(E, +)$ est un groupe abélien,

(ii) La loi externe est distributive par rapport à la loi interne de E : pour $\lambda \in \mathbb{K}$, $(x, y) \in E^2$,
$$\lambda(x + y) = \lambda x + \lambda y,$$

(iii) La loi externe est distributive par rapport à la loi interne de $\mathbb{K}$: pour $(\lambda, \mu) \in \mathbb{K}^2$, $x \in E$,
$$(\lambda + \mu)x = \lambda x + \mu x,$$

(iv) Si $1_{\mathbb{K}}$ désigne le neutre de $(\mathbb{K}\backslash\{0\}, \times)$, on a pour tout $x \in E$,
$$1_{\mathbb{K}} \cdot x = x,$$

(v) Enfin, pour tout $(\lambda, \mu) \in \mathbb{K}^2$ et tout $x \in E$,
$$\lambda(\mu x) = (\lambda\mu)x.$$

Les éléments de $\mathbb{K}$ sont appelés *scalaires* et les éléments de E des *vecteurs*.

De ces axiomes on peut déduire les conséquences suivantes :

Lemme I.2.2. *Pour tout $x \in E$ et tout $\lambda \in \mathbb{K}$:*

(1) Si $0_{\mathbb{K}}$ désigne le neutre de $(\mathbb{K}, +)$,
$$0_{\mathbb{K}} \cdot x = 0_E,$$

(2) On a
$$(-1_{\mathbb{K}}) \cdot x = -x,$$

(3) Si $\lambda x = 0_E$, alors $\lambda = 0_{\mathbb{K}}$ ou $x = 0_E$.

Les premiers exemples d'espaces vectoriels sur $\mathbb{R}$ sont le plan $\mathbb{R}^2$ et l'espace $\mathbb{R}^3$.

Exemple I.2.3. L'ensemble $\mathbb{K}^n$ des n-uplets d'éléments de $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel, où les lois sont données, pour $\lambda \in \mathbb{K}$, $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ et $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$ par

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n),$$

$$x + y = (x_1 + y_1, \ldots, x_n + y_n),$$

et l'élément neutre de $\mathbb{K}^n$ est

$$0_{\mathbb{K}^n} = (0, \ldots, 0).$$

Un autre exemple important dans ce cours est celui des polynômes à une indéterminée sur $\mathbb{K}$:

Exemple I.2.4. L'ensemble $\mathbb{K}[X]$ est un espace vectoriel sur $\mathbb{K}$ muni des lois $+$ et $\cdot$ définies, pour $\lambda \in \mathbb{K}$, $p(X) = \sum_{i=0}^{d} a_i X^i$ et $q(X) = \sum_{i=0}^{d'} b_i X^i$ par

$$\lambda \cdot p = \sum_{i=0}^{d} \lambda a_i X^i$$

et

$$p(X) + q(X) = \sum_{i=0}^{\max(d,d')} (a_i + b_i) X^i,$$

où l'on a posé $a_i = 0$ si $i > d$ et $b_i = 0$ si $i > d'$. Enfin, l'élément neutre de $\mathbb{K}[X]$ est le polynôme nul.

Dans la suite de ce chapitre, E désigne un $\mathbb{K}$-espace vectoriel.

Définition I.2.5. Un sous-ensemble non-vide $F \subset E$ est un sous-espace vectoriel de E si F est stable par les lois de composition interne et externe de E, et si, muni de la restriction de ces lois, F est lui-même un espace-vectoriel sur $\mathbb{K}$.

Proposition I.2.6. *$F \subset E$ est un sous-espace vectoriel de E si :*

(i) $F \neq \varnothing$

(ii) $\forall \lambda \in \mathbb{K}$, $\forall (x, y) \in F^2$, $\lambda x + y \in F$.

Exemples I.2.7. Toute droite qui passe par l'origine du plan $\mathbb{R}^2$ (resp. de l'espace $\mathbb{R}^3$) est un sous-espace vectoriel de $\mathbb{R}^2$ (resp. de l'espace $\mathbb{R}^3$). Tout plan contenant l'origine dans $\mathbb{R}^3$ est un sous-espace vectoriel de $\mathbb{R}^3$. En revanche, une droite affine qui ne contient pas l'origine n'est pas un sous-espace vectoriel du plan.

L'espace $\mathbb{R}_2[X]$ des polynômes de degré inférieur ou égal à 2 est un sous-espace vectoriel de $\mathbb{R}[X]$. L'espace des polynômes à coefficients réels de degré égal à 1 n'est pas un sous-espace vectoriel de $\mathbb{R}[X]$.

Une partie $\mathcal{P}$ de E n'est pas forcément un sous-espace vectoriel. Cependant, on peut construire un sous-espace vectoriel à partir de $\mathcal{P}$ on considérant le plus petit sous-espace vectoriel de E contenant $\mathcal{P}$:

DÉFINITION I.2.8. Soit $\mathcal{P}$ une partie de E. L'espace vectoriel engendré par $\mathcal{P}$, noté Vect $(\mathcal{P})$ ou $< \mathcal{P} >$ est l'ensemble des combinaisons linéaires finies d'éléments de $\mathcal{P}$, i.e. :

$$\mathrm{Vect}\,(\mathcal{P}) := \left\{ \sum_{i=1}^{m} \lambda_i v_i, m \in \mathbb{N}, \lambda_i \in \mathbb{K}, v_i \in \mathcal{P} \right\}.$$

C'est le plus petit sous-espace vectoriel de E contenant $\mathcal{P}$. On notera, pour $v \in E \backslash \{0\}$,

$$\mathbb{K}v = \mathrm{Vect}\,\{v\} = \{\lambda v, \lambda \in \mathbb{K}\}.$$

EXEMPLE I.2.9. Dans $\mathbb{R}^3$, si $\mathcal{P} = \{(1,0,0),(1,0,-2)\}$, on a

$$\begin{array}{rcc} \mathrm{Vect}\,\mathcal{P} & = & \{\lambda(1,0,0) + \mu(1,0,-2), (\lambda,\mu) \in \mathbb{R}^2\} \\ & = & \{(\lambda+\mu, 0, -2\mu), (\lambda,\mu) \in \mathbb{R}^2\}, \\ & = & \{(\lambda', 0, \mu'), (\lambda',\mu') \in \mathbb{R}^2\}, \\ & = & \{(x,y,z) \in \mathbb{R}^3, y = 0\}. \end{array}$$

Une autre construction classique est celle de somme d'espaces vectoriels :

DÉFINITION I.2.10. Soient $F_1, \ldots, F_r$ des sous-espaces vectoriels de E. On appelle somme des sous-espaces $F_1, \ldots, F_r$, notée $F_1+F_2+\ldots+F_r$, l'espace vectoriel engendré par $F_1 \cup \ldots \cup F_r$. Autrement dit,

$$F_1 + \ldots + F_r = \{x_1 + \ldots + x_r, x_i \in F_i, 1 \leqslant i \leqslant r\}.$$

EXEMPLE I.2.11. Dans $\mathbb{R}^3$, on considère les sous-espaces $F_1 = \mathrm{Vect}\,\{(1,0,0),(1,1,0)\}$ et $F_2 = \mathrm{Vect}\,\{(0,1,1),(0,0,1)\}$. On peut vérifier que $F_1 + F_2 = \mathbb{R}^3$. Soit $x = (1,1,1) \in \mathbb{R}^3$. On peut remarquer que x s'écrit $x = x_1 + x_2$ avec $x_1 = (1,0,0) \in F_1$ et $x_2 = (0,1,1) \in F_2$, mais aussi $x = x_1' + x_2'$, avec $x_1' = (1,1,0) \in F_1$ et $x_2' = (0,0,1) \in F_2$. L'écriture d'un vecteur x comme somme de vecteurs de F_1 et F_2 n'est donc pas unique.

Comme le suggère l'exemple précédent, l'écriture $x = x_1 + \ldots + x_r$ avec $x_i \in F_i$ n'est pas unique dans la somme des F_i. On distingue les sommes d'espaces où c'est le cas :

DÉFINITION I.2.12. Soient $F_1, \ldots, F_r$ des sous-espaces vectoriels de E. On dit qu'ils sont en somme directe si pour tout $(x_i)_{1 \leqslant i \leqslant r} \in F_1 \times \ldots \times F_r$, la condition

$$x_1 + \ldots + x_r = 0$$

implique

$$\forall i \in [\![1, r]\!],\ x_i = 0.$$

On note alors la somme

$$F_1 + \ldots + F_r = F_1 \oplus \ldots \oplus F_r.$$

De manière équivalente, tout élément x de $F_1 \oplus \ldots \oplus F_r$ s'écrit de manière unique comme somme $x = x_1 + \ldots + x_r$, $x_i \in F_i$, $1 \leqslant i \leqslant r$.

LEMME I.2.13. *Des sous-espaces vectoriels $F_1, \ldots, F_r$ de E sont en somme directe si et seulement si pour tout $i \in [\![1, r]\!]$,*

$$F_i \cap (F_1 + \ldots + F_{i-1} + F_{i+1} + \ldots + F_r) = \{0\}.$$

Décomposer l'espace E en somme directe $E = F_1 \oplus \ldots \oplus F_r$ de sous-espaces F_i permet souvent de simplifier son étude, par restriction aux sous-espaces F_i. Deux cas sont souvent rencontrés. Le premier, quand E est somme directe de deux sous-espaces :

Définition I.2.14. Soit F un sous-espace vectoriel de E. On appelle supplémentaire de F dans E tout sous-espace vectoriel $G \subset E$ vérifiant

$$E = F \oplus G.$$

Proposition I.2.15. *Soit F un sous-espace vectoriel de E. Alors F admet un supplémentaire dans E.*

Remarque I.2.16. En général, un sous-espace vectoriel admet de nombreux supplémentaires, on n'a pas unicité.

Exemples I.2.17. Dans $\mathbb{R}^2$, Vect $\{(1,0)\}$ admet pour supplémentaire Vect $\{(\lambda, 1)\}$, quelque soit λ.

Exercice I.2.18. Déterminer un supplémentaire de Vect $\{(1,0,1),(1,1,0)\}$ dans $\mathbb{R}^3$ et un supplémentaire de $\mathbb{R}_2[X]$ dans $\mathbb{R}[X]$.

Le second cas de décomposition suggère le concept de base, quand E est somme directe de sous-espaces les plus simples possibles, du type Vect $\{v_i\}$, pour $v_i \in E$.

2.2. Bases. La notion de base d'un espace vectoriel permet de se ramener à des calculs en coordonnées.

Définition I.2.19. Une famille $\mathcal{F} = \{v_i, i \in I\}$ de vecteurs de E est dite libre si pour tout $J \subset I$ fini et tout $(\lambda_j)_{j \in J} \in \mathbb{K}^J$, la condition

$$\sum_{j \in J} \lambda_j v_j = 0,$$

implique

$$\forall j \in J, \lambda_j = 0.$$

Autrement dit, une famille de vecteurs est libre s'il n'existe pas de combinaison linéaires nulles de ces vecteurs qui ne soit pas triviale. Une famille non libre est dite *liée*.

Exemple I.2.20. La famille $\{(1,0),(1,-1)\}$ de $\mathbb{R}^2$ est libre. En revanche, la famille $\{(1,-1),(-2,2)\}$ est liée.

Remarque I.2.21. Les faits suivants sont immédiats :

(1) Une famille de vecteurs $\{v_1, \ldots, v_r\}$ est liée s'il existe $\lambda_1, \ldots, \lambda_r$ non tous nuls tels que $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$.
(2) Si $\{v_i, i \in I\}$ est libre, alors pour tout $i \neq j$, $v_i \neq v_j$.
(3) $\varnothing$ est libre mais $\{0\}$ ne l'est pas.
(4) Toute sous-famille d'une famille libre est libre.
(5) Si $v \neq 0$, $\{v\}$ est libre.
(6) Une famille $\{v_1, \ldots, v_r\}$ est libre si et seulement si $\mathbb{K}v_1 + \ldots + \mathbb{K}v_r = \mathbb{K}v_1 \oplus \ldots \oplus \mathbb{K}v_r$.

Définition I.2.22. Une famille $\mathcal{F} = \{v_i, i \in I\}$ de vecteurs de E est dite génératrice de E si

$$\mathrm{Vect}\,(\mathcal{F}) = E.$$

Autrement dit, une famille $\mathcal{F}$ est génératrice si tout vecteur de E peut s'écrire comme combinaison linéaire finie d'éléments de $\mathcal{F}$.

Exemples I.2.23. La famille $\{(0,1),(1,-1)\}$ est génératrice de $\mathbb{R}^2$, tandis que la famille $\{1, X, X^2, X^3, \ldots\}$ est génératrice dans $\mathbb{R}[X]$.

Remarque I.2.24. Le lecteur pourra se convaincre des faits suivants :

(1) Si $\mathcal{F}$ est génératrice de E, alors toute famille $\mathcal{F}'$ contenant $\mathcal{F}$ est génératrice de E. La réciproque est fausse.

(2) Si pour $i \in [\![1, r]\!]$, $\mathcal{G}_i$ est une famille génératrice d'un sous espace vectoriel F_i de E, alors $\displaystyle\bigcup_{i=1}^{r} \mathcal{G}_i$ est une famille génératrice de $F_1 + \ldots + F_r$.

(3) Une famille $\{v_1, \ldots, v_r\}$ est génératrice si et seulement si $E = \mathbb{K}v_1 + \ldots + \mathbb{K}v_r$.

(4) Une famille génératrice n'est pas forcément libre. Une famille libre n'est pas nécessairement génératrice.

Définition I.2.25. Une famille $\mathcal{B}$ de vecteurs de E est une base si elle est libre et génératrice.

Exemples I.2.26. La *base canonique* de $\mathbb{K}^n$ est donnée par la famille $\{e_1, \ldots, e_n\}$ où pour tout i, $e_i = (0, 0, \ldots, 1, 0, \ldots,)$ où toutes les coordonnées de e_i sont nulles, sauf un 1 en i-ème position.

La famille $\{1, X, X^2, X^3, \ldots\}$ est une base de $\mathbb{K}[X]$.

Une base $\mathcal{B}$ permet d'écrire tout vecteur de E de manière unique comme combinaison linéaire d'éléments de $\mathcal{B}$.

Proposition I.2.27. *Soit $\mathcal{B} := \{e_1, \ldots, e_n\}$ une base de E. Alors pour tout $x \in E$ il existe un unique n-uplet $(\lambda_i) \in \mathbb{K}^n$ tel que*

$$x = \sum_{i=1}^{n} \lambda_i e_i.$$

Réciproquement, si

$$E = \mathbb{K}e_1 \oplus \ldots \oplus \mathbb{K}e_r,$$

alors $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base de E.

Les scalaires $(\lambda_i) \in \mathbb{K}^n$ dans la proposition précédente sont les *coordonnées* du vecteur x dans la base $\mathcal{B}$. On peut alors se ramener à des calculs en coordonnées dans $\mathbb{K}^n$, sous réserve d'avoir une base à disposition.

Théorème I.2.28. *Tout $\mathbb{K}$-espace vectoriel admet une base.*

Le théorème suivant est appelé “théorème de la base incomplète”, et donne un algorithme pour trouver une base à partir d'une famille génératrice.

Théorème I.2.29. *Soient $\mathcal{G}$ une famille génératrice finie et $\mathcal{L}$ une famille libre de E, avec $\mathcal{L} \subset \mathcal{G}$. Alors il existe une base $\mathcal{B}$ de E avec $\mathcal{L} \subset \mathcal{B} \subset \mathcal{G}$.*

Notons qu'une base n'est pas forcément de cardinal fini.

Définition I.2.30. On dit qu'un espace vectoriel est de dimension finie s'il possède une famille génératrice finie.

Dans le cas contraire, l'espace vectoriel est de *dimension infinie*. Ceci est équivalent à l'existence d'une famille libre infinie.

Théorème I.2.31. *Dans un espace vectoriel E de dimension finie, toutes les bases ont même cardinal. On appelle dimension de E ce cardinal, noté* $\dim_{\mathbb{K}}(E)$ *ou simplement* $\dim(E)$ *si le corps des scalaires est sous-entendu.*

Exemples I.2.32. L'espace $\mathbb{K}^n$ est de dimension n, l'espace $\mathbb{K}_n[X]$ de dimension $n+1$ et l'espace $\mathbb{K}[X]$ est de dimension infinie.

Le Théorème I.2.31 repose sur un résultat intermédiaire qui stipule que dans un espace vectoriel de dimension finie, le cardinal d'une famille libre est toujours inférieur au cardinal d'une famille génératrice.

Corollaire I.2.33. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Alors, pour $\mathcal{F}$ une famille de vecteurs de E :*

(1) $\mathcal{F}$ est une base de E si et seulement si $\mathcal{F}$ est libre et de cardinal n,

(2) $\mathcal{F}$ est une base de E si et seulement si $\mathcal{F}$ est génératrice et de cardinal n,

(3) Si $\mathcal{F}$ est génératrice, $\sharp\mathcal{F} \geqslant n$,

(4) Si $\mathcal{F}$ est libre, $\sharp\mathcal{F} \leqslant n$.

On a aussi :

Corollaire I.2.34. *Si E est un $\mathbb{K}$-espace vectoriel de dimension finie n, alors tout sous-espace vectoriel F de E est de dimension finie avec* $\dim(F) \leqslant n$. *De plus, si $\mathcal{F}$ est une base de F, il existe une base de E contenant $\mathcal{F}$.*

Remarque I.2.35. Dans les cas extrêmes, pour F un sous-espace vectoriel d'un espace vectoriel E de dimension finie :

(1) $\dim(F) = 0$ si et seulement si $F = \{0\}$,

(2) $\dim(F) = \dim(E)$ si et seulement si $F = E$.

On obtient aussi une caractérisation des sommes directes en termes de dimension :

Proposition I.2.36. *Soient $F_1, \ldots, F_r$ des sous espaces vectoriels de dimension finie de E. Alors*

$$\dim(F_1 + \ldots + F_r) \leqslant \dim(F_1) + \ldots + \dim(F_r)$$

avec égalité si et seulement si les F_i sont en somme directe. Dans ce cas, si $\mathcal{B}_i$ est une base de F_i pour tout $i \in [\![1, r]\!]$, $\bigcup_{i=1}^{r} \mathcal{B}_i$ est une base de $F_1 \oplus \ldots \oplus F_r$.

Exercice I.2.37. Dans $\mathbb{C}^3$, on considère l'espace $F = \{(x, y, z), x + y + z = 0\}$.

(1) Déterminer une base de F.

(2) En déduire la dimension de F.

(3) Soit $v \in \mathbb{C}^3$. Quelles sont les valeurs possible pour la dimension de $\mathrm{Vect}\{v\} + F$?

(4) L'espace $\mathrm{Vect}\{(1, j, j^2)\}$ est-il un supplémentaire de F dans $\mathbb{C}^3$ (on rappelle que $j = e^{\frac{2i\pi}{3}}$) ?

(5) Même question pour $\mathrm{Vect}\{(e^{i\frac{\pi}{4}}, 0, e^{-i\frac{\pi}{4}})\}$.

3. Applications linéaires, matrices

Les applications linéaires, et leurs représentations matricielles, forment l'objet d'étude de ce cours. Elles apparaissent d'abord historiquement en algèbre dans la résolution d'équations linéaires et en géométrie comme des changement de coordonnées que l'on utilise pour étudier les courbes ou les surfaces du plan et de l'espace. On les retrouve également en analyse comme approximations au premier ordre des fonctions différentiables (dérivées, ou différentielles en dimension supérieure), en théorie des probabilités (chaînes de Markov), en statistiques (méthodes de régression linéaire), etc.

3.1. Matrices. On commence par rappeler la notion de matrice. Soient m et n deux entiers non nuls. Une matrice de taille (m, n) à coefficients dans $\mathbb{K}$ est la donnée de n colonnes de taille m et dont les entrées sont dans $\mathbb{K}$, représentée de la façon suivante :

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & & \vdots & & a_{2n} \\ \vdots & & a_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix},$$

où tous les *coefficients* a_{ij} de la matrice sont dans $\mathbb{K}$. L'indice ij de a_{ij} signifie que a_{ij} est placé sur la i-ème ligne et la j-ème colonne de la matrice. On notera également une telle matrice par

$$[a_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}.$$

Exemples I.3.1. Les matrices suivantes sont des matrices de taille respectivement $(2, 2)$, $(2, 3)$ et $(3, 1)$ à coefficients réels :

$$\begin{bmatrix} 1 & 2 \\ \pi & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 8 \\ 3 \\ 12 \end{bmatrix}.$$

On dénote $\mathrm{M}_{m,n}(\mathbb{K})$ l'ensemble de toutes les matrices de taille (m, n) à coefficients dans $\mathbb{K}$. Les éléments de $\mathrm{M}_{m,1}(\mathbb{K})$ sont les vecteurs colonnes de tailles m et ceux de $\mathrm{M}_{1,n}$ les vecteurs lignes de tailles n. Si $n = m$, on parle de matrice carrée de taille n, et l'ensemble de ces matrices est noté $\mathrm{M}_n(\mathbb{K})$.

On peut multiplier par un élément $\lambda \in \mathbb{K}$ une matrice $A = [a_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n} \in \mathrm{M}_{m,n}(\mathbb{K})$ de la façon suivante :

$$\lambda \cdot A = [\lambda a_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n} = \begin{bmatrix} \lambda a_{11} & \dots & \lambda a_{1j} & \dots & \lambda a_{1n} \\ \lambda a_{21} & & \vdots & & \lambda a_{2n} \\ \vdots & & \lambda a_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mj} & \dots & \lambda a_{mn} \end{bmatrix}.$$

On peut également définir une addition sur les matrices. Pour $A = [a_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ et $B = [b_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ deux éléments de $\mathrm{M}_{m,n}(\mathbb{K})$, on pose

$$A + B := [a_{ij} + b_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n} = \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1j}+b_{1j} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & & \vdots & & a_{2n}+b_{2n} \\ \vdots & & a_{ij}+b_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mj}+b_{mj} & \dots & a_{mn}+b_{mn} \end{bmatrix}.$$

Enfin, si $p \in \mathbb{N}^*$ est un troisième entier naturels non nul, pour $A = [a_{ij}]_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n} \in \mathrm{M}_{m,n}(\mathbb{K})$ et $B = [b_{jk}]_{1\leqslant j\leqslant n, 1\leqslant k\leqslant p} \in \mathrm{M}_{n,p}(\mathbb{K})$, on peut définir la matrice produit $C = AB \in \mathrm{M}_{m,p}(\mathbb{K})$,

$$C = [c_{ik}]_{1\leqslant i\leqslant m, 1\leqslant k\leqslant p}$$

par la formule

$$\forall i, k \in [\![1, m]\!] \times [\![1, p]\!],\ c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Exemple I.3.2. Vérifier que l'on a

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 10 \\ 0 & 3 & 0 \end{bmatrix}$$

Proposition I.3.3. *L'ensemble* $(\mathrm{M}_{m,n}(\mathbb{K}), +, \cdot)$ *est un* $\mathbb{K}$*-espace vectoriel de dimension* mn.

Une base de cet espace est donnée par les matrices $E_{ij} = [e_{kl}]_{1\leqslant k\leqslant m, 1\leqslant l\leqslant n} \in \mathrm{M}_{m,n}(\mathbb{K})$ dont les entrées e_{kl} sont toutes nulles, sauf $e_{ij} = 1$, i.e. $e_{kl} = \delta_{ki}\delta_{lj}$ (on rappelle que δ_{ij} désigne le symbole de Kronecker [2]).

Par ailleurs, comme le produit de deux matrices carrées de taille n est encore une matrice carrée de taille n, on peut composer les produits, et on obtient :

Proposition I.3.4. *L'ensemble* $(\mathrm{M}_n(\mathbb{K}), +, \times)$ *est un anneau unitaire, non commutatif si* $n \geqslant 2$.

Dans la proposition, $\times$ désigne le produit matriciel. L'unité pour la somme est la *matrice nulle* (dont toutes les entrées sont nulles), et l'unité pour le produit est la *matrice identité*, notée Id_n et donnée par la matrice carrée de taille n qui possède des zéros partout en dehors de la diagonale, et des 1 sur la diagonale :

$$\mathrm{Id}_n := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & 1 & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

En fait, on peut vérifier que le produit interne $\times$ est *bi-linéaire* par rapport à la structure d'espace vectoriel sur $\mathrm{M}_n(\mathbb{K})$, c'est à dire que pour tout couple de scalaires $(\lambda, \mu) \in \mathbb{K}^2$ et tout triplet $(A, B, C) \in \mathrm{M}_n(\mathbb{K})$ on a

$$A \times (\lambda B + \mu C) = \lambda(A \times B) + \mu(A \times C)$$

2. $\delta_{ij} = 0$ is $i \neq j$ et 1 si $i = j$.

et
$$(\lambda B + \mu C) \times A = \lambda(B \times A) + \mu(C \times A).$$
Cette propriété supplémentaire implique :

PROPOSITION I.3.5. *L'ensemble* $(\mathrm{M}_n(\mathbb{K}), +, \cdot, \times)$ *est une algèbre*[3] *sur* $\mathbb{K}$.

Le calcul matriciel se généralise aisément aux "matrices par blocs". Soient $A \in \mathrm{M}_{m,n}(\mathbb{K})$ et $B \in \mathrm{M}_{n,p}(\mathbb{K})$. Si l'on décompose les entiers m, n et p en sommes $m = m_1 + \ldots + m_s$, $n = n_1 + \ldots + n_t$ et $p = p_1 + \ldots + p_u$, où les m_i, n_j et p_k sont des entiers non nuls, on peut décomposer les matrices A et B par blocs :
$$A = \begin{bmatrix} A_{11} & \ldots & A_{1j} & \ldots & A_{1t} \\ A_{21} & & \vdots & & A_{2t} \\ \vdots & & A_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ A_{s1} & \ldots & A_{sj} & \ldots & A_{st} \end{bmatrix},$$
et
$$B = \begin{bmatrix} B_{11} & \ldots & B_{1j} & \ldots & B_{1u} \\ B_{21} & & \vdots & & B_{2u} \\ \vdots & & B_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ B_{t1} & \ldots & B_{tj} & \ldots & B_{tu} \end{bmatrix},$$
où les entrées sont cette fois-ci elles mêmes des matrices $A_{kl} \in \mathrm{M}_{m_k,n_l}(\mathbb{K})$ et $B_{lq} \in \mathrm{M}_{n_l,p_q}(\mathbb{K})$. Les produits $A_{kl}B_{lq}$ sont bien définis et appartiennent à l'espace $\mathrm{M}_{m_k,p_q}(\mathbb{K})$. On obtient alors le produit par blocs $C = AB \in \mathrm{M}_{m,p}(\mathbb{K})$:
$$C = \begin{bmatrix} C_{11} & \ldots & C_{1j} & \ldots & C_{1u} \\ C_{21} & & \vdots & & C_{2u} \\ \vdots & & C_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ C_{s1} & \ldots & C_{sj} & \ldots & C_{su} \end{bmatrix},$$
où $C_{ij} \in \mathrm{M}_{m_i,p_j}(\mathbb{K})$ est donnée par la formule :
$$C_{ij} = \sum_{l=1}^{t} A_{il} B_{lj}.$$

EXERCICE I.3.6. (1) Si X et Y sont diagonales par blocs,
$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$
et
$$Y = \begin{bmatrix} E & F \\ G & H \end{bmatrix},$$

3. Une $\mathbb{K}$-algèbre $(A, +, \cdot, \times)$ est un $\mathbb{K}$-espace vectoriel $(A, +, \cdot)$ muni d'une loi de composition interne $\times$ qui est bilinéaire par rapport à $(+, \cdot)$.

que doivent vérifier les tailles des différentes matrices A, B, C, D, E, F, G et H pour que le calcul par bloc de XY et YX ait un sens ?

(2) Dans ce cas, montrer que $Z = XY$ vérifie

$$Z = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}.$$

(3) Calculer le produit de matrices diagonales par blocs.

(4) Montrer que si X et Y sont triangulaires supérieures par blocs, i.e. :

$$X = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

et

$$Y = \begin{bmatrix} D & E \\ 0 & F \end{bmatrix},$$

de tailles convenables, alors leur produit est triangulaire supérieur par bloc, plus précisément :

$$XY = \begin{bmatrix} AD & AE + BF \\ 0 & CF \end{bmatrix}.$$

Les matrices apparaissent naturellement dans ce cours pour les raisons suivantes : les vecteurs sont représentés par des matrices colonnes, les changements de bases se traduisent par des matrices de passage, et les applications linéaires d'un espace vectoriel sont représentées par des matrices carrées.

3.2. Matrices de changement de base. Dans toute la suite, E est un $\mathbb{K}$-espace vectoriel de dimension finie n et de base $\mathcal{B} = \{e_1, \dots, e_n\}$. Tout vecteur $x \in E$ s'écrit de manière unique

$$x = \sum_{i=1}^{n} x_i e_i,$$

où les $x_i \in \mathbb{K}$ sont les coordonnées de x dans la base $\mathcal{B}$. On va représenter chaque vecteur x par le *vecteur colonne* X de ses coordonnées :

$$[x]_{\mathcal{B}} := X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

et on note l'espace de tous les vecteurs colonnes ainsi obtenus par

$$\mathrm{M}_{n,1}(\mathbb{K}) := \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, (x_1, \dots, x_n) \in \mathbb{K}^n \right\}.$$

Bien entendu, le vecteur colonne obtenu pour représenter x dépend du choix de la base. On va introduire les matrices pour étudier l'effet d'un changement de base sur les coordonnées. Soit $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ une autre base de E. Alors on a

$$x = \sum_{i=1}^{n} x'_i e'_i,$$

et on obtient le vecteur colonne

$$X' := [x]_{\mathcal{B}'} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}.$$

Par ailleurs, comme $\mathcal{B}$ est une base de E, pour chaque $i \in [\![1, n]\!]$, il existe un unique n-uplet $(p_{1i}, \dots, p_{ni})$ de coordonnées de e'_i dans $\mathcal{B}$ telles que

$$e'_i = \sum_{k=1}^{n} p_{ki} e_k.$$

On en déduit

$$\begin{aligned} x &= \sum_{i=1}^{n} x'_i e'_i \\ &= \sum_{i=1}^{n} x'_i (\sum_{k=1}^{n} p_{ki} e_k) \\ &= \sum_{k=1}^{n} (\sum_{i=1}^{n} x'_i p_{ki}) e_k \end{aligned}$$

et donc, par unicité de l'écriture dans $\mathcal{B}$:

$$(1) \qquad \forall k \in [\![1, n]\!],\ x_k = \sum_{i=1}^{n} p_{ki} x'_i.$$

On introduit alors la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, notée $P_{\mathcal{B}}^{\mathcal{B}'}$ (parfois $P_{\mathcal{B}\to\mathcal{B}'}$) :

$$P = P_{\mathcal{B}}^{\mathcal{B}'} := [p_{kj}]_{1 \leqslant k,j \leqslant n} = \begin{bmatrix} p_{11} & \vdots & p_{1i} & \vdots & p_{1n} \\ p_{21} & \vdots & p_{2i} & \vdots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{k1} & \vdots & p_{ki} & \vdots & p_{kn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & \vdots & p_{ni} & \vdots & p_{nn} \end{bmatrix}.$$

Cette matrice a pour entrée p_{ki} en k-eme ligne et i-eme colonne. Elle est donc constituée par les n vecteurs colonnes donnés par l'écriture des vecteurs $e'_1, \dots, e'_n$ dans la base $\mathcal{B}$. C'est un élément de l'espace des matrices carrées de taille n :

$$\mathrm{M}_n(\mathbb{K}) = \{[a_{ij}]_{1 \leqslant i,j \leqslant n}, a_{ij} \in \mathbb{K}, 1 \leqslant i, j \leqslant n\}.$$

En termes de calculs matriciels, l'équation (1) devient simplement :

$$X = P_{\mathcal{B}}^{\mathcal{B}'} X'.$$

Soit maintenant $P' := P_{\mathcal{B}'}^{\mathcal{B}} = [p'_{kj}]_{1 \leqslant, k, j \leqslant n}$ donnée par les relations

$$e_i = \sum_{k=1}^{n} p'_{ki} e'_k.$$

On obtient alors, pour $i \in [\![1, n]\!]$:

$$\begin{aligned} e_i &= \sum_{k=1}^{n} p'_{ki} e'_k \\ &= \sum_{k=1}^{n} p'_{ki} (\sum_{j=1}^{n} p_{jk} e_j) \\ &= \sum_{j=1}^{n} (\sum_{k=1}^{n} p'_{ki} p_{jk}) e_j \end{aligned}$$

et encore une fois par unicité de la décomposition dans la base $\mathcal{B}$ on obtient :

$$\forall i, j \in [\![1, n]\!], \ \sum_{k=1}^{n} p_{jk} p'_{ki} = \delta_{ij}.$$

Cette relation, en termes matriciels, s'écrit

$$PP' = P_{\mathcal{B}}^{\mathcal{B}'} P_{\mathcal{B}'}^{\mathcal{B}} = \mathrm{Id}_n.$$

De même, on vérifie que $P'P = \mathrm{Id}_n$. Autrement dit, la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est l'inverse de la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

3.3. Applications linéaires. Dans cette section, E et F désignent deux espaces vectoriels sur $\mathbb{K}$.

Définition I.3.7. Une application $f : E \to F$ est une application linéaire si pour tout couple de scalaires $(\lambda, \mu) \in \mathbb{K}^2$ et tout couple de vecteurs $(x, y) \in E^2$, on a

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

On note $L_{\mathbb{K}}(E, F)$, ou simplement $L(E, F)$ l'espace de toutes les applications $\mathbb{K}$-linéaires de E vers F. Si $E = F$, on note $L(E) = L(E, E)$ (ou encore $\mathrm{End}_{\mathbb{K}}(E)$, ou $\mathrm{End}(E)$) et ses éléments sont appelés les endomorphismes de E.

Exemples I.3.8. L'application suivante est linéaire :

$$\begin{array}{rccc} \phi : & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x, y, z) & \mapsto & (x + y + z, x - 2z) \end{array}$$

En revanche, l'application

$$\begin{array}{rccc} \psi : & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x, y, z) & \mapsto & (x^2 + y, x - 2z^3) \end{array}$$

ne l'est pas, à cause des quantités quadratique x^2 et cubique z^3.

L'application dérivée est linéaire :

$$\begin{array}{rccc} D : & \mathbb{R}[X] & \to & \mathbb{R}[X] \\ & p(X) & \mapsto & p'(X). \end{array}$$

Soient $\lambda \in \mathbb{K}$ et $(f, g) \in L(E, F)^2$. On peut multiplier par un scalaire une application linéaire en posant pour tout $x \in E$:

$$(\lambda \cdot f)(x) := \lambda \cdot (f(x))$$

et on peut additionner deux applications linéaires en posant pour tout $x \in E$:

$$(f + g)(x) := f(x) + g(x).$$

Proposition I.3.9. *L'espace* $(L_{\mathbb{K}}(E,F),+,\cdot)$ *est un* $\mathbb{K}$*-espace vectoriel. De plus, si* E *est de dimension finie* n *et* F *de dimension finie* m*, alors* $L(E,F)$ *est de dimension finie* mn*.*

Si G est un troisième espace vectoriel sur $\mathbb{K}$, on peut composer les éléments $f \in L(E,F)$ par les éléments $g \in L(F,G)$ pour obtenir $g \circ f \in L(E,G)$, en posant pour tout $x \in E$:

$$g \circ f(x) = g(f(x)).$$

En particulier, si $E = F = G$, on peut itérer cette opération et définir, pour $f \in L(E)$,

$$f^k := f \circ f \circ \ldots \circ f$$

où f apparaît k-fois ($f^1 = f$, $f^2 = f \circ f$, etc). On pose par convention $f^0 = \mathrm{Id}_E$, où Id_E est le morphisme identité donné par :

$$\forall x \in E,\ \mathrm{Id}_E(x) = x.$$

Proposition I.3.10. *L'espace* $(L(E),+,\cdot,\circ)$ *est une* $\mathbb{K}$*-algèbre, non-commutative en général.*

Soit $f \in L(E,F)$.

Définition I.3.11. Le noyau de f, noté $\ker(f)$, est l'image réciproque de 0_F par f, soit

$$\ker(f) = \{x \in E,\ f(x) = 0\}.$$

L'image de f, notée $\mathrm{Im}(f)$, est l'image directe de E par f, soit

$$\mathrm{Im}(f) = f(E) = \{f(x), x \in E\} \subset F.$$

Les ensembles $\ker(f)$ et $\mathrm{Im}(f)$ sont des sous-espaces vectoriels de E et F respectivement.

Proposition I.3.12. *L'application linéaire* f *est injective si et seulement si* $\ker(f) = \{0_E\}$*. Elle est surjective si et seulement si* $\mathrm{Im}(f) = F$*.*

Exercice I.3.13. L'application ϕ de l'Exemple I.3.8 est-elle injective ? Surjective ?

Définition I.3.14. L'application linéaire f est un *isomorphisme* si elle est injective et surjective (i.e. f est linéaire et bijective). On note $\mathrm{GL}_{\mathbb{K}}(E)$ (ou simplement $\mathrm{GL}(E)$) l'ensemble des isomorphismes de E vers E. C'est un groupe pour la composition.

On suppose désormais que E est de dimension finie n. L'image de f est alors nécessairement de dimension finie inférieure à n.

Définition I.3.15. On définit le rang de f comme étant la dimension de son image.

On a alors le théorème du rang :

Théorème I.3.16. *On a l'égalité suivante :*

$$\dim(E) = \dim(\ker(f)) + \mathrm{rang}(f).$$

Corollaire I.3.17. *Soit* $f \in L(E,F)$*.* **On suppose que** $\dim(E) = \dim(F) < +\infty$*. Alors les assertions suivantes sont équivalentes :*
(i) f *est injective,*
(ii) f *est surjective,*
(iii) f *est un isomorphisme,*
(iv) Il existe une base $\mathcal{B}$ *de* E *telle que* $f(\mathcal{B})$ *soit une base de* F*,*
(v) Pour toute famille $\mathcal{B}$ *de* E*,* $\mathcal{B}$ *est une base si et seulement si* $f(\mathcal{B})$ *est une base de* F*.*

On va maintenant faire le lien avec les matrices. On fixe $\mathcal{B}_E = \{e_1, \ldots, e_n\}$ une base de E.

Lemme I.3.18. *Les applications*

$$\begin{array}{rcl} E & \to & \mathbb{K}^n \\ x = \sum_{i=1}^n x_i e_i & \mapsto & (x_1, \ldots, x_n) \end{array}$$

et

$$\begin{array}{rcl} E & \to & \mathrm{M}_{n,1}(\mathbb{K}) \\ x = \sum_{i=1}^n x_i e_i & \mapsto & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{array}$$

sont des isomorphismes de $\mathbb{K}$*-espaces vectoriels.*

On pourra donc identifier E à l'ensemble des vecteurs colonnes de taille n pour faire des calculs en coordonnées, une fois une base $\mathcal{B}_E$ choisie. Nous allons traduire la notion d'application linéaire en termes matricielles. Soient $u \in L(E, F)$ et $\mathcal{B}_F = \{f_1, \ldots, f_m\}$ une base de F (on a donc supposé ici $\dim(F) = m < +\infty$). On va pouvoir déterminer entièrement u à l'aide d'une matrice $A \in \mathrm{M}_{m,n}(\mathbb{K})$, appelée "matrice de u dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$, et notée $A = [u]_{\mathcal{B}_E}^{\mathcal{B}_F}$ (ou $[u]_{\mathcal{B}_E, \mathcal{B}_F}$, et simplement $[u]_{\mathcal{B}_E}$ si $E = F$). Pour cela, on décompose l'image par u de chaque vecteur de la base $\mathcal{B}_E$ dans la base $\mathcal{B}_F$:

$$\forall j \in [\![1, n]\!],\ \exists!\,(a_{ij}) \in \mathbb{K}^m,\ u(e_j) = \sum_{i=1}^m a_{ij} f_i.$$

On pose alors

$$[u]_{\mathcal{B}_E, \mathcal{B}_F} := A = [a_{ij}]_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n} \in \mathrm{M}_{m,n}(\mathbb{K}).$$

Ceci est motivé par les calculs suivants. Soit $x \in E$, de coordonnées $(x_1, \ldots, x_n)$ dans la base $\mathcal{B}_E$. On représente alors x par

$$X = [x]_{\mathcal{B}_E} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

et l'on calcule son image par u :

$$\begin{array}{rcl} u(x) & = & u(\sum_{j=1}^n x_j e_j) \\ & = & \sum_{j=1}^n x_j u(e_j) \\ & = & \sum_{j=1}^n x_j (\sum_{i=1}^m a_{ij} f_i) \\ & = & \sum_{i=1}^m (\sum_{j=1}^n a_{ij} x_j) f_i. \end{array}$$

Par unicité de la décomposition de $y := u(x)$ dans la base $\mathcal{B}_F$, on obtient les coordonnées $(y_1, \ldots, y_m)$ de y dans cette base :

$$\forall i \in [\![1, m]\!],\ y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Si on pose

$$Y = [y]_{\mathcal{B}_F} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

le vecteur de $\mathrm{M}_{m,1}(\mathbb{K})$ qui représente y dans $\mathcal{B}_F$, on a obtenu

$$Y = AX$$

autrement dit

$$[u(x)]_{\mathcal{B}_F} = [u]_{\mathcal{B}_E, \mathcal{B}_F} [x]_{\mathcal{B}_E}. \tag{2}$$

On voit que, modulo le choix de bases, le calcul des images de u se fait via des calculs matriciels. Dans ce dictionnaire matriciel, la composition des applications linéaires (resp. la somme) correspond au produit des matrices (resp. à la somme des matrices).

Théorème I.3.19. *Soient E, F et G trois espaces vectoriels de dimensions finies respectives m, n et p, et de bases respectives $\mathcal{B}_E$, $\mathcal{B}_F$ et $\mathcal{B}_G$.*

(1) L'application

$$\begin{array}{ccc} L_{\mathbb{K}}(E, F) & \to & \mathrm{M}_{m,n}(\mathbb{K}) \\ u & \mapsto & [u]_{\mathcal{B}_E}^{\mathcal{B}_F} \end{array}$$

est un isomorphisme d'espaces vectoriels.

(2) Pour tout $u \in L(E, F)$, $v \in L(F, G)$, on a

$$[v \circ u]_{\mathcal{B}_E}^{\mathcal{B}_G} = [v]_{\mathcal{B}_F}^{\mathcal{B}_G} [u]_{\mathcal{B}_E}^{\mathcal{B}_F}.$$

Dans le cas où $F = E$, on peut toujours composer les applications linéaires, et on obtient :

Théorème I.3.20. *L'application suivante est un isomorphisme d'algèbres*[4] *:*

$$\begin{array}{ccc} L_{\mathbb{K}}(E) & \to & \mathrm{M}_n(\mathbb{K}) \\ u & \mapsto & [u]_{\mathcal{B}_E}. \end{array}$$

Exercice I.3.21. Donner les matrices associées aux applications et bases suivantes :

(1) Dans la base canonique, $u : \mathbb{R}^3 \to \mathbb{R}^3$ donnée par $u(x, y, z) = (x + y + z, x, y + 9z)$.

(2) Dans la base canonique, $v : \mathbb{C}_n[X] \to \mathbb{C}_{n-1}[X]$ donnée par $v(P) = P'$.

(3) Dans la base canonique, avec

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

pour $w : \mathrm{M}_{2,3}(\mathbb{R}) \to \mathrm{M}_{2,3}(\mathbb{R})$ donnée par $w(M) = AM$.

4. C'est à la fois un isomorphisme d'espaces vectoriels et un isomorphisme d'anneaux, i.e. il respecte la structure multiplicative.

(4) Soit $u \in L(E,F)$. Quel est l'effet sur $[u]^{\mathcal{B}_F}_{\mathcal{B}_E}$ d'un changement de base
$$\{e_1,\dots,e_i,\dots e_j,\dots,e_n\} \longrightarrow \{e_1,\dots,e_j,\dots e_i,\dots,e_n\}?$$
Même question si on permute deux éléments de la base de F ?

Exercice I.3.22. Étant donné un espace vectoriel E, on appelle *dual* de E, noté $E^\vee$ (ou E^*), l'espace
$$E^\vee = L(E,\mathbb{K}).$$

(1) Montrer que E est de dimension finie si et seulement si $E^\vee$ est de dimension finie. Montrer que dans ce cas, $\dim(E) = \dim(E^\vee)$. (Indication : on pourra considérer $E \to (E^\vee)^\vee$ défini par $x \mapsto (u \to u(x))$)
(2) Si $\mathcal{B}_E = \{e_1,\dots,e_n\}$ est une base de E, montrer que l'on définit une base de $E^\vee$ par $\mathcal{B}_E^* := \{e_1^*,\dots,e_n^*\}$ donnée par $e_i^*(x) = x_i$ où x_i est la i-eme coordonnée de x dans $\mathcal{B}_E$. C'est la *base duale* de $\mathcal{B}_E$.
(3) Si $u \in L(E,F)$, on définit $u^* : F^\vee \to E^\vee$ par $u^*(f^*) = f^* \circ u$. Montrer que $u^* \in L(F^\vee,E^\vee)$.
(4) Soient $u \in L(E,F)$, $\mathcal{B}_E$ une base de E et $\mathcal{B}_F$ une base de F. Déterminer $[u^*]^{\mathcal{B}_E^*}_{\mathcal{B}_F^*}$ en fonction de $[u]^{\mathcal{B}_F}_{\mathcal{B}_E}$.

On va maintenant donner les "formules de changement de bases" qui explicitent la dépendance des matrices qui représentent des applications linéaires par rapport aux bases. Ces formules permettent en pratique de se ramener à des représentations matricielles plus simples pour faire les calculs, ou encore de montrer que des résultats obtenus en coordonnées ne dépendent pas des bases choisies. Soient E et F comme précédemment, munis de bases $\mathcal{B}_E, \mathcal{B}'_E$ pour E et $\mathcal{B}_F, \mathcal{B}'_F$ pour F. On pose $P = P^{\mathcal{B}'_E}_{\mathcal{B}_E}$ et $Q = P^{\mathcal{B}'_F}_{\mathcal{B}_F}$ les matrices de passages de $\mathcal{B}_E$ à $\mathcal{B}'_E$ et de $\mathcal{B}_F$ à $\mathcal{B}'_F$.

Proposition I.3.23. *Soit $u \in L(E,F)$. Si on pose $A = [u]^{\mathcal{B}_F}_{\mathcal{B}_E}$ et $A' = [u]^{\mathcal{B}'_F}_{\mathcal{B}'_E}$, on a*
$$A' = Q^{-1}AP,$$
autrement dit :
$$[u]^{\mathcal{B}'_F}_{\mathcal{B}'_E} = P^{\mathcal{B}_F}_{\mathcal{B}'_F}[u]^{\mathcal{B}_F}_{\mathcal{B}_E}P^{\mathcal{B}'_E}_{\mathcal{B}_E}. \tag{3}$$

Remarque I.3.24. La formule (3) peut paraître un peu étrange. Pour mieux la retenir, il faut remarquer que
$$P^{\mathcal{B}'_E}_{\mathcal{B}_E} = [\mathrm{Id}_n]^{\mathcal{B}_E}_{\mathcal{B}'_E},$$
c'est à dire que la matrice de changement de base de $\mathcal{B}_E$ vers $\mathcal{B}'_E$ représente l'application identité dans la base $\mathcal{B}'_E$ au départ et $\mathcal{B}_E$ à l'arrivée. La formule (3) devient alors :
$$[u]^{\mathcal{B}'_F}_{\mathcal{B}'_E} = [\mathrm{Id}_m]^{\mathcal{B}'_F}_{\mathcal{B}_F}[u]^{\mathcal{B}_F}_{\mathcal{B}_E}[\mathrm{Id}_n]^{\mathcal{B}_E}_{\mathcal{B}'_E}.$$

Définition I.3.25. On dénote par $\mathrm{GL}_n(\mathbb{K})$ le groupe des matrices de $\mathrm{M}_n(\mathbb{K})$ qui sont inversibles. Ce groupe est appelé le groupe linéaire (sur $\mathbb{K}$).

On rappelle que les matrices de changement de bases sont dans le groupe linéaire, et réciproquement tout élément de $\mathrm{GL}_n(\mathbb{K})$ correspond à un changement de base de $\mathbb{K}^n$.

Définition I.3.26. Deux matrices $A, A' \in \mathrm{M}_{m,n}(\mathbb{K})$ qui vérifient une relation du type
$$A' = Q^{-1}AP$$
pour $Q \in \mathrm{GL}_m(\mathbb{K})$ et $P \in \mathrm{GL}_n(\mathbb{K})$ sont appelées “matrices équivalentes”. Si de plus $m = n$ et $P = Q$, i.e. si
$$A' = P^{-1}AP,$$
on dit que A et A' sont “semblables”.

Deux matrices qui représentent la même application linéaire dans des bases différentes seront donc équivalentes, et réciproquement. L'effet d'un changement de base sur la matrice associée à un endomorphisme est de la remplacer par une matrice semblable, et réciproquement.

Exercice I.3.27. Soit $u : \mathbb{R}^2 \to \mathbb{R}^2$ défini par $u(x, y) = (2x + y, x + 2y)$.

(1) Déterminer la matrice A de u dans la base canonique.
(2) Montrer que $\{v_1, v_2\}$ est une base de $\mathbb{R}^2$, où $v_1 = (1, 1)$ et $v_2 = (1, -1)$.
(3) Déterminer la matrice de u dans $\{v_1, v_2\}$.
(4) En déduire que A est semblable à
$$\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

L'exercice précédent suggère que pour manipuler un endomorphisme, il est utile de pouvoir déterminer une base dans laquelle la forme de la matrice qui le représente est assez simple. C'est ce qu'on appelle la “réduction des endomorphismes”. Voici des exemples de formes assez simples :

Définition I.3.28. Une matrice $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$ est dite diagonale si tous ces coefficients non diagonaux sont nuls :
$$\forall (i, j),\ i \neq j, a_{ij} = 0.$$
On dira que A est triangulaire supérieure si tous les coefficients sous-diagonaux sont nuls :
$$\forall (i, j),\ i > j, a_{ij} = 0.$$
Enfin, on dira que A est triangulaire inférieure si tous les coefficients sur-diagonaux sont nuls :
$$\forall (i, j),\ i < j, a_{ij} = 0.$$

Exemples I.3.29. Les matrices suivantes sont respectivement diagonale :
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$
triangulaire supérieure :
$$\begin{bmatrix} 1 & -2 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$
et triangulaire inférieure :
$$\begin{bmatrix} 1 & 0 & 0 \\ 5 & 2 & 0 \\ 5 & 5 & 3 \end{bmatrix}.$$

3.4. Projecteurs. On termine cette section par des rappels sur une classe particulièrement simple d'application linéaire : les projecteurs. On fera une utilisation répétée de ces derniers dans les prochains chapitres, avec l'idée qu'un projecteur permet de se restreindre à un sous-espace.

DÉFINITION I.3.30. Un endomorphisme $p \in \mathrm{End}_{\mathbb{K}}(E)$ est un projecteur si

$$p \circ p = p.$$

Un projecteur induit une décomposition de E :

PROPOSITION I.3.31. *Soit $p \in \mathrm{End}_{\mathbb{K}}(E)$ un projecteur. Alors*

$$E = \ker(p) \oplus \mathrm{Im}(p).$$

De plus, on peut déterminer $\mathrm{Im}(p)$ *via*

$$\mathrm{Im}(p) = \ker(\mathrm{Id}_E - p),$$

et on a

$$\ker(p) = \mathrm{Im}(\mathrm{Id}_E - p).$$

On parle alors de projecteur sur l'espace $\mathrm{Im}(p)$ parallèlement à $\ker(p)$.

EXERCICE I.3.32. Démontrer la proposition précédente. Il est instructif de faire un dessin.

On peut généraliser cette situation à des décompositions de E en sommes directes de s-sous-espaces. On suppose que E est de dimension finie n et qu'il existe $F_1, \ldots, F_s$ des sous-espaces vectoriels de E tels que :

$$E = F_1 \oplus \ldots \oplus F_s.$$

Pour tout $i \in \{1, \ldots, s\}$, on définit une application linéaire :

$$\begin{array}{rccc} p_i : & E & \to & E \\ & x_1 + \ldots + x_s & \mapsto & x_i, \end{array}$$

où $x_1 + \ldots + x_s$ est l'unique décomposition d'un vecteur de E dans la somme directe $F_1 \oplus \ldots \oplus F_s$. L'endomorphisme p_i est appelée la projection sur F_i (dans la décomposition $E = F_1 \oplus \ldots \oplus F_s$).

PROPOSITION I.3.33. *Les endomorphismes $p_1, \ldots, p_s$ vérifient :*

(1) Pour tout $i \in \{1, \ldots, s\}$, $p_i \circ p_i = p_i$,

(2) Si $i \neq j$, $p_i \circ p_j = 0$,

(3) $\displaystyle\sum_{i=1}^{s} p_i = \mathrm{Id}_E$,

(4) Pour tout $i \in \{1, \ldots, s\}$, $\mathrm{Im}(p_i) = F_i$ et $\ker(p_i) = \displaystyle\bigoplus_{j \neq i} F_j$.

Réciproquement, on a

PROPOSITION I.3.34. *Soient $p'_1, \ldots, p'_s$ des endomorphismes de E qui vérifient*

(i) $\displaystyle\sum_{i=1}^{s} p'_i = \mathrm{Id}_E$,

(ii) pour $i \neq j$, $p'_i \circ p'_j = 0$.

Alors pour tout $i \in \{1,\ldots,s\}$, p'_i est un projecteur sur $F'_i = \mathrm{Im}(p'_i)$ et E est une somme directe :

$$E = F'_1 \oplus \ldots \oplus F'_s.$$

En représentation matricielle, si $\mathcal{B}_i$ est une base de F_i et si $n_i = \dim(F_i)$ pour tout $i \in \{1,\ldots,s\}$, Alors $\mathcal{B} := \bigcup_{i=1}^{s} \mathcal{B}_i$ est une base de E dans laquelle on a pour tout $j \in \{1,\ldots s\}$:

$$[p_j]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \vdots \\ \vdots & 0 & \mathrm{Id}_{n_j} & 0 \\ 0 & \ldots & 0 & 0 \end{bmatrix},$$

où Id_{n_j} apparaît dans la j-eme colonne de cette représentation par blocs.

4. Déterminants

Il est remarquable que la notion de déterminant soit historiquement antérieure à celle d'espace vectoriel. Elle apparaît déjà dans les travaux de Cramer en 1750 dans ses fameuses formules explicites qui donnent les solutions à un système linéaire. Le terme "déterminant" serait quand à lui dû à Cauchy qui le premier démontra certaines de ses propriétés algébriques. Dans ce cours, le déterminant sera essentiellement utilisé pour calculer le polynôme caractéristique d'un endomorphisme et déterminer ses valeurs propres (cf Chapitre II).

4.1. Définitions et premières propriétés. Soit $A = [a_{ij}]_{1\leqslant i,j\leqslant n} \in \mathrm{M}_n(\mathbb{K})$.

DÉFINITION I.4.1. Le déterminant de A, noté $\det(A)$, est l'élément de $\mathbb{K}$ défini par :

$$\det(A) = \sum_{\sigma\in\mathfrak{S}_n} \epsilon(\sigma)\Pi_{i=1}^{n} a_{\sigma(i)i}, \tag{4}$$

où $\epsilon(\sigma)$ désigne la signature[5] de σ et $\mathfrak{S}_n$ est le groupe des permutations de $\{1,\ldots,n\}$.

PROPOSITION I.4.2. *Le déterminant vérifie :*

(1) $\det(\mathrm{Id}_n) = 1$,

(2) Pour $A, B \in \mathrm{M}_n(\mathbb{K})$, $\det(AB) = \det(A)\det(B)$,

(3) $A \in \mathrm{GL}_n(\mathbb{K})$ *si et seulement si* $\det(A) \neq 0$, *et dans ce cas* $\det(A)^{-1} = \det(A^{-1})$,

(4) Si ${}^tA$ *est la transposée*[6] *de* A, $\det(A) = \det({}^tA)$,

(5) Si A *est triangulaire supérieure, triangulaire inférieure, ou diagonale,*

$$\det(A) = \Pi_{i=1}^{n} a_{ii}.$$

Les points (1) – (3) ci-dessus révèlent le fait que

$$\begin{array}{rccc} \det: & \mathrm{GL}_n(\mathbb{K}) & \to & \mathbb{K}^* \\ & A & \mapsto & \det(A) \end{array}$$

5. La signature est obtenu en décomposant σ en produit de transpositions. Si σ est le produit de r transpositions, alors sa signature est $(-1)^r$.

6. La transposée de A est la matrice de coefficients ${}^tA = [a_{ji}]_{1\leqslant i,j\leqslant n}$.

est un morphisme de groupes [7]. Ceci va nous permettre de définir le déterminant d'une application linéaire. En effet, si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, de matrice de passage $P = P_{\mathcal{B}}^{\mathcal{B}'}$, et si $u \in \mathrm{End}_{\mathbb{K}}(E)$, on sait que

$$[u]_{\mathcal{B}'} = P^{-1}[u]_{\mathcal{B}}P,$$

et donc $\det([u]_{\mathcal{B}'}) = \det(P^{-1})\det([u]_{\mathcal{B}})\det(P) = \det([u]_{\mathcal{B}})$. La quantité $\det([u]_{\mathcal{B}})$ ne dépend pas de la base choisie.

DÉFINITION I.4.3. Le déterminant de $u \in \mathrm{End}_{\mathbb{K}}(E)$, noté $\det(u)$, est le scalaire donné par

$$\det(u) = \det([u]_{\mathcal{B}}),$$

où $\mathcal{B}$ est n'importe quelle base de E.

COROLLAIRE I.4.4. *Les propriétés du déterminant des matrices se transportent aux applications linéaires qu'elles représentent. Ainsi*

(1) $\det(\mathrm{Id}_E) = 1$,

(2) Si $u, v \in \mathrm{End}_{\mathbb{K}}(E)$, $\det(u \circ v) = \det(u)\det(v)$,

(3) u est inversible si et seulement si $\det(u) \neq 0$.

On va donner un dernier point de vu sur le déterminant. On a défini les matrices de $\mathrm{M}_n(\mathbb{K})$ comme étant données par n vecteurs colonnes de tailles n. Si $A = [a_{ij}]_{1\leqslant i,j\leqslant n} \in \mathrm{M}_n(\mathbb{K})$, on notera [8] $A = [C_1, \dots, C_n]$ où pour tout $j \in [\![1, n]\!]$, C_j est la j-eme colonne de A :

$$C_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

On peut alors voir le déterminant comme une application

$$\begin{array}{rccc} \det : & \mathrm{M}_{n1}(\mathbb{K}) \times \ldots \times \mathrm{M}_{n1}(\mathbb{K}) & \rightarrow & \mathbb{K} \\ & [C_1, \dots, C_n] & \mapsto & \det([C_1, \dots, C_n]), \end{array}$$

ou encore comme une application de $\mathbb{K}^n \times \ldots \times \mathbb{K}^n$ vers $\mathbb{K}$.

PROPOSITION I.4.5. *Le déterminant, vu comme application sur les colonnes des matrices, vérifie :*

(1) il est n-linéaire : pour tous $(\lambda, \mu) \in \mathbb{K}^2$, $C_1, \dots, C_n \in \mathrm{M}_{n1}(\mathbb{K})^n$, $j \in \{1, \dots, n\}$, *et* $C'_j \in \mathrm{M}_{n1}(\mathbb{K})$, *on a*

$$\det(C_1, \dots, \lambda C_j + \mu C'_j, \dots, C_n) = \lambda \det(C_1, \dots, C_j, \dots, C_n) + \mu \det(C_1, \dots, C'_j, \dots, C_n),$$

(2) il est alterné : pour tous $C_1, \dots, C_n \in \mathrm{M}_{n1}(\mathbb{K})^n$, *si il existe* $i \neq j$ *avec* $C_i = C_j$, *alors*

$$\det([C_1, \dots, C_n]) = 0,$$

(3) il est normalisé : $\det(\mathrm{Id}_n) = 1$.

Par ailleurs, le déterminant est l'unique application de $\mathrm{M}_{n1}(\mathbb{K})^n$ *vers* $\mathbb{K}$ *qui vérifie* $(1) - (3)$.

7. On dénote $\mathbb{K}^*$ le groupe multiplicatif de $\mathbb{K}$, i.e. $\mathbb{K}\backslash\{0\}$.

8. Il y a un léger abus de notation ici, les crochets qui délimitent les C_j ne doivent pas apparaître dans l'expression qui définit A.

REMARQUE I.4.6. On obtient des résultats similaires si l'on travaille avec les lignes des matrices : la valeur du déterminant ne change pas si on ajoute à une ligne une combinaison linéaire des autres lignes. Ceci est très utile en pratique pour calculer les déterminants.

EXERCICE I.4.7. Calculer le déterminant des matrices suivantes en utilisant des opérations sur les lignes et les colonnes, et la propriété (5) de la Proposition I.4.2 :

$$\begin{bmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 3 \\ j & 0 & -2 \\ j^2 & -1 & -1 \end{bmatrix}.$$

Ce nouveau regard sur le déterminant permet de définir le déterminant de n vecteurs d'un espace vectoriel E de dimension n. Bien entendu, comme pour les applications linéaires, il faut fixer une base $\mathcal{B}$ de E pour donner un sens à l'expression. Soient alors $v_1, \dots, v_n$ une famille de n vecteurs de E. On peut lui associer la matrice A dont les colonnes sont les n vecteurs $v_1, \dots, v_n$ écrits dans la base $\mathcal{B}$:

$$A := [[v_1]_{\mathcal{B}}, \dots, [v_n]_{\mathcal{B}}],$$

et on définit alors *le déterminant de* $\{v_1, \dots, v_n\}$ *dans la base* $\mathcal{B}$ par :

$$(\det)_{\mathcal{B}}(v_1, \dots, v_n) := \det(A).$$

Comme pour les vecteurs colonnes, on a :

COROLLAIRE I.4.8. *L'application* $\det_{\mathcal{B}}$ *est l'unique application* $E^n \to \mathbb{K}$ *qui soit* n*-linéaire, alternée, et normalisée par* $\det_{\mathcal{B}}(\mathcal{B}) = 1$.

Malheureusement, contrairement au cas des applications linéaires, ce déterminant dépend de la base choisie. La formule de changement de base est donnée dans la proposition suivante :

PROPOSITION I.4.9. *Soient* $\mathcal{B}$ *et* $\mathcal{B}'$ *deux bases de* E. *On note la matrice de changement de base* $P = P_{\mathcal{B}}^{\mathcal{B}'}$. *Alors*

$$(\det)_{\mathcal{B}} = \det(P)(\det)_{\mathcal{B}'}.$$

Une utilité du déterminant d'une famille de vecteurs est qu'il détecte les bases :

PROPOSITION I.4.10. *Une famille* $\mathcal{F}$ *de* n *vecteurs de* E *est libre si et seulement si c'est une base si et seulement si* $(\det)_{\mathcal{B}}(\mathcal{F}) \neq 0$.

4.2. Calculs de déterminants. On commence par un mot sur la notation : lorsque on calcule le déterminant d'une matrice $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$, on notera :

$$\det(A) = \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & & \vdots & & a_{2n} \\ \vdots & & a_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{vmatrix}.$$

Afin de calculer un déterminant en pratique, on pensera à utiliser le caractère n-linéaire : ajouter à une colonne une combinaison linéaire des autres ne change pas le déterminant. D'autre part, le déterminant étant alterné, permuter deux colonnes change le signe du déterminant de la matrice. Ces remarques sont également valables si on travaille avec les lignes. On peut également calculer le déterminant de manière récursive via les *développements selon les lignes ou les colonnes* :

PROPOSITION I.4.11. *Soit $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$. On a, pour $i \in [\![1, n]\!]$, le développement selon la i-ème ligne :*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

et pour $j \in [\![1, n]\!]$, le développement selon la j-ème colonne :

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

où $A_{ij} \in \mathrm{M}_{n-1}(\mathbb{K})$ est la matrice carrée de taille $n-1$ obtenue en enlevant à A sa i-ème ligne et sa j-ème colonne.

EXERCICE I.4.12. Calculer le déterminant des matrices suivantes en développant selon les lignes ou les colonnes :

$$\begin{bmatrix} 0 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 5 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & -2 \\ 3 & 0 & -1 \end{bmatrix}.$$

Le terme $(-1)^{i+j} \det(A_{ij})$ de la Proposition I.4.11 est appelé le *cofacteur* d'indice (i, j) de A. On va donner au passage une propriété algébrique remarquable des cofacteurs.

DÉFINITION I.4.13. On appelle comatrice de la matrice $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$, la matrice $\mathrm{Com}(A) \in \mathrm{M}_n(\mathbb{K})$ définit par ses coefficients :

$$\mathrm{Com}(A)_{ij} := (-1)^{i+j} \det(A_{ij})$$

où $A_{ij} \in \mathrm{M}_{n-1}(\mathbb{K})$ est la matrice carrée de taille $n-1$ obtenue en enlevant à A sa i-ème ligne et sa j-ème colonne.

On a alors le fait suivant :

THÉORÈME I.4.14. *Pour toute matrice $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$ on a*

$$A \cdot {}^t\mathrm{Com}(A) = {}^t\mathrm{Com}(A) \cdot A = \det(A)\mathrm{Id}_n.$$

Ce résultat n'est pas vraiment utile pour calculer le déterminant, mais plutôt pour les informations qu'il donne sur l'inverse d'une matrice quand celle-ci existe :

COROLLAIRE I.4.15. *Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est inversible si et seulement si $\det(A) \neq 0$, et dans ce cas*

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

Enfin, le calcul par blocs déjà vu dans la Section 3.1 s'étend au calcul de déterminants.

PROPOSITION I.4.16. *Soit*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \in \mathrm{M}_n(\mathbb{K})$$

où $A \in \mathrm{M}_m(\mathbb{K})$ et $D \in \mathrm{M}_{n-m}(\mathbb{K})$. Alors

$$\det(M) = \det(A) \det(D).$$

REMARQUE I.4.17. On laisse au lecteur le soin d'adapter cette proposition au cas d'une matrice triangulaire supérieure par blocs plus générale.

Corollaire I.4.18. *Soit*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \in \mathrm{M}_n(\mathbb{K})$$

où $A \in \mathrm{M}_m(\mathbb{K})$ *et* $D \in \mathrm{M}_{n-m}(\mathbb{K})$. *Alors* M *est inversible si et seulement si* A *et* D *le sont, auquel cas*

$$M^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{bmatrix}.$$

Chapitre II

Vecteurs et valeurs propres, diagonalisation

Dans ce chapitre et les suivants, on fixe E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, $u \in L_{\mathbb{K}}(E) = \mathrm{End}_{\mathbb{K}}(E)$ et $A \in \mathrm{M}_n(\mathbb{K})$. Pour alléger les notations, on identifiera souvent $\mathbb{K}^n$ à $\mathbb{M}_{n,1}(\mathbb{K})$ à l'aide de l'isomorphisme

$$\begin{array}{ccc} \mathbb{K}^n & \to & \mathrm{M}_{n,1}(\mathbb{K}) \\ x = (x_1, \ldots, x_n) & \mapsto & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \end{array}$$

On va maintenant rentrer dans le vif du sujet : étant donné un endomorphisme $u \in \mathrm{End}_{\mathbb{K}}(E)$, est-il possible de déterminer une base de E dans laquelle la matrice associée à u soit la plus simple possible ? Idéalement, on cherche une base $\mathcal{B}$ dans laquelle u soit diagonale, c'est à dire telle qu'il existe $\lambda_1, \ldots, \lambda_n$ tels que

$$[u]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & 0 & & 0 \\ \vdots & 0 & \lambda_i & 0 & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{bmatrix}. \tag{5}$$

En effet, dans une telle base, en coordonnées, u se lira simplement

$$(x_1, \ldots, x_n) \mapsto (\lambda_1 x_1, \ldots, \lambda_n x_n).$$

Notez que demander que les λ_i soient tous égaux à 1 pour simplifier encore plus l'écriture de u n'est pas raisonnable, car dans ce cas on aurait $[u]_{\mathcal{B}} = \mathrm{Id}_n$, et par formule de changement de base, ceci serait vrai pour toute base, et u serait en fait égale à Id_E. En revanche, on verra que si $\mathbb{K} = \mathbb{C}$, demander l'existence d'une telle forme diagonale (5) est très raisonnable : en un certain sens, la plupart des endomorphismes d'un espace vectoriel complexe admettent une base dans laquelle leur représentation est diagonale. On dit alors d'un tel endomorphisme qu'il est *diagonalisable*, et il s'agit de déterminer une telle base $\mathcal{B} = \{e_1, \ldots, e_n\}$, ainsi que les valeurs λ_i. Notez dans ce cas que pour tout $i \in \{1, \ldots n\}$,

$$u(e_i) = \lambda_i e_i.$$

C'est cette équation en $\lambda \in \mathbb{K}$ et $x \in E$ qui sera notre point de départ :

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

1. Éléments propres d'un endomorphisme

La notion de valeur propre (elle aussi historiquement antérieure à celle d'espace vectoriel) est d'abord apparue à la fin du 18e siècle dans les travaux de Lagrange, Laplace ou Bernoulli sur les systèmes d'équations différentielles.

Définition II.1.1. Soit $\lambda \in \mathbb{K}$. On dit que λ est une valeur propre de u (resp. de A) s'il existe un vecteur non nul $x \neq 0$ dans E (resp. $X \in \mathrm{M}_{n,1}(\mathbb{K})\backslash\{0\}$) tel que

$$u(x) = \lambda x \text{ (resp. } AX = \lambda X).$$

L'ensemble des valeurs propres de u (resp. A) est appelé spectre de u (resp. spectre de A) et noté $\mathrm{Sp}_{\mathbb{K}}(u)$ (resp. $\mathrm{Sp}_{\mathbb{K}}(A)$).

Remarque II.1.2. On remarque que si $A = [u]_{\mathcal{B}}$ pour une base $\mathcal{B}$ de E, l'équation (2) implique que λ est valeur propre de u si et seulement si c'est une valeur propre de A. On a donc dans ce cas $\mathrm{Sp}(u) = \mathrm{Sp}(A)$.

Exemple II.1.3. Le spectre d'une matrice diagonale est facile à déterminer. Par exemple, soit

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \in \mathrm{M}_3(\mathbb{K}),$$

i.e. $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}$. Alors $\mathrm{Sp}_{\mathbb{K}}(D) = \{\lambda_1, \ldots, \lambda_3\}$. En effet, si (e_1, e_2, e_3) désigne la base canonique de $\mathbb{K}^3$, $De_i = \lambda_i e_i$ pour $i \in \{1,2,3\}$ et donc $\lambda_i \in \mathrm{Sp}_{\mathbb{K}}(D)$ pour tout $i \in \{1,2,3\}$. On verra par ailleurs que le cardinal du spectre est inférieur à la dimension de E, ce qui permet de conclure ici.

Exemple II.1.4. On va voir dans cet exemple que le spectre d'un endomorphisme peut être vide et qu'il dépend du corps de base dans lequel on considère l'endomorphisme. Soit $u : \mathbb{R}^2 \to \mathbb{R}^2$ défini par sa matrice dans la base canonique de $\mathbb{R}^2$:

$$A = [u]_{\mathcal{B}_{can}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Supposons que $\lambda \in \mathbb{R}$ soit valeur propre de u. On a alors l'existence de $x = (x_1, x_2) \in \mathbb{R}^2\backslash\{(0,0)\}$ tel que $u(x) = \lambda x$. On a alors l'équation matricielle $AX = \lambda X$, où $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$:

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

De manière équivalente, (x_1, x_2) vérifie le système

$$\begin{cases} x_2 &= \lambda x_1 \\ -x_1 &= \lambda x_2 \end{cases}.$$

Si $\lambda = 0$, ce système impose $x_1 = x_2 = 0$, ce qui est impossible par définition d'une valeur propre, et si $\lambda \neq 0$ alors nécessairement on obtient $\lambda^2 = -1$ qui n'a pas de solution sur $\mathbb{R}$, d'où $\mathrm{Sp}_{\mathbb{R}}(u) = \varnothing$. En revanche, si on considère $u : \mathbb{C}^2 \to \mathbb{C}^2$ défini par la même matrice A dans la base canonique de $\mathbb{C}^2$ cette fois-ci (i.e. ici $\mathbb{K} = \mathbb{C}$), on obtient

$$\mathrm{Sp}_{\mathbb{C}}(u) = \{i, -i\}.$$

En effet, l'équation $u(x) = ix$ admet la solution $(1, i)$ et l'équation $u(x) = -ix$ admet la solution $(1, -i)$.

EXERCICE II.1.5. Montrer que u est non injective si et seulement si $0 \in \mathrm{Sp}_{\mathbb{K}}(u)$.

DÉFINITION II.1.6. Soit $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ (resp. $\lambda \in \mathrm{Sp}_{\mathbb{K}}(A)$).

(1) Le sous-espace propre associé à λ est le sous-espace vectoriel de E (resp. de $\mathbb{K}^n$) :

$$E_\lambda := \ker(u - \lambda \mathrm{Id}_E) = \{x \in E,\ u(x) = \lambda x\}$$

(resp.

$$\ker(A - \lambda \mathrm{Id}_n) = \{X \in \mathbb{K}^n,\ AX = \lambda X\}).$$

(2) Les éléments de $\ker(u - \lambda \mathrm{Id}_n)$ (resp. de $\ker(A - \lambda \mathrm{Id}_n)$) sont appelés les vecteurs propres associés à λ.

REMARQUE II.1.7. Un vecteur propre peut être nul. On a donc

$$\lambda \in \mathrm{Sp}_{\mathbb{K}}(u) \Leftrightarrow \ker(u - \lambda \mathrm{Id}_E) \neq \{0\}$$

et

$$\lambda \in \mathrm{Sp}_{\mathbb{K}}(A) \Leftrightarrow \ker(A - \lambda \mathrm{Id}_n) \neq \{0\}.$$

EXEMPLES II.1.8. Dans l'exemple II.1.3, en supposant les λ_i deux à deux distincts, on a

$$\ker(D - \lambda_i \mathrm{Id}_3) = \mathrm{Vect}\,(e_i).$$

Dans l'exemple II.1.4, sur le corps $\mathbb{C}$, on a $\mathrm{Sp}_{\mathbb{C}}(u) = \{i, -i\}$, et un calcul donne

$$E_i = \mathrm{Vect}_{\mathbb{C}}(e_1 + ie_2)\ \text{, et } E_{-i} = \mathrm{Vect}_{\mathbb{C}}(e_1 - ie_2).$$

On remarque que $\mathcal{B} = \{e_1 + ie_2, e_1 - ie_2\}$ est une base de $\mathbb{C}^2$ dans laquelle $[u]_{\mathcal{B}}$ est diagonale

$$[u]_{\mathcal{B}} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

et $\mathbb{C}^2$ se décompose :

$$\mathbb{C}^2 = E_i \oplus E_{-i}.$$

Le résultat suivant est trivial mais important :

LEMME II.1.9. *Si $A = [u]_{\mathcal{B}}$ pour une base $\mathcal{B}$ de E, alors pour tout $\lambda \in \mathrm{Sp}_K(u) = \mathrm{Sp}_K(A)$, l'application*

$$\begin{array}{ccc} \ker(u - \lambda \mathrm{Id}_E) & \to & \ker(A - \lambda \mathrm{Id}_n) \\ x & \mapsto & [x]_{\mathcal{B}} \end{array}$$

est un isomorphisme.

Étudier les valeurs propres et espaces propres de u revient donc à étudier ceux de $[u]_{\mathcal{B}}$ pour une base $\mathcal{B}$ de E.

Les sous-espaces propres d'un endomorphisme ont le bon goût d'être en somme directe :

THÉORÈME II.1.10. *Soient $\lambda_1, \ldots, \lambda_s \in \mathrm{Sp}_{\mathbb{K}}(u)$ des valeurs propres de u deux à deux distinctes. Les sous-espaces propres associés $E_{\lambda_i} = \ker(u - \lambda_i \mathrm{Id}_E)$ sont en somme directe dans E, i.e.*

$$\sum_{i=1}^{s} \ker(u - \lambda_i \mathrm{Id}_E) = \bigoplus_{i=1}^{s} \ker(u - \lambda_i \mathrm{Id}_E) \subset E.$$

On a bien sûr un énoncé analogue pour les espaces-propres associés au valeurs propres de $A \in \mathrm{M}_n(\mathbb{K})$.

Démonstration. On va montrer l'énoncé par récurrence sur le nombre de valeurs propres distinctes s. Le résultat est directe si $s = 1$. On suppose alors qu'il est vrai pour $s - 1 \in \mathbb{N}^*$ valeurs propres distinctes et on va le démontrer pour s valeurs propres distinctes. Soit $(v_1, \ldots, v_s) \in E_{\lambda_1} \times \ldots \times E_{\lambda_s}$ tel que

$$v_1 + \ldots + v_s = 0. \tag{6}$$

Il s'agit de montrer que les v_i sont tous nuls. On applique u à l'équation (6) pour obtenir (u est linéaire et les $v_i \in E_{\lambda_i}$) :

$$\lambda_1 v_1 + \ldots + \lambda_s v_s = 0. \tag{7}$$

On multiplie également (6) par λ_s et on a :

$$\lambda_s v_1 + \ldots + \lambda_s v_s = 0. \tag{8}$$

On soustrait alors (8) à (7) :

$$(\lambda_1 - \lambda_s)v_1 + \ldots + (\lambda_{s-1} - \lambda_s)v_{s-1} = 0.$$

Or pour tout $i \in \{1, \ldots, s-1\}$, $(\lambda_i - \lambda_s)v_i \in E_{\lambda_i}$ (les E_{λ_i} sont des sous-espaces vectoriels de E), donc par hypothèse de récurrence,

$$\forall i \in \{1, \ldots, s-1\},\ (\lambda_i - \lambda_s)v_i = 0.$$

Cependant, les λ_i sont supposées deux à deux distinctes, on a donc, pour $i \in \{1, \ldots, s-1\}$, $\lambda_i - \lambda_{s-1} \neq 0$ et $v_i = 0$. Finalement, avec (6) on conclut $v_s = 0$. $\square$

Corollaire II.1.11. *Sous les mêmes hypothèses que dans le Théorème* II.1.10*, on a*

$$\sum_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \dim_{\mathbb{K}}(\ker(u - \lambda \mathrm{Id}_E)) = \dim_{\mathbb{K}}(\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E)) \leqslant \dim_{\mathbb{K}}(E),$$

avec égalité si et seulement si

$$\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E) = E.$$

Ainsi, une fois connu le spectre d'un endomorphisme, les espaces propres associés à ses valeurs propres nous fournissent un sous-espace vectoriel de E sur lequel u admet une forme particulièrement simple :

Exercice II.1.12. Soit $F \subset E$ un sous-espace vectoriel de E *stable par* u, i.e. tel que $u(F) \subset F$. On appelle *restriction de* u *à* F lâendomorphisme $u_{|F} \in L(F)$ défini, pour $x \in F$, par $u_{|F}(x) = u(x)$.

(1) Soit $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$. Montrer que $u(E_\lambda) \subset E_\lambda$.
(2) En déduire que $\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E)$ est stable par u.
(3) On fixe une base $\mathcal{B}_\lambda$ de E_λ. Montrer que la restriction de u à E_λ a pour matrice $\lambda \mathrm{Id}_{n_\lambda}$, avec $n_\lambda = \dim(E_\lambda)$.
(4) En déduire la matrice de la restriction de u à $\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} E_\lambda$ dans une base de la forme $\bigcup_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \mathcal{B}_\lambda$.

Les espaces propres peuvent être calculés via la résolution de systèmes d'équations linéaires. Il reste alors à déterminer les valeurs propres. Elles seront obtenues comme racines d'un polynôme.

2. Polynôme caractéristique

Déterminer si λ est une valeur propre de u revient à savoir si l'équation $u(x) - \lambda x = 0$ a une solution non nulle en x, ce qui est équivalent au fait que $u - \lambda \mathrm{Id}_E$ soit non injective, i.e. non bijective car E est de dimension finie. Ceci motive la définition suivante :

DÉFINITION II.2.1. Soit T une indéterminée sur $\mathbb{K}$. On appelle polynôme caractéristique de A le polynôme [1]

$$\chi_A(T) = \det(T \cdot \mathrm{Id}_n - A) \in \mathbb{K}[T].$$

REMARQUE II.2.2. Dans le chapitre I on a vu les notions de matrices à coefficients dans un corps et de déterminants de ces matrices. On peut étendre sans difficulté ces notions à des matrices à coefficients dans un anneau commutatif (e.g. $\mathbb{K}[T]$). Le déterminant d'une telle matrice est défini par la même formule (4). Il est alors $\mathbb{K}[T]$-n-linéaire par rapport à ses colonnes (ou ses lignes), alterné, vérifie les formules de développement selon les lignes et les colonnes (Proposition I.4.11), ainsi que le résultat du Théorème I.4.14. Les preuves de ces faits sont exactement les mêmes que dans le cas des matrices de $\mathrm{M}_n(\mathbb{K})$.

Notez que si $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K})$, alors

$$\begin{aligned} T\mathrm{Id}_n - A &= \begin{bmatrix} T & 0 & \dots & 0 \\ 0 & T & 0 & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \dots & 0 & T \end{bmatrix} - \begin{bmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix} \\ &= \begin{bmatrix} T - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & T - a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ -a_{n1} & \dots & \dots & T - a_{nn} \end{bmatrix}, \end{aligned}$$

ce qui est une matrice de $\mathrm{M}_n(\mathbb{K}[T])$, et $\chi_A(T)$, son déterminant, est bien un élément de $\mathbb{K}[T]$.

EXEMPLE II.2.3. Si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{M}_2(\mathbb{C}),$$

alors

$$\chi_A(T) = T^2 - (a + d)T + (ad - bc).$$

Le développement du polynôme caractéristique fait intervenir des quantités bien connues [2] :

PROPOSITION II.2.4. *On a*

$$\chi_A(T) = T^n - \mathrm{tr}(A)T^{n-1} + \ldots + (-1)^n \det(A).$$

DÉMONSTRATION. Le coefficients d'ordre 0 de $\chi_A(T)$ est obtenu en évaluant $\chi_A(T)$ en 0, ce qui donne bien $(-1)^n \det(A)$. Pour les termes d'ordre n et $n - 1$, un développement selon la première colonne donne :

$$\chi_A(T) = (T - a_{11}) \det(T \cdot \mathrm{Id}_n - A_{11}) + R(T),$$

1. Le polynôme caractéristique est parfois défini comme étant $\det(A - T\mathrm{Id}_n)$.

2. On rappelle que la trace d'une matrice (resp. d'un endomorphisme), notée $\mathrm{tr}(A)$, est le scalaire donné par la somme de ses éléments diagonaux (resp. la somme des éléments diagonaux de n'importe quelle matrice qui le représente.)

où $R(T)$ est un polynôme de degré inférieur ou égal à $n-2$. Par récurrence sur n, on obtient alors que le coefficient de degré n de χ_A est 1 et celui de degré $(n-1)$ est $-a_{11} - \mathrm{tr}(A_{11}) = -\mathrm{tr}(A)$. □

Remarque II.2.5. Ainsi, pour $n = 2$, on a simplement

$$\chi_A(T) = T^2 - \mathrm{tr}(A)T + \det(A).$$

La proposition suivante va nous permettre d'étendre la définition de déterminant aux endomorphismes :

Proposition II.2.6. *Soient A et A' deux matrices semblables de $\mathrm{M}_n(\mathbb{K})$ (i.e. il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A' = P^{-1}AP$). Alors*

$$\chi_A(T) = \chi'_A(T).$$

Démonstration. Par définition, et avec les propriétés du déterminant (morphisme de groupes multiplicatifs) :

$$\begin{array}{rcc} \chi_{A'}(T) & = & \det(T\mathrm{Id}_n - A') \\ & = & \det(T\mathrm{Id}_n - P^{-1}AP) \\ & = & \det(TP^{-1}P - P^{-1}AP) \\ & = & \det(P^{-1}(T\mathrm{Id}_n - A)P) \\ & = & \det(P^{-1})\det(T\mathrm{Id}_n - A)\det(P) \\ & = & \det(T\mathrm{Id}_n - A) \\ & = & \chi_A(T). \end{array}$$

□

Ainsi, deux matrices qui représentent le même endomorphisme dans deux bases différentes ont le même polynôme caractéristique. La définition suivante a donc un sens :

Définition II.2.7. On définit le polynôme caractéristique de u, noté $\chi_u(T)$, comme étant le polynôme de $\mathbb{K}[T]$ donné par

$$\chi_u(T) = \chi_{[u]_{\mathcal{B}}}(T)$$

pour n'importe quelle base $\mathcal{B}$ de E.

Remarque II.2.8. On utilisera parfois la notation $\det(X\mathrm{Id}_E - u)$. Cela dit il faudra être prudent avec le sens donné à l'objet “$X \cdot \mathrm{Id}_E$”.

On a maintenant une caractérisation des valeurs propres en termes de racines du polynôme caractéristique :

Théorème II.2.9. *Soit $\lambda \in \mathbb{K}$. Les assertions suivantes sont équivalentes[3] :*
(i) λ est valeur propre de u, i.e. $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$,
(ii) $u - \lambda\mathrm{Id}_E$ est non injectif,
(iii) $u - \lambda\mathrm{Id}_E$ est non bijectif,
(iv) λ est racine de χ_u, i.e. $\chi_u(\lambda) = 0$.

Remarque II.2.10. Le même résultat est valable pour une matrice $A \in \mathrm{M}_n(\mathbb{K})$, on a l'équivalence entre :

3. Notons au passage que les points (ii) et (iii) ne sont pas équivalents en dimension infinie, où les notions de valeurs propres ($(u - \lambda\mathrm{Id}_E)$ non injectif) et de spectre (ensemble des scalaires pour lesquels $u - \lambda\mathrm{Id}_E$ est non inversible) sont différentes.

(i') $\lambda \in \mathrm{Sp}_{\mathbb{K}}(A)$,
(ii') $A - \lambda \mathrm{Id}_n$ est non inversible,
(iii') $\chi_A(\lambda) = 0$.
De plus, si $A = [u]_{\mathcal{B}}$ pour une base $\mathcal{B}$ de E, alors les assertions du Théorème II.2.9 sont équivalentes à $(i') - (iii')$.

Démonstration. C'est assez direct :

$$\begin{aligned} \lambda \in \mathrm{Sp}_{\mathbb{K}}(u) &\iff \exists x \in E,\ x \neq 0,\ u(x) - \lambda x = 0 \\ &\iff u - \lambda \mathrm{Id}_E \text{ non injectif} \\ &\iff u - \lambda \mathrm{Id}_E \text{ non bijectif} \\ &\iff \det(u - \lambda \mathrm{Id}_E) = 0. \end{aligned}$$

□

La conséquence suivante est bien entendu valable pour $A \in \mathrm{M}_n(\mathbb{K})$ également :

Corollaire II.2.11. *Le cardinal de* $\mathrm{Sp}_{\mathbb{K}}(u)$ *est inférieur ou égal à* n :

$$\sharp \mathrm{Sp}_{\mathbb{K}}(u) \leqslant \dim_{\mathbb{K}}(E).$$

Démonstration. Le polynôme χ_u est de degré n, donc il possède au plus n racines [4]. □

Exemple II.2.12. Les valeurs propres d'une matrice triangulaire supérieure sont données par ses éléments diagonaux.

On dispose alors d'un programme pour déterminer les valeurs propres et espaces propres d'un endomorphisme u (ou d'une matrice A) :

(1) Calculer son polynôme caractéristique via un calcul de déterminant,
(2) Factoriser ce polynôme sur $\mathbb{K}$ et en déduire ses racines (et donc le spectre de u ou A),
(3) Pour chacune des racines, déterminer l'espace propre $\ker(u - \lambda \mathrm{Id}_E)$ associé via la résolution d'un système d'équations linéaires.

Exercice II.2.13. Soit

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

(1) Montrer que $\chi_A(T) = T^3 - 1$.
(2) En déduire que $\mathrm{Sp}_{\mathbb{R}}(A) = \{1\}$ et que $\mathrm{Sp}_{\mathbb{C}}(A) = \{1, j, j^2\}$.
(3) Déterminer dans les deux cas les espaces propres de A.

On remarque ici que l'influence du corps sur lequel on travaille sur le spectre de l'endomorphisme se traduit par la possibilité de décomposer ou non le polynôme caractéristique en produit de facteurs de degré 1.

Définition II.2.14. Un polynôme $p(T) \in \mathbb{K}[T]$ est dit scindé sur $\mathbb{K}$ s'il se décompose sur $\mathbb{K}$ comme produit de facteurs de degré un, i.e. il existe $a \in \mathbb{K}$ et $(\alpha_1, \ldots, \alpha_s) \in \mathbb{K}^s$ tels que [5]

$$p(T) = a \Pi_{i=1}^{s} (T - \alpha_i).$$

4. Ce fait, vu en L1 pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, s'étend de la même manière à tout corps. Ceci repose sur l'existence d'une division euclidienne sur $\mathbb{K}[X]$ et le fait que $p(a) = 0$ si et seulement si $(X - a)$ divise p.

5. On peut démontrer que tout corps $\mathbb{K}$ admet une "clôture algébrique", c'est à dire un corps $\mathbb{L}$ (minimal pour l'inclusion) qui contienne $\mathbb{K}$ et dans lequel tous les polynômes soient scindés. Dans le cas de $\mathbb{R}$, il s'agit de $\mathbb{C}$.

Exemples II.2.15. Le polynôme $X^2 + 1$ n'est pas scindé sur $\mathbb{R}$ mais il est scindé sur $\mathbb{C}$. Le polynôme $(X - 2)^2$ est scindé sur $\mathbb{R}$.

Définition II.2.16. On dira que l'endomorphisme u (resp. la matrice A) est scindé (resp. scindée) si son polynôme caractéristique l'est.

Remarque II.2.17. L'endomorphisme u est scindé sur $\mathbb{K}$ si et seulement si

$$\chi_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T - \lambda)^{n_\lambda}$$

où n_λ est la multiplicité de λ dans χ_u, notée $n_\lambda = \mathrm{mult}(\lambda, \chi_u)$.

Exemple II.2.18. Si A est triangulaire supérieure (ou inférieure), alors A est scindée.

Lemme II.2.19. *Si $\mathbb{K} = \mathbb{C}$, tout endomorphisme est scindé. Si $\mathbb{K} = \mathbb{R}$, $u \in \mathrm{End}_{\mathbb{K}}(E)$ est scindé si et seulement si χ_u n'a que des racines réelles.*

Démonstration. Le corps $\mathbb{C}$ est algébriquement clos donc tout polynôme sur $\mathbb{C}$ est scindé. $\square$

Le polynôme caractéristique contient également une information importante sur les espaces propres d'un endomorphisme (ou d'une matrice) :

Proposition II.2.20. *Soit $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$. Alors*

$$1 \leqslant \dim(E_\lambda) \leqslant \mathrm{mult}(\lambda, \chi_u).$$

Démonstration. Soit $\mathcal{B}_\lambda$ une base de E_λ que l'on complète en une base $\mathcal{B}$ de E. Dans cette base, la matrice de u est diagonale par blocs :

$$[u]_{\mathcal{B}} = \begin{bmatrix} \lambda \mathrm{Id}_{n_\lambda} & B \\ 0 & C \end{bmatrix}$$

où l'on a posé $n_\lambda = \dim(E_\lambda)$ et où $C \in \mathrm{M}_{n-n_\lambda}(\mathbb{K})$. On en déduit que

$$\chi_u(T) = (T - \lambda)^{n_\lambda} \cdot \chi_C(T),$$

et donc $n_\lambda \leqslant \mathrm{mult}(\lambda, \chi_u)$.

$\square$

Un corollaire très utile dans la recherche des espaces propres est le suivant :

Corollaire II.2.21. *Si $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ est de multiplicité 1 dans le polynôme χ_u, i.e. $\mathrm{mult}(\lambda, \chi_u) = 1$, alors $\dim_{\mathbb{K}}(\ker(u - \lambda \mathrm{Id}_E)) = 1$.*

Ainsi, pour calculer un espace propre associé à une valeur propre de multiplicité 1, il suffit de déterminer un vecteur propre de cet espace.

Exemples II.2.22. Soient A et B les matrices de $\mathrm{M}_2(\mathbb{R})$:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

En dimension $n = 2$, la formule

$$\chi_M(T) = T^2 - \mathrm{tr}(M)T + \det(M)$$

donne directement

$$\chi_A(T) = T^2 - 4T + 3 \text{ et } \chi_B(T) = T^2 - 2T + 1$$

que l'on factorise :
$$\chi_A(T) = (T-3)(T-1) \text{ et } \chi_B(T) = (T-1)^2.$$
Pour A, $\mathrm{Sp}_{\mathbb{R}}(A) = \{1,3\}$, et $\mathrm{mult}(1,\chi_A) = \mathrm{mult}(3,\chi_A) = 1$, d'où $\dim_{\mathbb{R}}(\ker(A-\mathrm{Id}_2)) = \dim_{\mathbb{R}}(\ker(A-3\mathrm{Id}_2)) = 1$. On a donc
$$\mathbb{R}^2 = \ker(A-\mathrm{Id}_2) \oplus \ker(A-3\mathrm{Id}_2),$$
et un calcul direct donne $\ker(A-\mathrm{Id}_2) = \mathrm{Vect}\,\{(1,-1)\}$ et $\ker(A-3\mathrm{Id}_2) = \mathrm{Vect}\,\{(1,1)\}$. Pour B en revanche, $\mathrm{Sp}_{\mathbb{R}}(B) = \{1\}$, et $\mathrm{mult}(1,\chi_B) = 2$. Un calcul montre par ailleurs que $\ker(B-\mathrm{Id}_2) = \mathrm{Vect}\,(e_1)$, et $\dim(\ker(B-\mathrm{Id}_2)) < \mathrm{mult}(1,\chi_B)$. Dans cet exemple, on a
$$\bigoplus_{\lambda\in\mathrm{Sp}_{\mathbb{R}}(B)} \ker(B-\lambda\mathrm{Id}_2) \subsetneq \mathbb{R}^2,$$
même si B est scindé sur $\mathbb{R}$.

On termine par cette proposition utile pour calculer les polynômes caractéristiques des matrices triangulaires par blocs. La preuve est laissée en exercice au lecteur.

Proposition II.2.23. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ une matrice triangulaire supérieure par blocs :*
$$A = \begin{bmatrix} A_{11} & \dots & A_{1i} & \dots & A_{1r} \\ 0 & \ddots & & & A_{2r} \\ \vdots & 0 & A_{ii} & & \vdots \\ \vdots & & 0 & \ddots & \vdots \\ 0 & \dots & \dots & 0 & A_{rr} \end{bmatrix}.$$
Alors

(1) Le polynôme caractéristique de A se décompose :
$$\chi_A(T) = \Pi_{i=1}^r \chi_{A_{ii}}(T),$$
(2) Le spectre vérifie :
$$\mathrm{Sp}_{\mathbb{K}}(A) = \bigcup_{i=1}^r \mathrm{Sp}_{\mathbb{K}}(A_{ii}),$$
(3) A est scindée sur $\mathbb{K}$ si et seulement si pour tout $i \in \{1,\dots r\}$, A_{ii} est scindée sur $\mathbb{K}$.

3. Diagonalisation

On donne dans cette section deux caractérisations importantes des endomorphismes *diagonalisables.*

Définition II.3.1. Un endomorphisme $u \in \mathrm{End}_{\mathbb{K}}(E)$ est diagonalisable s'il existe une base $\mathcal{B}$ de E dans laquelle $[u]_{\mathcal{B}}$ est diagonale. Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est diagonalisable si elle est semblable à une matrice diagonale, i.e. il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit diagonale.

Remarque II.3.2. On verra souvent des exemples d'endomorphismes définis par des matrices à coefficients réels. On pourra voir ces applications comme endomorphismes de $\mathbb{R}^n$ ou bien de $\mathbb{C}^n$. On parlera alors d'endomorphismes diagonalisables "sur $\mathbb{R}$" ou "sur $\mathbb{C}$". La distinction sera cruciale car certains endomorphismes ne sont pas scindés sur $\mathbb{R}$.

Une première caractérisation des endomorphismes diagonalisables est la suivante :

Théorème II.3.3. *Soit* $u \in \mathrm{End}_{\mathbb{K}}(E)$. *Les assertions suivantes sont équivalentes :*
(i) u est diagonalisable,
(ii) $[u]_{\mathcal{B}}$ *est diagonalisable pour une base* $\mathcal{B}$ *de* E*,*
(iii) $[u]_{\mathcal{B}}$ *est diagonalisable pour toute base* $\mathcal{B}$ *de* E*,*
(iv) E *est somme directe des espaces propres de* u *:*

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E),$$

(v) E *possède une base formée de vecteurs propres de* u*.*

Démonstration. (i), (ii) et (iii) sont équivalents en vertu du fait que si $\mathcal{B}, \mathcal{B}'$ sont deux bases de E de matrice de passage $P = P_{\mathcal{B}}^{\mathcal{B}'}$, alors $[u]_{\mathcal{B}'} = P^{-1}[u]_{\mathcal{B}}P$ et réciproquement si $P \in \mathrm{GL}_n(\mathbb{K})$, alors P est une matrice de passage de $\mathcal{B}$ à une autre base de E [6].

Pour $(iii) \Rightarrow (iv)$, soit $\mathcal{B}$ une base de E dans laquelle $[u]_{\mathcal{B}}$ soit diagonale, égale à

$$[u]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & 0 & & 0 \\ \vdots & 0 & \lambda_i & 0 & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{bmatrix}.$$

Il est alors clair que les vecteurs de $\mathcal{B}$ sont des vecteurs propres de u. On a alors pour tout $e \in \mathcal{B}$,

$$e \in \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E)$$

et donc

$$E = \mathrm{Vect}\,(\mathcal{B}) \subset \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E).$$

Pour $(iv) \Rightarrow (v)$, il suffit de considérer une base $\mathcal{B}_\lambda$ pour chaque espace propre de u. La réunion $\mathcal{B} = \bigcup_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \mathcal{B}_\lambda$ est une base de E formée de vecteurs propres.

Enfin, $(v) \Rightarrow (i)$ est claire, car dans une telle base $\mathcal{B}$ formée de vecteurs propres, la matrice $[u]_{\mathcal{B}}$ est diagonale. □

Dans la suite, si $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$, on utilisera la notation

$$\mathrm{Diag}(\lambda_1, \dots, \lambda_n) := \begin{bmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & 0 & & 0 \\ \vdots & 0 & \lambda_i & 0 & \vdots \\ \vdots & & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{bmatrix}.$$

Proposition II.3.4. *Si un endomorphisme* $u \in \mathrm{End}_{\mathbb{K}}(E)$ *diagonalisable, alors il est scindé sur* $\mathbb{K}$*. La réciproque est fausse.*

6. via la formule $e'_j = \sum_{i=1}^n p_{ij} e_i$.

DÉMONSTRATION. Si u est diagonalisable, on peut fixer une base $\mathcal{B}$ dans laquelle $[u]_{\mathcal{B}}$ soit diagonale, et il existe $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$, telles que

$$[u]_{\mathcal{B}} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n).$$

D'où

$$\chi_u(T) = \Pi_{i=1}^n (T - \lambda_i)$$

est scindé. Le contre-exemple

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

prouve que la réciproque est fausse. En effet, $\chi_A(T) = T^2$ et donc $\mathrm{Sp}_{\mathbb{K}}(A) = \{0\}$, mais on a clairement $\mathbb{K}^2 \neq \ker(A)$ (ce qui serait le cas si A était diagonalisable). □

Du Théorème II.3.3 on déduit un critère efficace pour déterminer si un endomorphisme est diagonalisable :

COROLLAIRE II.3.5. *Si $u \in \mathrm{End}_{\mathbb{K}}(E)$ (resp. $A \in \mathrm{M}_n(\mathbb{K})$) possède n valeurs propres distinctes, alors u (resp. A) est diagonalisable.*

Autrement dit, si χ_u (resp. χ_A), admet n racines distinctes sur $\mathbb{K}$, alors u (resp. A) est diagonalisable.

DÉMONSTRATION. En effet, supposons $\mathrm{Sp}_{\mathbb{K}}(u) = \{\lambda_1, \ldots, \lambda_n\}$, avec $\lambda_i \neq \lambda_j$ pour $i \neq j$. Comme

$$\bigoplus_{i=1}^n \ker(u - \lambda_i \mathrm{Id}_E) \subset E,$$

on a

$$\sum_{i=1}^n \dim(\ker(u - \lambda_i \mathrm{Id}_E)) \leqslant n.$$

Or par définition des espaces propres, pour tout $i \in [\![1, n]\!]$,

$$1 \leqslant \dim(\ker(u - \lambda_i \mathrm{Id}_E)),$$

et donc

$$n \leqslant \sum_{i=1}^n \dim(\ker(u - \lambda_i \mathrm{Id}_E)) \leqslant n.$$

D'où

$$n = \dim(\bigoplus_{i=1}^n \ker(u - \lambda_i \mathrm{Id}_E))$$

et

$$\bigoplus_{i=1}^n \ker(u - \lambda_i \mathrm{Id}_E) = E.$$

□

REMARQUE II.3.6. La réciproque est fausse, comme le montre l'exemple $u = \mathrm{Id}_E$ si $n \geqslant 2$.

Ce critère se généralise de la façon suivante :

THÉORÈME II.3.7. *Soit $u \in \mathrm{End}_{\mathbb{K}}(E)$. Alors u est diagonalisable sur $\mathbb{K}$ si et seulement si*

$$\chi_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} (T - \lambda)^{\dim_{\mathbb{K}}(E_\lambda)},$$

c'est à dire si et seulement si les deux critères suivants sont vérifiés :

(i) $\chi_u(T)$ *est scindé sur* $\mathbb{K}$*, et*

(ii) $\forall \lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$*, on a* $\dim_{\mathbb{K}}(E_\lambda) = \mathrm{mult}_{\mathbb{K}}(\lambda, \chi_u)$.

De même, une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est diagonalisable si et seulement si χ_A est scindé sur $\mathbb{K}$ et pour tout élément $\lambda \in \mathrm{Sp}_{\mathbb{K}}(A)$, $\dim_{\mathbb{K}}(\ker(A - \lambda \mathrm{Id}_n)) = \mathrm{mult}(\lambda, \chi_A)$.

Démonstration. Supposons que u soit diagonalisable. On note $\mathrm{Sp}_{\mathbb{K}}(u) = \{\lambda_1, \ldots, \lambda_s\}$ le spectre de u. On a alors

$$E = \bigoplus_{i=1}^{s} E_{\lambda_i}$$

et si pour tout $i \in \{1, \ldots, s\}$ on fixe une base (de vecteurs propres) $\mathcal{B}_i$ de E_{λ_i}, on obtient une base $\mathcal{B} = \bigcup_{i=1}^{s} \mathcal{B}_i$ de E dans laquelle

$$[u]_{\mathcal{B}} = \mathrm{Diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1 \text{ fois}}, \underbrace{\lambda_2, \ldots, \lambda_2}_{n_2 \text{ fois}}, \ldots, \underbrace{\lambda_s, \ldots, \lambda_s}_{n_s \text{ fois}})$$

où $n_i = \dim(E_{\lambda_i})$. On calcul le polynôme caractéristique de cette matrice

$$\chi_u(T) = \Pi_{i=1}^{s}(T - \lambda_i)^{n_i}$$

et on a bien $n_i = \mathrm{mult}(\lambda_i, \chi_u)$.

Réciproquement, supposons que le polynôme caractéristique vérifie

$$\chi_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T - \lambda)^{\dim_{\mathbb{K}}(E_\lambda)}.$$

Ce polynôme est de degré $n = \dim_{\mathbb{K}}(E)$, donc

$$\sum_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \dim_{\mathbb{K}}(E_\lambda) = \dim_{\mathbb{K}}(E).$$

Ceci implique

$$\dim_{\mathbb{K}}(\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} E_\lambda) = \dim_{\mathbb{K}}(E)$$

et donc

$$\bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} E_\lambda = E,$$

i.e. u est diagonalisable. □

Notez au passage le résultat suivant qui provient du fait que le polynôme caractéristique ne dépend pas de la base choisie pour le calcul :

Proposition II.3.8. *Soit* $u \in \mathrm{End}_{\mathbb{K}}(E)$ *un endomorphisme diagonalisable de polynôme caractéristique*

$$\chi_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T - \lambda)^{n_\lambda},$$

où $n_\lambda = \dim(E_\lambda)$*. Si* $\mathcal{B}$ *est une base de* E *dans laquelle* $[u]_{\mathcal{B}}$ *est diagonale,*

$$D = [u]_{\mathcal{B}} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$$

alors nécessairement les λ_i *sont les valeurs propres de* u*, chaque valeur propre* $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ *apparaissant* n_λ*-fois dans* D*.*

C'est à dire que si u est diagonalisable, il existe un nombre fini de matrices diagonales qui représentent u. Ces différentes formes sont obtenues en permutant les éléments diagonaux d'une forme diagonale donnée. Attention cependant, une même forme diagonale peut être obtenue dans un nombre infini de bases distinctes (ce sera même toujours le cas si $\mathbb{K}$ est infini). Matriciellement, pour tout $A \in \mathrm{M}_n(\mathbb{K})$, l'ensemble

$$\{P^{-1}AP \text{ diagonale }, P \in \mathrm{GL}_n(\mathbb{K})\}$$

est fini (il peut être vide), mais l'ensemble

$$\{P \in \mathrm{GL}_n(\mathbb{K}),\ P^{-1}AP = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)\}$$

peut être infini s'il est non vide.

Exemple II.3.9. Posons

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & -3 & 2 \\ -2 & 1 & -4 \end{bmatrix} \in \mathrm{M}_3(\mathbb{R}),$$

que l'on voit comme un endomorphisme u de $\mathbb{R}^3$ dans la base canonique. On calcul (par exemple en utilisant les transformations sur les lignes et les colonnes $L_3 \to L_3 + L_2$ puis $C_3 \to C_3 - C_2$, puis en développant selon la dernière ligne)

$$\chi_A(T) = (T+1)(T+2)^2.$$

Le spectre de A est $\mathrm{Sp}_{\mathbb{R}}(A) = \{-1, -2\}$. On s'intéresse d'abord à E_{-2} pour déterminer si A est diagonalisable :

$$\begin{aligned} \ker(A + 2\mathrm{Id}_3) &= \left\{ X \in \mathbb{R}^3, \begin{bmatrix} 4 & -2 & 4 \\ 2 & -1 & 2 \\ -2 & 1 & -2 \end{bmatrix} \cdot X = 0 \right\} \\ &= \left\{ (x,y,z) \in \mathbb{R}^3,\ 2x - y + 2z = 0 \right\}. \end{aligned}$$

Donc $\dim_{\mathbb{R}}(E_{-2}) = 2 = \mathrm{mult}_{\mathbb{R}}(-2, \chi_A)$. On a également $\mathrm{mult}_{\mathbb{R}}(-1, \chi_A) = 1$ et $\dim_{\mathbb{R}}(E_{-1}) = 1$. En conclusion, A est diagonalisable. Une base de E_{-2} est donnée par $\{v_1 = (1, 0, -1), v_2 = (1, 2, 0)\}$ et une base de E_{-1} par $\{v_3 = (-2, -1, 1)\}$. Dans la base $\mathcal{B} = \{v_1, v_2, v_3\}$ on a :

$$[u]_{\mathcal{B}} = P^{-1}AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

où

$$P = P_{\mathcal{B}_{can}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & -2 \\ 0 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

4. Trigonalisation

Comme on l'a vu sur certains exemples, un endomorphisme n'est pas nécessairement diagonalisable. On cherche alors une autre forme matricielle simple, mais moins restrictive, pour effectuer les calculs : une forme triangulaire (supérieure ou inférieure).

Définition II.4.1. On dit que u (resp. A) est trigonalisable sur $\mathbb{K}$ s'il existe une base $\mathcal{B}$ de E (resp. $P \in \mathrm{GL}_n(\mathbb{K})$) telle que $[u]_{\mathcal{B}}$ (resp. $P^{-1}AP$) soit triangulaire supérieure.

Remarque II.4.2. Par définition, u est trigonalisable si et seulement si pour une (et donc pour toute) base $\mathcal{B}$, $[u]_{\mathcal{B}}$ est trigonalisable.

Lemme II.4.3. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ une matrice triangulaire inférieure. Alors A est semblable à une matrice triangulaire supérieure.*

Démonstration. On pose $A = [a_{ij}]_{1 \leqslant i,j \leqslant n}$. Pour $\mathcal{B}_{can} = \{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{K}^n$ on a $A = [f]_{\mathcal{B}_{can}}$, avec pour tout $j \in \{1, \ldots, n\}$:

$$f(e_j) = \sum_{i=j}^{n} a_{ij} e_j.$$

Dans la base $\mathcal{B} = \{e_n, \ldots, e_1\}$, l'endomorphisme f vérifie :

$$[f]_{\mathcal{B}} = \begin{bmatrix} a_{nn} & a_{n,n-1} & \cdots & a_{n1} \\ 0 & a_{n-1,n-1} & & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & a_{11} \end{bmatrix}.$$

□

Remarque II.4.4. Par le lemme qui précède, on peut remplacer triangulaire supérieure par triangulaire inférieure dans la Définition II.4.1.

Lemme II.4.5. *Soit $\mathcal{B} = \{v_1, \ldots, v_n\}$ une base de E. On pose, pour $j \in [\![1, n]\!]$,*

$$V_j := \operatorname{Vect}\{v_1, \ldots, v_j\} = \mathbb{K}v_1 \oplus \ldots \oplus \mathbb{K}v_j.$$

La matrice $[u]_{\mathcal{B}}$ est triangulaire supérieure si et seulement si pour tout $j \in [\![1, n]\!]$, $u(V_j) \subset V_j$.

Démonstration. Il suffit de remarquer que $u(V_j) \subset V_j$ est équivalent à $u(v_i) \in \mathbb{K}v_1 \oplus \ldots \oplus \mathbb{K}v_j$ pour tout $i \in [\![1, j]\!]$. La traduction matricielle donne le résultat. □

On a alors une caractérisation des endomorphismes trigonalisables :

Théorème II.4.6. *Un endomorphisme $u \in \mathrm{End}_{\mathbb{K}}(E)$ (resp. une matrice $A \in \mathrm{M}_n(\mathbb{K})$) est trigonalisable si et seulement si χ_u (resp. χ_A) est scindé sur $\mathbb{K}$.*

Un corollaire immédiat :

Corollaire II.4.7. *Sur $\mathbb{C}$, tous les endomorphismes sont trigonalisables. Sur $\mathbb{R}$, un endomorphisme est trigonalisable si et seulement si toutes ses valeurs propres sont réelles*[7].

Démonstration du Théorème II.4.6. Supposons $A \in \mathrm{M}_n(\mathbb{K})$ trigonalisable. Il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A' := P^{-1}AP$ soit triangulaire supérieure, et on a $\chi_A(T) = \chi_{A'}(T)$. La forme triangulaire supérieure de A' permet de conclure que $\chi_{A'}$, et donc χ_A, est scindé.

Réciproquement, supposons χ_u scindé. On va démontrer par récurrence sur $n = \dim_{\mathbb{K}}(E)$ que u est trigonalisable. Si $n = 1$ c'est clair. Supposons le résultat vrai au rang $n - 1 \in \mathbb{N}^*$, et montrons le au rang n. Comme χ_u est scindé et de degré $n \geqslant 2$, il admet une racine $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$. On a donc l'existence d'un vecteur propre non nul $e_\lambda \in E_\lambda$. Soit alors $F \subset E$ un

7. De manière générale, quitte à passer à une clôture algébrique de $\mathbb{K}$, on pourra toujours trigonaliser une matrice.

supplémentaire de Vect $\{e_\lambda\}$ dans E. Si $\mathcal{B}_F$ est une base de F, la matrice de u dans la base $\mathcal{B} = \{e_\lambda\} \cup \mathcal{B}_F$ est de la forme :

$$[u]_{\mathcal{B}} = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}$$

où $B \in \mathrm{M}_{n-1}(\mathbb{K})$. On a alors $\chi_u(T) = (T-\lambda)\chi_B(T)$ et donc B est scindé car $\chi_u(T)$ l'est. Par hypothèse de récurrence, la matrice B est trigonalisable : il existe $P_B \in \mathrm{GL}_{n-1}(\mathbb{K})$ telle que $B' := P_B^{-1}BP_B$ soit triangulaire supérieure. On défini alors par blocs la matrice :

$$P := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P_B & \\ 0 & & & \end{bmatrix}.$$

On a

$$P^{-1} := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P_B^{-1} & \\ 0 & & & \end{bmatrix}$$

et un calcul par blocs donne

$$P^{-1}[u]_{\mathcal{B}}P = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & P_B^{-1}BP_B & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{bmatrix}$$

qui est triangulaire supérieure, donc u est trigonalisable. □

Exercice II.4.8. Soit

$$A = \frac{1}{3}\begin{bmatrix} 2 & -1 & -4 \\ -3 & 0 & 0 \\ -2 & -2 & 1 \end{bmatrix}$$

(1) Déterminer le polynôme caractéristique de A.

(2) En déduire que $\mathrm{Sp}_{\mathbb{R}}(A) = \{-1, 1\}$.

(3) Pourquoi A est-elle trigonalisable sur $\mathbb{R}$?

(4) Étudier E_1 et montrer que A n'est pas diagonalisable.

(5) Trigonaliser A.

Solution II.4.9. On pose

$$B = \begin{bmatrix} 2 & -1 & -4 \\ -3 & 0 & 0 \\ -2 & -2 & 1 \end{bmatrix}$$

et on calcule :

$$\begin{aligned}
\chi_A(T) &= \det(T \cdot \mathrm{Id}_3 - A) \\
&= \tfrac{1}{3^3} \det(3T \cdot \mathrm{Id}_3 - B) \\
&= \tfrac{1}{3^3} \begin{vmatrix} 3T-2 & 1 & 4 \\ 3 & 3T & 0 \\ 2 & 2 & 3T-1 \end{vmatrix} \\
&\underset{L_1 \leftarrow L_1+L_2+L_3}{=} \tfrac{1}{3^3} \begin{vmatrix} 3T+3 & 3T+3 & 3T+3 \\ 3 & 3T & 0 \\ 2 & 2 & 3T-1 \end{vmatrix} \\
&= \tfrac{T+1}{3} \begin{vmatrix} 1 & 1 & 1 \\ 1 & T & 0 \\ 2 & 2 & 3T-1 \end{vmatrix} \\
&\underset{L_3 \leftarrow L_3-2L_1}{=} \tfrac{T+1}{3} \begin{vmatrix} 1 & 1 & 1 \\ 1 & T & 0 \\ 0 & 0 & 3T-3 \end{vmatrix} \\
&= (T+1)(T-1)^2.
\end{aligned}$$

On en déduit la réponse aux questions $(1) - (3)$: le spectre est l'ensemble des racines du polynômes caractéristique, A est trigonalisable car χ_A est scindé sur $\mathbb{R}$.

Pour E_1, on résout $(x, y, z) \in \ker(A - \mathrm{Id}_3)$ et on obtient

$$E_1 = \{(x, y, z) \in \mathbb{R}^3,\ x + y = z = 0\} = \mathrm{Vect}\,\{(1, -1, 0)\}.$$

En particulier, $\dim(E_1) = 1 < \mathrm{mult}(1, \chi_A)$ et donc A n'est pas diagonalisable.

Pour trigonaliser A on commence par déterminer une base de E_{-1}. Comme $\mathrm{mult}(-1, \chi_1) = 1$, $\dim_{\mathbb{R}}(E_{-1}) = 1$ et il suffit de déterminer un vecteur directeur de E_{-1} :

$$E_{-1} = \mathrm{Vect}_{\mathbb{R}}\{(1, 1, 1)\}.$$

On pose alors $v_1 = (1, 1, 1)$, $v_2 = (1, -1, 0)$ et on cherche v_3 de telle sorte que $\mathcal{B} = \{v_1, v_2, v_3\}$ soit une base de $\mathbb{R}^3$. Si $P = P_{\mathcal{B}_{can}}^{\mathcal{B}}$, on a

$$P^{-1}AP = \begin{bmatrix} -1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & c \end{bmatrix}.$$

En effet, si f est l'endomorphisme de $\mathbb{R}^3$ représenté par A dans $\mathcal{B}_{can}$, $P^{-1}AP$ n'est autre que la matrice de f dans $\mathcal{B}$. De plus, $\chi_{P^{-1}AP} = \chi_A$ et donc nécessairement $c = 1$. On prend v_3 le plus simple possible, i.e. $v_3 = (0, 0, 1)$. On calcul alors a et b via l'équation :

$$f(v_3) = av_1 + bv_2 + v_3.$$

On écrit v_1, v_2 et v_3 dans $\mathcal{B}_{can}$, ce qui donne le système :

$$\left\{\begin{array}{ccccccc} -\frac{4}{3} & = & a & + & b & & \\ 0 & = & a & - & b & & \\ \frac{1}{3} & = & a & & & + & 1 \end{array}\right.$$

On trouve $a = b = -\frac{2}{3}$ et dans la base $\mathcal{B}$ on a

$$[f]_{\mathcal{B}} = P^{-1}AP = \begin{bmatrix} -1 & 0 & -\frac{2}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix}.$$

Chapitre III

Polynômes d'endomorphismes

Le polynôme caractéristique χ_u joue un rôle clé dans la réduction d'un endomorphisme u. Une des raisons est qu'il est un *polynôme annulateur* de u : par le théorème de Cayley-Hamilton, $\chi_u(u) = 0$. Dans ce chapitre, on va donner un sens à la notion de polynôme en u, et établir un lien entre la structure d'anneau de $\mathbb{K}[T]$ et les décompositions de E et de ses sous-espaces en sommes directes. On verra apparaître la notion de polynôme minimal, un diviseur de χ_u dont la factorisation donnera un nouveau critère pour déterminer si un endomorphisme est diagonalisable.

1. L'algèbre $\mathbb{K}[u]$ et le théorème de Cayley-Hamilton

On commence par introduire les algèbres $\mathbb{K}[u]$ et $\mathbb{K}[A]$ pour donner un sens à $\chi_A(A)$, avant de démontrer le théorème de Cayley-Hamilton.

1.1. Polynômes d'endomorphismes. Soit $p(T) \in \mathbb{K}[T]$, défini par ses coefficients :

$$p(T) = \sum_{i=0}^{d} a_i T^i,$$

où d est le degré de p. On rappelle que pour $i \in \mathbb{N}^*$,

$$A^i = \underbrace{A \cdot A \dots A}_{i \text{ fois}}$$

et de même :

$$u^i = \underbrace{u \circ u \circ \dots \circ u}_{i \text{ fois}},$$

avec les conventions

$$A^0 = \mathrm{Id}_n$$

et

$$u^0 = \mathrm{Id}_E.$$

Définition III.1.1. On définit le polynôme d'endomorphisme :

$$p(u) := \sum_{i=0}^{d} a_i u^i \in \mathrm{End}_{\mathbb{K}}(E),$$

et le polynôme de matrice :

$$\sum_{i=0}^{d} a_i A^i \in \mathrm{M}_n(\mathbb{K}).$$

On laisse la preuve de la proposition suivante en exercice :

Proposition III.1.2. *L'application*

$$\begin{array}{rccc} \Phi_u : & \mathbb{K}[T] & \to & \mathrm{End}_{\mathbb{K}}(E) \\ & p(T) & \mapsto & p(u) \end{array}$$

est un morphisme d'algèbres, i.e.

(i) Φ_u *est une application* $\mathbb{K}$*-linéaire,*

(ii) Pour tout $(p, q) \in \mathbb{K}[T]^2$, $\Phi_u(pq) = \Phi_u(p) \circ \Phi_u(q)$.

Remarque III.1.3. On a de même que l'application suivante est un morphisme d'algèbres :

$$\begin{array}{rccc} \Phi_A : & \mathbb{K}[T] & \to & \mathrm{M}_n(\mathbb{K}) \\ & p(T) & \mapsto & p(A). \end{array}$$

Définition III.1.4. On appelle l'image de Φ_u, notée $\mathbb{K}[U]$ (resp. de Φ_A, notée $\mathbb{K}[A]$) l'espace des polynômes en u (resp. en A). C'est une $\mathbb{K}$-algèbre, et si $A = [u]_{\mathcal{B}}$ pour une base $\mathcal{B}$ de E,

$$\begin{array}{ccc} \mathbb{K}[u] & \to & \mathbb{K}[A] \\ p(u) & \mapsto & p(A) \end{array}$$

est un isomorphisme.

La proposition suivante sera utilisée à plusieurs reprises (on dira que deux polynômes en u (ou en A) commutent) :

Proposition III.1.5. *L'algèbre* $\mathbb{K}[u]$ *(resp.* $\mathbb{K}[A]$*) est commutative, i.e.*

$$\forall (u_1, u_2) \in \mathrm{Im}(\Phi_u),\ u_1 \circ u_2 = u_2 \circ u_1.$$

Démonstration. En effet, cela découle de la commutativité de $\mathbb{K}[T]$. Par définition il existe $p_1, p_2 \in \mathbb{K}[T]$ tels que $u_1 = p_1(u)$ et $u_2 = p_2(u)$. On a alors

$$\begin{array}{rcl} u_1 \circ u_2 & = & p_1(u) \circ p_2(u) \\ & = & \Phi_u(p_1) \circ \Phi_u(p_2) \\ & = & \Phi_u(p_1 p_2) \\ & = & \Phi_u(p_2 p_1) \\ & = & \Phi_u(p_2) \circ \Phi_u(p_1) \\ & = & u_2 \circ u_1. \end{array}$$

□

1.2. Théorème de Cayley-Hamilton. Ce théorème fondamental fut d'abord obtenu par Cayley, qui se contenta d'en donner une preuve en dimension 2 et 3 (!).

Théorème III.1.6 (Théorème de Cayley-Hamilton). *Le polynôme caractéristique est un polynôme annulateur, i.e. pour toute matrice* $A \in \mathrm{M}_n(\mathbb{K})$ *(resp. tout endomorphisme* $u \in \mathrm{End}_{\mathbb{K}}(E)$*),* $\chi_A(A) = 0$ *(resp.* $\chi_u(u) = 0$*).*

Remarque III.1.7. On prendra garde à la signification de $\chi_A(A)$. Comme $\chi_A \in \mathbb{K}[T]$, $\chi_A(A)$ est le polynôme en A donné par $\Phi_A(\chi_A)$. Par exemple, si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{M}_2(\mathbb{K}),$$

alors

$$\chi_A(T) = T^2 - (a + d)T + (ad - bc)$$

et donc
$$\chi_A(A) = A^2 - (a+d)A + (ad - bc)\mathrm{Id}_2 \in \mathrm{M}_2(\mathbb{K}).$$
On ne peut pas remplacer T par A dans
$$\begin{vmatrix} T-a & -b \\ -c & T-d \end{vmatrix}$$
car l'expression
$$\begin{bmatrix} A-a & -b \\ -c & A-d \end{bmatrix}$$
n'a pas de sens.

Démonstration. On part de la relation
$$(T\mathrm{Id}_n - A) \cdot {}^t\mathrm{Com}(T\mathrm{Id}_n - A) = \chi_A(T)\mathrm{Id}_n. \tag{9}$$
Notons que la matrice $T \cdot \mathrm{Id}_n - A$ ne fait intervenir que des polynômes de degrés au plus 1, c'est un élément de $\mathrm{M}_n(\mathbb{K}_1[T])$, tandis que la matrice $\mathrm{Com}(T\mathrm{Id}_n - A)$ fait intervenir des polynômes de degré au plus $n-1$. On développe ${}^t\mathrm{Com}(T\mathrm{Id}_n - A) \in \mathrm{M}_n(\mathbb{K}_{n-1}[T])$ selon le degré : il existe $B_0, \ldots, B_{n-1} \in \mathrm{M}_n(\mathbb{K})$ telles que
$${}^t\mathrm{Com}(T\mathrm{Id}_n - A) = \sum_{i=0}^{n-1} B_i T^i.$$
De même on développe $\chi_A(T)\mathrm{Id}_n$: il existe $b_0, \ldots, b_n \in \mathbb{K}^{n+1}$ tels que
$$\chi_A(T) = \sum_{i=0}^{n} b_i T^i.$$
On a alors avec l'équation (9) :
$$(T \cdot \mathrm{Id}_n - A)(B_{n-1} \cdot T^{n-1} + \ldots + B_0) = b_n\mathrm{Id}_n \cdot T^n + \ldots + b_0\mathrm{Id}_n.$$
On développe le terme de gauche de cette égalité :
$$B_{n-1} \cdot T^n + \sum_{i=1}^{n-1} (B_{n-1-i} - A \cdot B_{n-i}) \cdot T^{n-i} - A \cdot B_0 = b_n\mathrm{Id}_n \cdot T^n + \ldots + b_0\mathrm{Id}_n.$$
On identifie les coefficients en T^i, pour $i \in \{0, \ldots, n\}$, ce qui donne :
$$\left\{\begin{array}{ccc} B_{n-1} & = & b_n\mathrm{Id}_n \\ B_{n-2} - A \cdot B_{n-1} & = & b_{n-1}\mathrm{Id}_n, \\ \vdots & = & \vdots \\ B_0 - A \cdot B_1 & = & b_1\mathrm{Id}_n, \\ -A \cdot B_0 & = & b_0\mathrm{Id}_n. \end{array}\right. \tag{10}$$
On procède alors de la façon suivante : par récurrence sur $i \geqslant 1$, on fait l'opération $L_i \leftarrow L_i + A \cdot L_{i-1}$ sur les lignes du système (10). Pour $i = n+1$ on obtient :
$$b_n A^n + b_{n-1}A^{n-1} + \ldots + b_1 A + b_0\mathrm{Id}_n = 0.$$
Comme les (b_i) sont les coefficients de $\chi_A(T)$, le résultat est démontré. □

Corollaire III.1.8. *Soit $A \in \mathrm{M}_n(\mathbb{K})$ une matrice inversible. Alors son inverse A^{-1} est un polynôme en A.*

DÉMONSTRATION. Posons

$$\chi_A(T) = T^n + a_{n-1}T^{n-1} + \ldots + a_1 T + a_0.$$

Par le théorème de Cayley-Hamilton,

$$A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 \mathrm{Id}_n = 0$$

donc

$$A \cdot (A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 \mathrm{Id}_n) = -a_0 \mathrm{Id}_n.$$

On rappelle que $a_0 = (-1)^n \det(A) \neq 0$ car A est inversible, d'où

$$-\frac{1}{a_0} A \cdot (A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 \mathrm{Id}_n) = \mathrm{Id}_n$$

et

$$\begin{array}{rcl} A^{-1} & = & -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 \mathrm{Id}_n) \\ & = & \Phi_A(-\frac{1}{a_0}(T^{n-1} + a_{n-1}T^{n-2} + \ldots + a_1)). \end{array}$$

□

EXERCICE III.1.9. Soit

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

(1) Calculer χ_A.

(2) En déduire A^2 et A^{-1}.

DEVOIR MAISON III.1.10. On va donner une autre démonstration du théorème de Cayley-Hamilton. Soit $A \in \mathrm{M}_n(\mathbb{K})$. On suppose ici que A est scindée (c'est par exemple le cas si $\mathbb{K} = \mathbb{C}$). On souhaite montrer que $\chi_A(A) = 0$.

(1) Montrer que si $B = P^{-1}AP$ est une matrice semblable à A, alors

$$\chi_B(B) = P^{-1}\chi_A(A)P.$$

(2) En déduire que l'on peut supposer A triangulaire supérieure pour prouver le théorème.

(3) On suppose donc A triangulaire supérieure. Soit $\mathcal{B}_{can} = \{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{K}^n$.

  (a) Montrer que pour tout $i \in \{2, \ldots, n\}$ on a $(A - a_{ii}\mathrm{Id}_n)(e_i) \in \mathbb{K}e_1 \oplus \ldots \oplus \mathbb{K}e_{i-1}$.

  (b) Montrer que pour tout $i, j \in \{1, \ldots, n\}^2$,

$$(A - a_{ii}\mathrm{Id}_n)(A - a_{jj}\mathrm{Id}_n) = (A - a_{jj}\mathrm{Id}_n)(A - a_{ii}\mathrm{Id}_n).$$

  (c) Déduire de (a) et (b), par récurrence sur j :

$$\forall k \in \{1, \ldots, j\},\ (A - a_{11}\mathrm{Id}_n)(A - a_{22}\mathrm{Id}_n)\ldots(A - a_{jj}\mathrm{Id}_n)(e_k) = 0.$$

(4) Montrer que

$$\chi_A(A) = (A - a_{11}\mathrm{Id}_n)(A - a_{22}\mathrm{Id}_n)\ldots(A - a_{nn}\mathrm{Id}_n).$$

(5) Conclure.

2. Polynôme minimal

On va faire des rappels sur certaines propriétés de l'anneau $\mathbb{K}[T]$. Ces propriétés vont se traduire par l'existence d'un polynôme minimal μ_u tel que tout polynôme annulateur de u soit un multiple de μ_u. On verra alors que les valeurs propres sont également les racines de μ_u, et que sa factorisation est intimement liée à la diagonalisation de u.

2.1. Quelques propriétés de l'anneau $\mathbb{K}[T]$. L'anneau $\mathbb{K}[T]$ possède de nombreuses propriétés algébriques en commun avec $\mathbb{Z}$.

DÉFINITION III.2.1. Soient $(a, b) \in \mathbb{K}[T]$. On dit que b divise a (ou que a est divisible par b) s'il existe $q \in \mathbb{K}[T]$ tel que $a = bq$.

EXEMPLE III.2.2. Le polynôme T divise $T^3 + T$.

De manière plus générale, il existe sur $\mathbb{K}[T]$ une *division euclidienne*[1](On renvoit à [**RW1**] pour une démonstration) :

THÉORÈME III.2.3. *Soient* $(a, b) \in \mathbb{K}[T]^2$, $b \neq 0$. *Il existe un unique couple* $(q, r) \in \mathbb{K}[T]^2$ *avec* $\deg(r) < \deg(b)$ *et tel que*

$$a = bq + r.$$

Avec les notations du théorème précédent, le polynôme q est appelé le *quotient* et r le *reste* de la division euclidienne de a par b. On rappelle au passage que par convention, $\deg(0) = -\infty$.

DÉFINITION III.2.4. Soients $p_1, \dots, p_r$ des éléments non tous nuls de $\mathbb{K}[T]$. On appelle pgcd (plus grand commun diviseur) des $(p_i)_{1 \leqslant i \leqslant r}$, noté *pgcd*$(p_1, \dots, p_r)$, un polynôme p qui vérifie :
(i) Pour tout $i \in \{1, \dots, r\}$, p divise p_i,
(ii) Si $q \in \mathbb{K}[T]$ divise chacun des p_i, $1 \leqslant i \leqslant r$, alors q divise p

PROPOSITION III.2.5. *Soient* $p_1, \dots, p_r$ *des éléments non tous nuls de* $\mathbb{K}[T]$. *Le pgcd des* $(p_i)_{1 \leqslant i \leqslant r}$ *existe et est unique à un multiple de* $\mathbb{K}^*$ *près.*

DÉMONSTRATION. On va utiliser des techniques issues de la théorie des anneaux principaux. Soit

$$I := \{\sum_{i=1}^{r} a_i p_i,\ a_i \in \mathbb{K}[T],\ i = 1 \dots r\}.$$

L'ensemble $I \subset \mathbb{K}[T]$ est un "idéal" de $\mathbb{K}[T]$, c'est à dire qu'il vérifie[2] :
(i) I est un sous-groupe additif de $(\mathbb{K}[T], +)$,
(ii) Si $a \in \mathbb{K}[T]$ et $p \in I$, alors $ap \in I$.
Le fait que $\mathbb{K}[T]$ soit euclidien implique que tout idéal de $\mathbb{K}[T]$ est principal, autrement dit, il existe $p \in \mathbb{K}[T]$ tel que

$$I = (p) := \{ap, a \in \mathbb{K}[T]\}.$$

Pour démontrer cela, on considère l'ensemble des degrés des éléments non nuls de I :

$$\mathcal{D} := \{\deg(q),\ q \in I \backslash \{0\}\}.$$

1. Ceci fait de $\mathbb{K}[T]$, comme $\mathbb{Z}$, un anneau "principal". La définition du polynôme minimal peut être revue dans ce contexte plus abstrait : c'est le polynôme unitaire générateur de l'idéal $\ker(\Phi_u)$. Le lecteur intéressé pourra consulter [**RW1**, **RW2**].

2. On laisse au lecteur le soin de vérifier ces points.

Alors $\mathcal{D}$ est une partie (non vide car les p_i sont supposés non tous nuls) de $\mathbb{N}$ et admet donc un plus petit élément $n \in \mathcal{D}$. On fixe alors $p \in I\backslash\{0\}$ avec $\deg(p) = n$. Montrons comme annoncé que

$$I = (p).$$

Il est clair que $(p) \subset I$, car $p \in I$ et par l'axiome (ii) des idéaux. Réciproquement, si $b \in I\backslash\{0\}$, on effectue la division euclidienne de b par p :

$$b = bq + r$$

où $(q, r) \in \mathbb{K}[T]^2$ avec $\deg(r) < \deg(p)$. Mais alors $r = b - pq \in I$ par les axiomes (i) et (ii). Ceci contredit la minimalité du degré de p, sauf si $r = 0$. D'où p divise b et $b \in (p)$. On a donc $I = (p)$.

On va montrer que $p = pgcd(p_1, \dots, p_r)$. Tout d'abord, pour tout $i \in \{1, \dots, r\}$, $p_i \in I$ et comme $I = (p)$, p divise p_i. Notons de plus que $p \in I$ et donc on peut fixer $(\mathrm{a}_i)_{1 \leqslant i \leqslant r} \in \mathbb{K}[T]^r$ tels que

$$p = \sum_{i=1}^{r} \mathrm{a}_i p_i.$$

Si maintenant q divise les polynômes p_i, on en déduit que q divise p. p est bien un pgcd recherché.

Enfin, reste à montrer que ce pgcd est unique à un scalaire non nul près. Soit alors $\tilde{p}$ un autre pgcd des $(p_i)_{1 \leqslant i \leqslant r}$. Par définition, p divise $\tilde{p}$, et $\tilde{p}$ divise p. Les polynômes p et $\tilde{p}$ sont donc de même degré, et la division de l'un par l'autre (il existe $b \in \mathbb{K}[T]$ tel que $p = b\tilde{p}$) montre qu'ils diffèrent d'une constante multiplicative [3]. □

Exercice III.2.6. Déterminer les pgcd des familles suivantes :

(1) $(X - a), (X - b)$ où $a, b \in \mathbb{K}^2$,

(2) $X^2 - 2X + 1$ et $X - 1$.

Définition III.2.7. Soient $p_1, \dots, p_r$ des éléments non tous nuls de $\mathbb{K}[T]$. On dit qu'ils sont premiers entre eux si leur pgcd est une constante, i.e. si

$$pgcd(p_1, \dots, p_r) \in \mathbb{K}^*.$$

On a alors le Théorème de Bezout, version polynomiale :

Théorème III.2.8 (Théorème de Bézout). *Soient $p_1, \dots, p_r$ des éléments non tous nuls de $\mathbb{K}[T]$. Ces polynômes sont premiers entre eux si et seulement si il existe $(\mathrm{a}_i)_{1 \leqslant i \leqslant r} \in \mathbb{K}[T]^r$ tels que*

$$\mathrm{a}_1 p_1 + \ldots + \mathrm{a}_r p_r = 1.$$

Démonstration. Le fait que ces polynômes soient premiers entre eux implique (cf preuve de l'existence du pgcd) que l'idéal qu'ils engendrent est $\mathbb{K}[T]$ et donc l'existence des $(\mathrm{a}_i)_{1 \leqslant i \leqslant r} \in \mathbb{K}[T]^r$ tels que

$$\mathrm{a}_1 p_1 + \ldots + \mathrm{a}_r p_r = 1.$$

Réciproquement, tout diviseur des $(p_i)_{1 \leqslant i \leqslant r}$ doit diviser $1 = \mathrm{a}_1 p_1 + \ldots + \mathrm{a}_r p_r$, et donc doit être une constant non nulle. Le résultat suit. □

3. On dit qu'ils sont *associés*

Corollaire III.2.9. *Si $p_1, \dots, p_r$ sont des éléments non tous nuls de $\mathbb{K}[T]$, il existe $(a_i)_{1 \leqslant i \leqslant r} \in \mathbb{K}[T]^r$ tels que*

$$a_1 p_1 + \dots + a_r p_r = pgcd(p_1, \dots, p_r).$$

Démonstration. Il suffit de remarquer que si $p := pgcd(p_1, \dots, p_r)$, alors les polynômes p/p_i sont premiers entre eux. □

Enfin, comme pour les entiers, on peut décomposer les polynômes en produits de facteurs irréductibles.

Définition III.2.10. Un polynôme $p \in \mathbb{K}[T]$ est dit irréductible (ou premier) si
(i) p n'est pas une constante ($p \notin \mathbb{K}$),
(ii) Les seuls diviseurs de p sont les constantes et les polynômes du type λp, où $\lambda \in \mathbb{K}$.

Exemple III.2.11. Les irréductibles de $\mathbb{R}[X]$ et $\mathbb{C}[X]$ doivent être connus :
- Sur $\mathbb{C}$, les polynômes irréductibles sont les polynômes de degré 1, de la forme $aT + b$ où $(a, b) \in \mathbb{C}^2$.
- Sur $\mathbb{R}$, les polynômes irréductibles sont les polynômes de degré 1 et ceux de degré 2 dont le discriminant est négatif.

Théorème III.2.12. *Soit $p \in \mathbb{K}[T]$ un polynôme non constant. Alors p admet une décomposition en produit de facteurs irréductibles, c'est à dire qu' il existe $s \in \mathbb{N}^*$, des polynômes irréductibles $(p_i)_{1 \leqslant i \leqslant s} \in \mathbb{K}[T]^s$ deux à deux distincts, et des entiers non nuls $(n_i)_{1 \leqslant i \leqslant s} \in \mathbb{N}^s$ tels que*

$$p = \Pi_{i=1}^{s} p_i^{n_i}.$$

De plus, si on a une autre décomposition de p en produits d'irréductibles :

$$p = \Pi_{i=1}^{r} q_i^{m_i},$$

alors $s = r$, et il existe une permutation $\sigma \in \mathfrak{S}_s$ de $\{1, \dots, s\}$ telle que pour tout $i \in \{1, \dots, s\}$, p_i et $q_{\sigma(i)}$ sont égaux à un scalaire multiplicatif près, et $n_i = m_{\sigma(i)}$.

Démonstration. Pour l'existence, on procède par récurrence sur le degré de p : si p est irréductible on peut conclure, et sinon on trouve une décomposition de p en produit de deux polynômes non constants de degrés inférieurs à celui de p, auxquels on applique l'hypothése de récurrence. Notez que tous les polynômes de degré 1 sont irréductibles dans [4] $\mathbb{K}[T]$.

Pour l'unicité, on peut procéder par récurrence sur s et utiliser le lemme de Gauss. On renvoit à [**RW1**] pour les détails. □

Ce Théorème nous sera surtout utile en pratique pour déterminer une factorisation en produit de facteurs premiers entre eux des polynômes caractéristiques et minimaux que l'on va rencontrer.

Corollaire III.2.13. *Si $p, q \in \mathbb{K}[T]^2$ se décomposent en produits de facteurs irréductibles de la façon suivante (on autorise ici $m_i = 0$ et $n_i = 0$) :*

$$p = \Pi_{i=1}^{s} p_i^{n_i}$$

et

$$q = \Pi_{i=1}^{s} p_i^{m_i},$$

4. Attention, ceci n'est pas vrai dans $\mathbb{A}[X]$ pour un anneau $\mathbb{A}$ en général. Par exemple $2X + 2 \in \mathbb{Z}[X]$ n'est pas irréductible dans $\mathbb{Z}[X]$.

alors

$$pgcd(p,q) = \Pi_{i=1}^{s} p_i^{\min(n_i,m_i)}.$$

En particulier, p et q sont premiers entre eux si et seulement si ils n'ont aucun facteur irréductible en commun.

2.2. Application : le polynôme minimal. On revient à l'étude de nos endomorphismes. On rappelle que $u \in \mathrm{End}_{\mathbb{K}}(E)$ et $A \in \mathrm{M}_n(\mathbb{K})$.

DÉFINITION III.2.14. On appelle idéal annulateur de u (ou de Φ_u) l'ensemble de $\mathbb{K}[T]$ défini par :

$$I_u := \ker(\Phi_u) = \{p \in \mathbb{K}[T],\ p(u) = 0\}.$$

Ses éléments sont appelés polynômes annulateurs de u.

REMARQUE III.2.15. De même, on a l'idéal annulateur de A (ou de Φ_A) :

$$I_A := \ker(\Phi_A) = \{p \in \mathbb{K}[T],\ p(A) = 0\} \subset \mathbb{K}[T]$$

ainsi que les polynômes annulateurs de A.

Le lemme suivant est immédiat.

LEMME III.2.16. *Les ensembles I_u et I_A sont des idéaux de $\mathbb{K}[T]$.*

THÉORÈME III.2.17. *Il existe un unique polynôme unitaire μ_u (resp. μ_A) tel que*

$$I_u = (\mu_u) = \{a\mu_u,\ a \in \mathbb{K}[T]\},$$

(resp. $I_A = (\mu_A)$).

DÉMONSTRATION. La preuve est exactement la même que pour montrer l'existence de $p \in \mathbb{K}[T]$ tel que $I = (p)$ dans la démonstration de la Proposition III.2.5 (le caractère unitaire impose l'unicité). Il faut tout de même s'assurer que $I_u \neq (0)$, c'est à dire qu'il existe $\psi \in \mathbb{K}[T]\backslash\{0\}$ tel que $\psi(u) = 0$. Mais comme $\mathrm{End}_{\mathbb{K}}(E)$ est un espace vectoriel de dimension n^2, la famille $\{\mathrm{Id}_E, u, u^2, \ldots, u^{n^2}\}$ est liée, d'où le résultat. □

DÉFINITION III.2.18. Le polynôme μ_u (resp. μ_A) est appelé polynôme minimal de u (resp. de A).

REMARQUE III.2.19. Par définition, μ_u est le polynôme unitaire de plus petit degré qui annule u. Si $\psi \in \mathbb{K}[T]$ est un autre polynôme annulateur de u, alors μ_u divise ψ.

Par le théorème de Cayley-Hamilton :

COROLLAIRE III.2.20. *Le polynôme caractéristique appartient à I_u et est divisible par μ_u.*

Ainsi, les racines de μ_u sont toutes des valeurs propres de u. En fait on a le résultat remarquable suivant (valable également pour μ_A et $\mathrm{Sp}_{\mathbb{K}}(A)$) :

THÉORÈME III.2.21. *Soit $\lambda \in \mathbb{K}$. Alors $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ si et seulement si $\mu_u(\lambda) = 0$.*

DÉMONSTRATION. Il s'agit de montrer que toute valeur propre de u est racine de μ_u. Soit donc $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$. Il existe $x \neq 0$ un vecteur tel que $u(x) = \lambda x$. On calcul alors pour tout $\psi \in \mathbb{K}[T]$:

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

On a en particulier

$$0 = \mu_u(u)(x) = \mu_u(\lambda)x,$$

et comme x est non nul, $\mu_u(\lambda) = 0$. □

On en déduit les corollaires suivants :

Corollaire III.2.22. *Les racines de μ_u et χ_u sont les mêmes. Ainsi, u est scindé sur $\mathbb{K}$ si et seulement si μ_u l'est.*

Démonstration. Le premier énoncé découle du théorème qui précède. Pour le second, on va se contenter des cas $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, qui sont triviaux. □

On laisse au lecteur le soin d'énoncer les résultats analogues pour les matrices.

Exemple III.2.23. Supposons que $\dim_{\mathbb{R}}(E) = 6$ et que χ_u se décompose en facteurs premiers comme

$$\chi_u(T) = (T-1)(T-2)(T^2+T+1)^2.$$

Alors nécessairement,

$$\mu_u \in \{(T-1)(T-2)(T^2+T+1), (T-1)(T-2)(T^2+T+1)^2\}.$$

En effet, χ_u n'est pas scindé donc μ_u non plus, et 1 et 2 doivent être racines de μ_u. On identifie alors les seuls diviseurs de χ_u possibles.

Exemples III.2.24. On laisse en exercice la preuve des faits suivants :

(1) Si $A = 0$, alors $\mu_A(T) = T$ et $\chi_A(T) = T^n$.

(2) Si $A = \mathrm{Id}_n$, alors $\mu_A(T) = T - 1$ et $\chi_A(T) = (T-1)^n$.

(3) Si

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

alors $\chi_A(T) = \mu_A(T) = T^2$.

(4) Si

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

alors $\chi_A(T) = (T-1)^3$. Nécessairement,

$$\mu_A \in \{(T-1)^3, (T-1)^2, (T-1)\}.$$

Comme $A \neq \mathrm{Id}_3$, on calcule $(T-1)^2(A) = 0$ et $\mu_A(T) = (T-1)^2$.

Exercice III.2.25. Soit $A \in \mathrm{M}_n(\mathbb{R})$. On peut également voir A comme un élément de $\mathrm{M}_n(\mathbb{C})$. On dénote par $A^{\mathbb{C}}$ l'extension de A au corps des complexes. On va montrer que $\chi_A = \chi_{A^{\mathbb{C}}}$ et $\mu_A = \mu_{A^{\mathbb{C}}}$.

(1) Montrer en revenant à la définition que $\chi_A = \chi_{A^{\mathbb{C}}}$.

(2) Montrer que $\overline{\mu_{A^{\mathbb{C}}}} \in I_{A^{\mathbb{C}}}$.

(3) En déduire que $\mu_{A^{\mathbb{C}}}$ est à coefficients réels.

(4) Conclure que $\mu_A = \mu_{A^{\mathbb{C}}}$.

On finit cette section par un théorème fondamental qui donne une dernière caractérisation des endomorphismes diagonalisables. On a besoin pour cela de la notion suivante :

Définition III.2.26. Un polynôme $p \in \mathbb{K}[T]$ est dit simplement scindé (sur $\mathbb{K}$) si il est scindé et n'admet que des racines simples. Autrement dit, il existe $(\lambda_1, \ldots, \lambda_d) \in \mathbb{K}^d$, $\lambda_i \neq \lambda_j$ pour $i \neq j$, et $a \in \mathbb{K}$ tels que

$$p(X) = a\Pi_{i=1}^d (X - \lambda_i).$$

Théorème III.2.27. *Un endomorphisme $u \in \mathrm{End}_{\mathbb{K}}(E)$ est diagonalisable si et seulement si son polynôme minimal μ_u est simplement scindé.*

Remarque III.2.28. Bien entendu on a la version matricielle de ce résultat : $A \in \mathrm{M}_n(\mathbb{K})$ est diagonalisable si et seulement si μ_A est simplement scindé.

On reporte la preuve de ce résultat à la Section 3.3, car sa démonstration (le sens réciproque) nécessite le "Lemme des noyaux" qui est l'objet de la section suivante.

3. Lemme des noyaux

Dans cette partie on étudie de manière systématique le lien entre factorisation d'un polynôme annulateur de u et décomposition de E en somme directe.

3.1. Espaces stables et polynômes. On a déjà rencontré la notion d'espace stable :

Définition III.3.1. Un sous-espace vectoriel $F \subset E$ (resp. $F \subset \mathbb{K}^n$) est dit stable (ou invariant) par u (resp. par A) si $u(F) \subset F$ (resp. $A \cdot F \subset F$) [5].

Proposition III.3.2. *Soient $u, v \in \mathrm{End}_{\mathbb{K}}(E)$. On suppose que u et v commutent, i.e. $u \circ v = v \circ u$. Alors $\mathrm{Im}(v)$ et $\ker(v)$ sont stables par u.*

Démonstration. Soit $y \in \mathrm{Im}(v)$, $y = v(x)$ pour $x \in E$. Alors $u(y) = v(u(x))$ car u et v commutent, et donc $u(y) \in \mathrm{Im}(v)$.

Soit $x \in \ker(v)$. Alors $v(u(x)) = u(v(x)) = u(0) = 0$ car $u \circ vv = v \circ u$ et $v(x) = 0$. D'où le résultat. □

Le corollaire suivant sera particulièrement utile dans ce cours :

Corollaire III.3.3. *Si $v \in \mathbb{K}[u]$, $\mathrm{Im}(v)$ et $\ker(v)$ sont stables par u, et $\mathrm{Im}(u)$ et $\ker(u)$ sont stables par v.*

Démonstration. L'algèbre $\mathbb{K}[u]$ est commutative, et le résultat se déduit de la proposition précédente. □

Exemple III.3.4. Pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, u et $u - \lambda \mathrm{Id}_E$ commutent, donc E_λ est stable par u.

Remarque III.3.5. Si $F \subset E$ est un sous-espace vectoriel stable par u, et si G est un supplémentaire de F dans E, dans la décomposition $E = F \oplus G$, i.e. dans une base de E de la forme $\mathcal{B}_F \cup \mathcal{B}_G$ où $\mathcal{B}_F, \mathcal{B}_G$ sont des bases de F et G, la matrice par bloc de u est de la forme :

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}.$$

Si de plus G est stable par u, $C = 0$.

La proposition suivante donne un premier lien entre la décomposition de E en sous-espaces stables et la factorisation de ses polynômes annulateurs :

5. La notation $A \cdot F$ signifie $\{A \cdot X, X \in F\}$.

PROPOSITION III.3.6. *Soit $F \subset E$ un sous-espace vectoriel stable par u. On pose $u_F = u_{|F} \in \mathrm{End}_{\mathbb{K}}(F)$ la restriction de u à F. Alors χ_{u_F} divise χ_u et μ_{u_F} divise μ_u.*

DÉMONSTRATION. On fixe G un supplémentaire de F dans E, ainsi que $\mathcal{B}_F$ et $\mathcal{B}_G$ des bases de F et G. On dispose alors d'une base $\mathcal{B} = \mathcal{B}_E \cup \mathcal{B}_F$ de E dans laquelle $[u]_{\mathcal{B}}$ est triangulaire supérieure par blocs :

$$[u]_{\mathcal{B}} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

avec $B = [u_F]_{\mathcal{B}_F}$ par construction. Le calcul par blocs donne

$$\det(T\mathrm{Id}_E - u) = \det(T\mathrm{Id}_{\dim(F)} - B) \cdot \det(T\mathrm{Id}_{n-\dim(F)} - D),$$

d'où χ_{u_F} divise χ_u. Par récurrence sur k, on obtient la forme des matrices de u^k dans $\mathcal{B}$:

$$[u^k]_{\mathcal{B}} = \begin{bmatrix} B^k & * \\ 0 & D^k \end{bmatrix}.$$

Donc pour tout polynôme $p \in \mathbb{K}[T]$,

$$[p(u)]_{\mathcal{B}} = \begin{bmatrix} p(B) & * \\ 0 & p(D) \end{bmatrix},$$

et si la matrice $[p(u)]_{\mathcal{B}}$ est nulle, alors $p(B) = 0$. En particulier, $\mu_u(B) = 0$, et comme $B = [u_F]_{\mathcal{B}_F}$, μ_{u_F} divise μ_u par minimalité. □

3.2. Le lemme des noyaux. On conseil au lecteur de relire la Section 3.4 du Chapitre I avant d'étudier cette section.

THÉORÈME III.3.7 (Lemme des noyaux). *Soit $p(T) = p_1(T)p_2(T)$, où $(p, p_1, p_2) \in \mathbb{K}[T]^3$. Alors*

(1) On a

$$\ker(pgcd(p_1, p_2)(u)) = \ker(p_1(u)) \cap \ker(p_2(u)),$$

(2) Si de plus $pgcd(p_1, p_2) = 1$, alors

$$\ker(p(u)) = \ker(p_1(u)) \oplus \ker(p_2(u)),$$

(3) Si de plus $pgcd(p_1, p_2) = 1$, et $p(u) = 0$, alors

(a) E se décompose en somme directe

$$E = \ker(p_1(u)) \oplus \ker(p_2(u)),$$

(b) Les projecteurs $E \to \ker(p_i(u))$, $i = 1, 2$, sont des polynômes en u,

(c) $\mathrm{Im}(p_1(u)) = \ker(p_2(u))$ et $\mathrm{Im}(p_2(u)) = \ker(p_1(u))$.

Le corollaire suivant sera la forme la plus utilisée de ce résultat :

COROLLAIRE III.3.8. *Soit $p \in I_u$ un polynôme annulateur de u. Si p se décompose $p(T) = p_1(T) \dots p_s(T)$, où $(p_1, \dots, p_s) \in \mathbb{K}[T]^s$ sont premiers entre eux 2 à 2, alors*

$$E = \bigoplus_{i=1}^{s} \ker(p_i(u))$$

et les projecteurs $E \to \ker(p_i(u))$ sont des polynômes en u.

Remarque III.3.9. Ainsi, le corollaire III.3.8 s'applique à toute décomposition de χ_u ou μ_u en produit de facteurs premiers entre eux 2 à 2.

Démonstration. On procède par récurrence sur s. Le cas $s = 1$ est impliqué par $p \in I_u$. Supposons le résultat vrai au rang $s \in \mathbb{N}^*$ et montrons le au rang $s+1$. On pose $q := p_1 \dots p_s$ de telle sorte que $p = qp_{s+1}$. Les polynômes q et p_{s+1} sont premiers entre eux, et on peut appliquer le lemme des noyaux :

$$E = \ker(q(u)) \oplus \ker(p_{s+1}(u)).$$

De plus, les projecteurs sur ces deux espaces sont des polynômes en u. L'espace $F := \ker(q(u))$ est stable par u et q annule par définition la restriction de u à cet espace. On peut alors appliquer l'hypothèse de récurrence à $u_F = u_{|F}$, et on obtient la décomposition souhaitée (notez que le point (1) du lemme des noyaux implique que $\ker(p_i(u)) = \ker(p_i(u_F))$ pour $i \neq s+1$). Comme la restriction de u à F est donnée par la composition de u avec la projection sur F, et comme cette projection est un polynôme en u, on obtient la forme polynomiale de tous les projecteurs par récurrence également. □

Démonstration du lemme des noyaux. On pose $q := pgcd(p_1, p_2)$. Par le théorème de Bézout, on peut fixer $(a_1, a_2) \in \mathbb{K}[T]^2$ tel que

$$q = a_1 p_1 + a_2 p_2. \tag{11}$$

On va montrer le point (1) :

$$\ker(q(u)) = \ker(p_1(u)) \cap \ker(p_2(u)).$$

Par définition du pgcd, on peut fixer $b_i \in \mathbb{K}[T]$, $i = 1, 2$, tel que

$$p_i = b_i q.$$

Mais alors $p_i(u) = b_i(u) \circ q(u)$ et $\ker(q(u)) \subset \ker(p_1(u)) \cap \ker(p_2(u))$. Réciproquement, l'équation (11) donne $q(u) = a_1(u) \circ p_1(u) + a_2(u) \circ p_2(u)$, d'où l'on déduit que $\ker(p_1(u)) \cap \ker(p_2(u)) \subset \ker(q(u))$.

On suppose désormais que $pgcd(p_1, p_2) = 1$, et l'équation (11) implique alors

$$\mathrm{Id}_E = a_1(u) \circ p_1(u) + a_2(u) \circ p_2(u), \tag{12}$$

avec $(a_1(u), a_2(u), p_1(u), p_2(u)) \in \mathbb{K}[u]^4$ qui commutent entre eux.

On peut alors montrer le point (2). Notez que par (1),

$$\ker(p_1(u)) \cap \ker(p_2(u)) = \ker(\mathrm{Id}_E) = \{0_E\}$$

et donc $\ker(p_1(u))$ et $\ker(p_2(u))$ sont en somme directe. Comme $p = p_1 p_2$, on a $\ker(p_i(u)) \subset \ker(p(u))$, $i = 1, 2$. Reste à montrer que

$$\ker(p(u)) \subset \ker(p_1(u)) + \ker(p_2(u)).$$

On définit pour cela les endomorphismes :

$$\mathbf{p}_i := a_i(u) \circ p_i(u), \ i \in \{1, 2\}.$$

Soit $x \in \ker(p(u))$. Alors par (12),

$$x = \mathbf{p}_1(x) + \mathbf{p}_2(x).$$

De plus, $\mathbf{p}_1(x) \in \ker(p_2(u))$. En effet,

$$\begin{array}{rcl} p_2(u) \circ \mathbf{p}_1(x) & = & p_2(u) \circ a_1(u) \circ p_1(u)(x) \\ & = & a_1(u) \circ p_1(u) \circ p_2(u)(x) \\ & = & a_1(u) \circ p(u)(x) \\ & = & 0 \end{array}$$

car $x \in \ker(p(u))$. De même on démontre que $\mathbf{p}_2(x) \in \ker(p_1(u))$, d'où le résultat.

En addition on suppose maintenant que p est un polynôme annulateur de u. On va démontrer le point (3). Par (2), comme $p(u) = 0$,

$$E = \ker(p_1(u)) \bigoplus \ker(p_2(u)).$$

On rappelle l'équation (12) :

$$\mathrm{Id}_E = \mathbf{p}_1 + \mathbf{p}_2.$$

Comme $p(u) = 0$, on a également

$$\mathbf{p}_1 \circ \mathbf{p}_2 = 0,$$

ce qui implique par la Proposition I.3.34 que $\mathbf{p}_1$ et $\mathbf{p}_2$, sont des projecteurs et que

$$E = \mathrm{Im}(\mathbf{p}_1) \bigoplus \mathrm{Im}(\mathbf{p}_2).$$

On a déjà vu dans la preuve de (2) que $\mathrm{Im}(\mathbf{p}_1) \subset \ker(p_2(u))$. Pour l'inclusion inverse, si $p_2(u)(x) = 0$, alors $\mathbf{p}_2(x) = 0$ et on a $x = \mathbf{p}_1(x) \in \mathrm{Im}(\mathbf{p}_1)$. Les projecteurs sur les espaces $\ker(p_1(u))$ et $\ker(p_2(u))$ sont donc les applications $\mathbf{p}_2$ et $\mathbf{p}_1$ qui sont bien des polynômes en u.

Pour conclure, il reste à montrer que $\ker(p_2(u)) = \mathrm{Im}(p_1(u))$. On a clairement par définition de $\mathbf{p}_1$ que $\ker(p_2(u)) = \mathrm{Im}(\mathbf{p}_1) \subset \mathrm{Im}(p_1(u))$. La réciproque s'obtient alors simplement du fait que $p_2(u) \circ p_1(u) = p(u) = 0$. $\square$

Remarque III.3.10. Pour mettre en pratique le lemme des noyaux, si p est un polynôme annulateur de u, on procède ainsi :

(1) On commence par factoriser $p = p_1 \dots p_r$ en produit de polynômes deux à deux premiers entre eux, avec pour tout i, p_i une puissance d'un polynôme irréductible.

(2) On décompose en éléments simples

$$\frac{1}{p(T)} = \sum_{i=1}^{r} \frac{a_i(T)}{p_i(T)}$$

où pour tout $i \in \{1, \dots, r\}$, $\deg(a_i) < \deg(p_i)$. On en déduit en multipliant par p :

$$1 = \sum_{i=1}^{r} a_i q_i,$$

où $q_i := p/p_i$ est le quotient de p par p_i.

(3) On a alors

$$\mathrm{Id}_E = \sum_{i=1}^{r} \mathbf{p}_i,$$

avec $\mathbf{p}_i = a_i(u) \circ q_i(u)$. On peut alors vérifier que les $(\mathbf{p}_i)_{1 \leqslant i \leqslant r}$ sont bien les projecteurs associés à la décomposition du lemme des noyaux

$$E = \bigoplus_{i=1}^{s} \ker(p_i(u)).$$

Bien évidement, plus le polynôme annulateur est de petit degré, et plus les calculs sont aisés. En pratique, on sait toujours calculer le polynôme caractéristique. Le polynôme minimal sera un diviseur de ce polynôme, et si on peut l'identifier, lui appliquer le lemme des noyaux sera bien plus simple.

On va illustrer la méthode décrite dans la remarque précédente par un exemple.

Exemple III.3.11. Soit $u : \mathbb{R}^3 \to \mathbb{R}^3$ défini par

$$[u]_{\mathcal{B}_{can}} = A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & -3 & 2 \\ -2 & 1 & -4 \end{bmatrix} \in \mathrm{M}_3(\mathbb{R}),$$

On a déjà calculé le polynôme caractéristique de cette matrice (Exemple II.3.9) :

$$\chi_A(T) = (T+1)(T+2)^2.$$

Ici on décompose $\chi_A = p_1 p_2$ où $p_A(T) = (T+1)$ et $p_2(T) = (T+2)^2$. La décomposition en éléments simples (théorique) de $\frac{1}{\chi_A}$ est de la forme

$$\frac{1}{\chi_A} = q(T) + \frac{\mathrm{a}_1(T)}{T+1} + \frac{\mathrm{b}_1(T)}{T+2} + \frac{\mathrm{b}_2(T)}{(T+2)^2}$$

où q est la partie entière, $\deg(a_1) < \deg(T+1)$ et $\deg(b_i) < \deg(T+2)$. Bien entendu, la partie entière est nulle, a_1, b_1 et b_2 sont des constantes, et on peut mettre les derniers termes sur le même dénominateur pour obtenir l'existence de $\mathrm{a}_2(T) \in \mathbb{R}_1[T]$ tel que

$$\frac{1}{\chi_A} = \frac{\mathrm{a}_1(T)}{T+1} + \frac{\mathrm{a}_2(T)}{(T+2)^2}.$$

On peut obtenir les coefficients par identification, et on trouve

$$\frac{1}{\chi_A} = \frac{1}{3}\left(\frac{1}{(T+1)} - \frac{T+1}{(T+2)^2}\right).$$

On multiplie par χ_A et on évalue en A :

$$\mathrm{Id}_3 = \frac{1}{3}\left((A+2\mathrm{Id}_3)^2 - (A+\mathrm{Id})^2\right).$$

D'après le lemme des noyaux,

$$\mathbb{R}^3 = \ker(A+\mathrm{Id}) \bigoplus \ker((A+2\mathrm{Id}_3)^2)$$

et les projecteurs sont donnés par :

$$\begin{array}{ccc} \mathbb{R}^3 & \to & \ker(A+\mathrm{Id}) \\ X & \mapsto & \frac{1}{3}(A+2\mathrm{Id}_3)^2 X \end{array}$$

et

$$\begin{array}{ccc} \mathbb{R}^3 & \to & \ker((A+2\mathrm{Id})^2) \\ X & \mapsto & -\frac{1}{3}(A+\mathrm{Id})^2 X. \end{array}$$

On peut calculer directement ces projecteurs à l'aide du calcul matriciel. On reviendra sur cet exemple à la section suivante.

3.3. Retour sur la diagonalisation. On fini ce chapitre par la preuve du Théorème III.2.27. On rappelle son énoncé :

THÉORÈME III.3.12. *Les assertions suivantes sont équivalentes :*
(i) u est diagonalisable,
(ii) μ_u est simplement scindé.

DÉMONSTRATION. Supposons u diagonalisable. On pose

$$p(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T - \lambda).$$

On sait que u, et donc μ_u est scindé, et que de plus les valeurs propres de u sont les racines de μ_u. Donc p divise μ_u. D'autre part, $p(u) = 0$. En effet, comme u est diagonalisable,

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} E_\lambda,$$

et il suffit de montrer que pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ et tout $x \in E_\lambda$, $p(u)(x) = 0$. Mais si $x \in E_\lambda$, $(T - \lambda \mathrm{Id}_E)(u)(x) = u(x) - \lambda x = 0$. On a bien $p \in I_u$ et donc μ_u divise p. Comme p et μ_u sont unitaires, ils sont égaux et μ_u est simplement scindé.

Pour la réciproque, supposons μ_u simplement scindé. On a alors

$$\mu_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T - \lambda)$$

car les racines de μ_u sont exactement les valeurs propres de u. Comme les $(T - \lambda)$, $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, sont deux à deux premiers entre eux, le lemme des noyaux permet de conclure :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \ker(u - \lambda \mathrm{Id}_E).$$

□

COROLLAIRE III.3.13. *Un endomorphisme $u \in \mathrm{End}_{\mathbb{K}}(E)$ est diagonalisable si et seulement si il annule un polynôme simplement scindé sur $\mathbb{K}$.*

DÉMONSTRATION. Si u est diagonalisable il annule μ_u qui est simplement scindé par le théorème précédent. Réciproquement, si I_u contient un polynôme simplement scindé p, par définition μ_u divise p et est donc lui-même simplement scindé. Le théorème précédent implique alors que u est diagonalisable. □

Avec la Proposition III.3.6 on en déduit :

COROLLAIRE III.3.14. *Supposons u diagonalisable. Alors la restriction de u à tout sous-espace vectoriel stable par u est diagonalisable.*

EXEMPLE III.3.15. On revient sur l'exemple

$$[u]_{\mathcal{B}_{can}} = A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & -3 & 2 \\ -2 & 1 & -4 \end{bmatrix} \in \mathrm{M}_3(\mathbb{R}).$$

On a vu (Exemple II.3.9) que A était diagonalisable [6]. Son polynôme minimal est donc simplement scindé et égal à

$$\mu_u(T) = (T+1)(T+2).$$

6. Notez que pour prouver que A est diagonalisable nous n'avons pas besoin de calculer une base de vecteurs propres. Pour cet exemple il était facile de remarquer que pour toute valeur propre λ on avait $\dim(E_\lambda) = \mathrm{mult}(\lambda, \chi_A)$.

On procède alors à la décomposition en éléments simples de son inverse :

$$\frac{1}{\mu_u(T)} = \frac{1}{T+1} - \frac{1}{T+2}$$

ce qui donne

$$\mathrm{Id}_3 = (A + 2\mathrm{Id}_3) - (A + \mathrm{Id}_3).$$

D'après le lemme des noyaux,

$$\mathbb{R}^3 = \ker(A + \mathrm{Id}) \oplus \ker(A + 2\mathrm{Id}_3)$$

et les projecteurs sont donnés par :

$$\begin{array}{rccc} \mathbf{p}_{-1} : & \mathbb{R}^3 & \rightarrow & \ker(A + \mathrm{Id}) \\ & X & \mapsto & (A + 2\mathrm{Id}_3)X \end{array}$$

et

$$\begin{array}{rccc} \mathbf{p}_{-2} : & \mathbb{R}^3 & \rightarrow & \ker((A + 2\mathrm{Id})) \\ & X & \mapsto & -(A + \mathrm{Id})X. \end{array}$$

On calcul aisément

$$[\mathbf{p}_{-1}]_{\mathcal{B}_{can}} = \begin{bmatrix} 4 & -2 & 4 \\ 2 & -1 & 2 \\ -2 & 1 & -2 \end{bmatrix}$$

et

$$[\mathbf{p}_{-2}]_{\mathcal{B}_{can}} = \begin{bmatrix} -3 & 2 & -4 \\ -2 & 2 & -2 \\ 2 & -1 & 3 \end{bmatrix}.$$

On en déduit une base de vecteurs propres de u, car pour $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, $\mathrm{Im}(\mathbf{p}_\lambda) = E_\lambda$, et l'image de $\mathbf{p}_\lambda$ est engendrée par les vecteurs colonnes de la matrice $[\mathbf{p}_\lambda]_{\mathcal{B}_{can}}$. Ainsi, $E_{-1} = \mathrm{Vect}\,\{(2, 1, -1)\}$ et $E_{-2} = \mathrm{Vect}\,\{(-3, -2, 2), (2, 2, -1)\}$. On a alors, si on pose

$$P = \begin{bmatrix} 2 & -3 & 2 \\ 1 & -2 & 2 \\ -1 & 2 & -1 \end{bmatrix},$$

$$P^{-1}AP = \mathrm{Diag}(-1, -2, -2).$$

EXERCICE III.3.16. Soit

$$B = \begin{bmatrix} 5 & -1 & -1 \\ 0 & 6 & 3 \\ -2 & -2 & 1 \end{bmatrix} \in \mathrm{M}_3(\mathbb{R}).$$

(1) Montrer que

$$\chi_B(T) = (T - 3)^2(T - 6).$$

(2) En déduire

$$\mathbb{R}^3 = \ker((B - 3\mathrm{Id}_3)^2) \oplus \ker(B - 6\mathrm{Id}_3).$$

(3) Déterminer les projecteurs associés à cette décomposition.

(4) En déduire une base de $\ker((B - 3\mathrm{Id}_3)^2)$ et de $\ker(B - 6\mathrm{Id}_3)$.

EXERCICE III.3.17. Soit $A \in \mathrm{M}_n(\mathbb{C})$. Montrer que s'il existe $k \in \mathbb{N}$ tel que $A^k = \mathrm{Id}_n$, alors A est diagonalisable.

Chapitre IV

Réduction de Jordan et décomposition de Dunford

Comme on l'a vu dans les chapitres précédents, un endomorphisme n'est pas forcément diagonalisable. Il est pour cela nécessaire et suffisant que son polynôme minimal soit simplement scindé. Cela dit, sur $\mathbb{C}$ (ou sur $\mathbb{K}$ algébriquement clos), tout endomorphisme est scindé, et ceci implique que tout endomorphisme est trigonalisable. On va étudier dans ce chapitre les endomorphismes trigonalisables. On verra en particulier la forme de Jordan et la décomposition de Dunford pour ces endomorphismes.

On conserve les notations : E est un $\mathbb{K}$-espace vectoriel de dimension finie n, $u \in \mathrm{End}_{\mathbb{K}}(E)$ et $A \in \mathrm{M}_n(\mathbb{K})$.

1. Espaces caractéristiques

On commence par étudier une décomposition de E en "sous-espaces caractéristiques" induite par le lemme des noyaux.

Définition IV.1.1. Soit $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ (resp. $\lambda \in \mathrm{Sp}_{\mathbb{K}}(A)$). On appelle espace caractéristique associé à λ le sous-espace vectoriel $F_\lambda \subset E$ (resp. $F_\lambda \subset \mathbb{K}^n$) défini par

$$F_\lambda = \bigcup_{j \geqslant 0} \ker(u - \lambda \mathrm{Id}_E)^j$$

(resp. $\bigcup_{j\geqslant 0} \ker(A - \lambda \mathrm{Id}_n)^j$).

Ainsi, par définition, on a pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$

$$\{0\} \subset E_\lambda \subset F_\lambda \subset E.$$

Lemme IV.1.2. *Soit $f \in \mathrm{End}_{\mathbb{K}}(E)$. Alors pour tout $j \in \mathbb{N}$,*

$$\ker(f^j) \subset \ker(f^{j+1}).$$

Démonstration. Si $x \in \ker f^j$, on calcule $f^{j+1}(x) = f(f^j(x)) = f(0) = 0$. □

On déduit de ce lemme appliqué à $f = u - \lambda \mathrm{Id}_E$:

$$\{0\} \subset \ker(u - \lambda \mathrm{Id}_E) \subset \ldots \subset \ker(u - \lambda \mathrm{Id}_E)^j \subset \ker(u - \lambda \mathrm{Id}_E)^{j+1} \subset \ldots \subset E.$$

Proposition IV.1.3. *Soit $\lambda \in \mathrm{Sp}_{\mathbb{K}}(E)$. On pose $m_\lambda := \mathrm{mult}(\lambda, \mu_u)$. On a alors*

(1) $\ker(u - \lambda \mathrm{Id}_E)^{m_\lambda - 1} \subsetneq \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda}$,

(2) $\forall j \geqslant m_\lambda$, $\ker(u - \lambda \mathrm{Id}_E)^j = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda}$.

En conclusion,

$$F_\lambda = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda}.$$

On dit que la suite d'espaces $\ker(u - \lambda \mathrm{Id}_E)^j$ *stationne à partir du rang* m_λ. On montrera dans la prochaine section que pour $j \leqslant m_\lambda - 1$, on a

$$\ker(u - \lambda \mathrm{Id}_E)^j \subsetneq \ker(u - \lambda \mathrm{Id}_E)^{j+1}.$$

DÉMONSTRATION. On applique le lemme des noyaux à

$$\mu_u(T) = (T - \lambda)^{m_\lambda} f(T)$$

où $f(T)$ et $(T - \lambda)$ sont premiers entre eux. On obtient :

$$E = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda} \oplus \ker(f(u)). \tag{13}$$

Pour montrer (1), on sait que $\ker(u - \lambda \mathrm{Id}_E)^{m_\lambda - 1} \subset (u - \lambda \mathrm{Id}_E)^{m_\lambda}$. Supposons par l'absurde que $\ker(u - \lambda \mathrm{Id}_E)^{m_\lambda - 1} = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda}$. On obtient alors $(T - \lambda)^{m_\lambda - 1} f(T)$ est un polynôme annulateur de u, ce qui contredit la minimalité de μ_u. En effet, si $x \in E$ s'écrit $x = y + z$ dans la décomposition (13)

$$\begin{aligned} (u - \lambda)^{m_\lambda - 1} \circ f(u)(x) &= (u - \lambda)^{m_\lambda - 1} \circ f(u)(y) + (u - \lambda)^{m_\lambda - 1} \circ f(u)(z) \\ &= f(u) \circ (u - \lambda)^{m_\lambda - 1}(y) + (u - \lambda)^{m_\lambda - 1} \circ f(u)(z) \\ &= 0 + 0 \end{aligned}$$

car $f(u), (u - \lambda \mathrm{Id}_E) \in \mathbb{K}[u]$ commutent.

Pour le point (2), on remarque que si $j \geqslant m_\lambda$, le polynôme $(T - \lambda)^j f(T)$ annule u. Du lemme des noyaux on déduit une autre décomposition de E :

$$E = \ker(u - \lambda \mathrm{Id}_E)^j \oplus \ker(f(u)),$$

et donc $\dim(\ker(u - \lambda \mathrm{Id}_E)^j) = \dim(\ker(u - \lambda \mathrm{Id}_E)^{m_\lambda})$, d'où légalité des espaces car $\ker(u - \lambda \mathrm{Id}_E)^{m_\lambda} \subset \ker(u - \lambda \mathrm{Id}_E)^j$. □

COROLLAIRE IV.1.4. *Un endomorphisme scindé* $u \in \mathrm{End}_{\mathbb{K}}(E)$ *est diagonalisable si et seulement si pour tout* $\lambda \in \mathrm{Sp}_{\mathbb{K}}(E)$, $E_\lambda = F_\lambda$.

DÉMONSTRATION. Sous ces hypothèses, u est diagonalisable si et seulement si μ_u est simplement scindé, si et seulement si pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, $m_\lambda = 1$. D'où le résultat. □

THÉORÈME IV.1.5. *Soit* $\lambda \in \mathrm{Sp}_{\mathbb{K}}(E)$. *On pose* $m_\lambda := \mathrm{mult}(\lambda, \mu_u)$ *et* $n_\lambda = \mathrm{mult}(\lambda, \chi_u)$. *Alors*

$$\dim(F_\lambda) = n_\lambda.$$

DÉMONSTRATION. On utilise la décomposition (13) de la proposition précédente :

$$E = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda} \oplus \ker(f(u)) = F_\lambda \oplus \ker(f(u))$$

où $\mu_= (T - \lambda)^{m_\lambda} f$, $f(\lambda) \neq 0$ et où les espaces F_λ et $\ker(f(u))$ sont stables par u (ce sont des noyaux de polynômes en u). Dans une base adaptée à cette décomposition (notez que $u_{|F_\lambda}$ est trigonalisable car scindé), la matrice de u prend la forme

$$\begin{bmatrix} A_\lambda & 0 \\ 0 & B \end{bmatrix}$$

où $B \in \mathrm{M}_{n - \dim(F_\lambda)}(\mathbb{K})$ et $A_\lambda \in \mathrm{M}_{\dim(F_\lambda)}(\mathbb{K})$ qui représente $u_{|F_\lambda}$ est triangulaire supérieure avec uniquement λ sur la diagonale. On en déduit le polynôme caractéristique :

$$\chi_u(T) = \chi_{A_\lambda}(T)\chi_B(T) = (T - \lambda)^{\dim(F_\lambda)}\chi_B(T).$$

D'autre part $f(T)$ annule la restriction de u à $\ker(f(u))$ donc les racines de χ_B sont des zéros de f. Comme $f(\lambda) \neq 0$, λ n'est pas racine de χ_B et donc $n_\lambda = \mathrm{mult}(\lambda, \chi_u) = \dim(F_\lambda)$. □

Pour résumer la situation, si u est scindé, ce qui est toujours le cas sur $\mathbb{C}$, on peut écrire

$$\mu_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T-\lambda)^{m_\lambda},$$
$$\chi_u(T) = \Pi_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}(T-\lambda)^{n_\lambda},$$

et on a pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, $m_\lambda \leqslant n_\lambda$,

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} F_\lambda = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda}$$

où

$$\dim(F_\lambda) = n_\lambda.$$

Par ailleurs les F_λ sont stables par u, donc dans une base de la forme $\mathcal{B} = \bigcup_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \mathcal{B}_\lambda$, où $\mathcal{B}_\lambda$ est une base de F_λ, on a

$$[u]_{\mathcal{B}} = \begin{bmatrix} A_{\lambda_1} & 0 & \dots & 0 \\ 0 & A_{\lambda_2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & A_{\lambda_r} \end{bmatrix}$$

où $A_\lambda = [u_{|F\lambda}]_{\mathcal{B}_\lambda}$. Enfin, $u_\lambda := u_{|F_\lambda}$ est trigonalisable (il annule $(T-\lambda)^{m_\lambda}$ qui est scindé) et on peut choisir une base $\mathcal{B}_\lambda$ dans laquelle A_λ soit triangulaire supérieure :

$$A_\lambda = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & \lambda & * & \vdots \\ \vdots & 0 & \ddots & * \\ 0 & \dots & 0 & \lambda \end{bmatrix}.$$

Si on pose $f := (u - \lambda \mathrm{Id}_E)_{|F_\lambda}$, on a $f^{m_\lambda} = 0$. Afin de déterminer une forme spécifique pour A_λ, on est donc amené à étudier des endomorphismes qui annule X^m, $m \in \mathbb{N}^*$.

2. Endomorphismes nilpotents

Définition IV.2.1. On dit que u (resp. A) est nilpotent (resp. nilpotente) d'indice m si $u^m = 0$ et $u^{m-1} \neq 0$ (resp. $A^m = 0$ et $A^{m-1} \neq 0$).

Remarque IV.2.2. Un endomorphisme est donc nilpotent d'indice $m \in \mathbb{N}^*$ si T^m est son polynôme minimal. Il est alors scindé et sa seule valeur propre est 0. De plus, u est nilpotent si et seulement si $u^n = 0$ (ou $n = \dim(E)$).

On énonce le résultat suivant pour un endomorphisme. On laisse au lecteur le soin de le formuler dans le cas des matrices.

Proposition IV.2.3. *Soit $v \in \mathrm{End}_{\mathbb{K}}(E)$ un endomorphisme nilpotent d'indice m. Alors la suite $(\ker(v^j))_{j \in \mathbb{N}^*}$ est croissante pour l'inclusion, strictement croissante de 0 à m puis stationnaire à E.*

Démonstration. On sait déjà d'après la section précédente que cette suite d'espaces est croissante et stationne à E, ainsi que $\ker(v^{m-1}) \subsetneq \ker v^m$. Reste à montrer qu'elle est strictement croissante pour $j \leqslant m-1$.

Supposons alors que $\ker v^j = \ker v^{j+1}$ où $j \leqslant m-2$. On a alors $v^{m-j-1}(E) \subset \ker v^{j+1} = \ker v^j$. Mais alors T^{m-1} annule u, ce qui est absurde. □

Remarque IV.2.4. Notons les faits suivants :

(1) Une matrice $A \in \mathrm{M}_n(\mathbb{K})$ est nilpotente si et seulement si A est semblable à une matrice triangulaire supérieure avec des zéros sur la diagonale.

(2) Si v est nilpotent d'indice m,

$$\{0\} \subsetneq \ker v \subsetneq \ldots \subsetneq \ker v^j \subsetneq \ldots \subsetneq \ker v^m = E.$$

La matrice nilpotente par excellence est la suivante :

Définition IV.2.5. On définit le bloc de Jordan principal J_m, $m \in \mathbb{N}^*$, par

$$J_m := \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \ldots & \ldots & \ldots & 0 \end{bmatrix} \in \mathrm{M}_m(\mathbb{K}).$$

Proposition IV.2.6. *Le bloc de Jordan principal est une matrice nilpotente d'indice m.*

Démonstration. Un calcul direct montre que

$$J_m^2 := \begin{bmatrix} 0 & & & \\ 0 & & J_{m-1} & \\ \vdots & & & \\ 0 & \ldots & \ldots & 0 \end{bmatrix},$$

et une récurrence permet de conclure. □

On a alors le théorème suivant :

Théorème IV.2.7. *Soit $v \in \mathrm{End}_{\mathbb{K}}(E)$ un endomorphisme nilpotent d'indice m. Alors il existe $(m_1, \ldots, m_s) \in (\mathbb{N}^*)^s$ tel que $m_1 + \ldots + m_s = n$ et $m = \max\{m_1, \ldots, m_s\}$, ainsi qu'une base $\mathcal{B}$ de E telle que*

$$[v]_{\mathcal{B}} = \mathrm{Diag}(J_{m_1}, \ldots, J_{m_s}).$$

Exemple IV.2.8. Donner la matrice $\mathrm{Diag}(J_2, J_1, J_3)$.

Démonstration. On montre le résultat par récurrence sur la dimension de E.

Soit $x \in E$ tel que $v^{m-1}(x) \neq 0$. On peut vérifier que la famille

$$\mathcal{B}_F := \{v^{m-1}(x), \ldots, v(x), x\}$$

est libre. De plus, le sous-espace $F \subset E$ de E engendré par cette famille est stable par u. Supposons trouvé un supplémentaire G à F stable par u. Alors dans la décomposition $E = F \oplus G$, on aura

$$[v]_{\mathcal{B}} = \begin{bmatrix} J_m & 0 \\ 0 & * \end{bmatrix}$$

où $\mathcal{B}$ est une base du type $\mathcal{B}_F \cup \mathcal{B}_G$. On peut alors appliquer l'hypothèse de récurrence à $v_{|G}$ et conclure.

Reste donc à montrer l'existence de G. Pour cela on peut passer au dual $E^\vee$. Soit $\phi \in E^\vee$ telle que $\phi(v^{m-1}(x)) \neq 0$. On définit cette fois-ci

$$\mathcal{B}_G := \{\phi \circ v^{m-1}, \ldots, \phi \circ v, \phi\},$$

et on vérifie de la même manière que c'est une famille libre de $E^\vee$. On pose alors G l'orthogonal de Vect $(\mathcal{B}_G)$. Comme Vect $\mathcal{B}_G$ est stable par v^*, G est stable par v. De plus, la dimension de G est égale à la dimension de E moins celle de F. Reste donc à montrer que $F \cap G = \{0\}$, ce qui peut être vérifié directement. □

Remarque IV.2.9. La forme donnée dans le théorème ci-dessus est unique à permutation des blocs près. En effet, par le point (2) de la Remarque IV.2.4, si l'on pose $d_j := \dim(\ker v^j)$, on remarque que $d_j - d_{j-1}$ est le nombre de blocs de Jordan de la forme J_l avec $l \geqslant j$ qui apparaissent dans l'écriture $\mathrm{Diag}(J_{m_1}, \dots, J_{m_s})$. Ces blocs de Jordan sont donc uniquement déterminés par les dimensions des espaces $\ker v^j$, et donc par v.

Exercice IV.2.10. Soit
$$A := \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

(1) Montrer que A est nilpotente d'indice 3.

(2) Montrer qu'il n'existe pas de matrice $B \in \mathrm{M}_3(\mathbb{R})$ avec $B^2 = A$.

Exercice IV.2.11. Soit $A \in \mathrm{M}_4(\mathbb{R})$ une matrice nilpotente. Montrer que A est semblable à l'une des matrices suivantes :
$$\begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0 \end{bmatrix}.$$

3. Réduction de Jordan

Des deux sections précédentes on déduit le résultat suivant :

Théorème IV.3.1 (Réduction de Jordan). *Soit $u \in \mathrm{End}_{\mathbb{K}}(E)$ scindé sur $\mathbb{K}$. On pose*
$$\forall \lambda \in \mathrm{Sp}_{\mathbb{K}}(u),\ F_\lambda = \ker(u - \lambda \mathrm{Id}_E)^{m_\lambda},$$
où $m_\lambda = \mathrm{mult}(\lambda, \mu_u)$. Alors, pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, il existe une base $\mathcal{B}_\lambda$ de F_λ (de cardinal $\mathrm{mult}(\lambda, \chi_u)$) telle que dans $\mathcal{B} := \bigcup_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \mathcal{B}_\lambda$ on ait
$$[u]_{\mathcal{B}} = \mathrm{Diag}(J_{\lambda_1, m_{\lambda_1,1}}, \dots, J_{\lambda_1, m_{\lambda_1,s_1}}, J_{\lambda_2, m_{\lambda_2,1}}, \dots, J_{\lambda_2, m_{\lambda_2,s_2}}, \dots, J_{\lambda_r, m_{\lambda_r,s_r}}),$$
où pour tout $(\lambda, i) \in \mathbb{K} \times \mathbb{N}$, $J_{\lambda,i} := \lambda \mathrm{Id}_i + J_i$.

Remarque IV.3.2. Les blocs du type $J_{\lambda,i}$ sont appelés blocs de Jordan. La décomposition en blocs de Jordan est unique à permutation des blocs près.

Ainsi, sur $\mathbb{C}$, toute matrice A est semblable à une matrice diagonale par blocs où tous les blocs sont des blocs de Jordan. De plus, cette forme diagonale par blocs est unique à permutation des blocs près. On obtient alors une description des "classes de similitudes" des matrices de $\mathrm{M}_n(\mathbb{C})$ (i.e. des classes d'équivalence sous la relation $A \sim B$ si A et B sont semblables).

En pratique, la méthode pour déterminer la "forme de Jordan" d'une matrice A (c'est à dire déterminer $P \in \mathrm{GL}(\mathbb{K})$ telle que $P^{-1}AP$ soit diagonale par blocs de Jordan) est la suivante :

(1) Déterminer le polynôme minimal de A,

(2) Décomposer ce polynôme en produit de facteurs premiers entre eux deux à deux de type $(T-\lambda)^{m_\lambda}$,

(3) Pour chacun des $F_\lambda = \ker(A-\lambda \mathrm{Id}_n)^{m_\lambda}$, procéder de la façon suivante (on note $d_j = \dim(\ker(A-\lambda\mathrm{Id}_n)^j)$:

 (a) Déterminer une famille libre $\mathcal{F}_1$ de $\delta_{m_\lambda} := d_{m_\lambda} - d_{m_\lambda - 1}$ vecteurs $\{x_1, \dots, x_{\delta_{m_\lambda}}\}$ de

$$\ker(A-\lambda\mathrm{Id}_n)^{m_\lambda} \setminus \ker(A-\lambda\mathrm{Id}_n)^{m_\lambda - 1}.$$

 (b) On applique $(A-\lambda\mathrm{Id}_E)$ à cette famille, puis on complète

$$\{(A-\lambda\mathrm{Id}_n)x_1, \dots, (A-\lambda\mathrm{Id}_n)x_{\delta_{m_\lambda}}\}$$

 en une famille libre $\mathcal{F}_2$ de $d_{m_\lambda-1} - d_{m_\lambda-2}$ vecteurs de

$$\ker(A-\lambda\mathrm{Id}_n)^{m_\lambda} \setminus \ker(A-\lambda\mathrm{Id}_n)^{m_\lambda - 1}.$$

 (c) On itère ce procédé, jusqu'à obtenir ainsi une base de F_λ en prenant la réunion : $\cup_{i=1}^{m_\lambda}\mathcal{F}_i$. Dans cette base on obtient la forme de Jordan désirée.

Exercice IV.3.3. Donner la forme de Jordan de la matrice

$$A = \begin{bmatrix} -2 & 2 & -4 & 9 \\ -2 & 2 & -3 & 6 \\ 1 & -1 & 2 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Solution IV.3.4. On commence par calculer

$$\chi_A(T) = T(T-1)^3.$$

Le lemme des noyaux donne

$$\mathbb{R}^4 = F_1 \oplus F_0$$

où F_λ est l'espace caractéristique associé à la valeur propre λ, de dimension $n_\lambda = \mathrm{mult}(\lambda, \chi_A)$. On en déduit que A est semblable à l'une des matrices suivantes selon la valeur de $m_1 = \mathrm{mult}(1, \mu_A)$:

$$m_1 = 1 : \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, m_1 = 2 : \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, m_1 = 3 : \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Il s'agit alors d'identifier m_1. On calcule

$$A - \mathrm{Id}_4 = \begin{bmatrix} -3 & 2 & -4 & 9 \\ -2 & 1 & -3 & 6 \\ 1 & -1 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

et on vérifie que $\dim(\ker(A-\mathrm{Id}_4)) = 2 < \dim(F_1)$, donc $m_1 \geqslant 2$ (car A n'est pas diagonalisable). On calcul alors

$$(A-\mathrm{Id}_4)^2 = \begin{bmatrix} 1 & 0 & 2 & -3 \\ 1 & 0 & 2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

de rang 1, et donc de noyau de dimension $3 = \dim(F_1)$. On a donc $F_1 = \ker((A - \mathrm{Id}_4)^2)$ et $m_1 = 2$. Pour déterminer une matrice de changement de base P, on commence par déterminer

$$E_0 = \mathrm{Vect}\,\{v_4 = (1,1,0,0)\},$$

puis

$$E_1 = \{(-2z + 3t, -z, z, t), (z,t) \in \mathbb{R}^2\}.$$

On choisit alors $v_3 \in \ker((A - \mathrm{Id}_4)^2)\backslash\ker(A - \mathrm{Id}_4)$, par exemple $v_3 = (1,0,1,1)$, on pose $v_2 = (A - \mathrm{Id}_4)v_3 = (2,1,-1,0)$, et enfin on complète $\{v_2\}$ en une base de $\ker(A - \mathrm{Id}_4)$ par $v_1 = (3,0,0,1)$. On a alors une base $\mathcal{B} = \{v_1, \ldots, v_4\}$ et une matrice $P = P_{\mathcal{B}_{can}}^{\mathcal{B}}$:

$$P = \begin{bmatrix} 3 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

telles que

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

Exercice IV.3.5. Montrer que si $A \in \mathrm{M}_n(\mathbb{C})$, alors A est semblable à sa transposée.

4. Décomposition de Dunford

Après avoir déterminé les classes de similitudes des matrices à l'aide de la réduction de Jordan, nous allons nous intéresser à une autre décomposition utile, celle de Dunford. Ici, nous n'allons pas chercher une nouvelle base, mais décomposer directement la matrice (ou l'endomorphisme) de départ.

On utilisera les deux lemmes suivants :

Lemme IV.4.1. *Soient $(u,v) \in \mathrm{End}_{\mathbb{K}}(E)$ deux endomorphismes diagonalisables qui commutent. Alors u et v sont simultanément diagonalisables.*

Démonstration. Comme u est diagonalisable,

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} E_\lambda,$$

où $E_\lambda = \ker(u - \lambda \mathrm{Id}_E)$. De plus, pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, E_λ est stable par v car u et v commutent. La restriction de v à E_λ est encore diagonalisable (son polynôme minimal divise celui de v qui est simplement scindé par hypothèse). Si pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$ on fixe une base $\mathcal{B}_\lambda$ de E_λ qui rend $[v_{|E_\lambda}]$ diagonale, on obtient une base $\mathcal{B} = \bigcup_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \mathcal{B}_\lambda$ dans laquelle $[u]_{\mathcal{B}}$ et $[v]_{\mathcal{B}}$ sont diagonales. □

Remarque IV.4.2. Notez que ce lemme implique que la somme de deux endomorphismes diagonalisables qui commutent est diagonalisable. En effet, il suffit de considérer une base dans laquelle les deux endomorphismes en question sont représentés par des matrices diagonales, la somme de ces matrices est encore diagonale.

Lemme IV.4.3. *Soient $(u,v) \in \mathrm{End}_{\mathbb{K}}(E)$ deux endomorphismes nilpotents qui commutent. Alors $u + v$ est nilpotent.*

DÉMONSTRATION. Comme u et v commutent, la formule du binôme de Newton est valable. On pose alors m le maximum des indices de nilpotence de u et v, et on calcule

$$(u+v)^{2m} = \sum_{i=0}^{2m} \binom{2m}{i} u^i\, v^{2m-i} = 0$$

car pour $i \geqslant m$, $u^i = 0$, et pour $i \leqslant m$, $v^{2m-i} = v^{m-i+m} = 0$. □

THÉORÈME IV.4.4 (Décomposition de Dunford). *Soit $u \in \mathrm{End}_{\mathbb{K}}(E)$ un endomorphisme scindé. Alors il existe un unique couple (d, v) tel que*

(i) d est diagonalisable et v nilpotent,

(ii) d et v commutent,

(iii) $u = d + v$.

De plus, si $E = \bigoplus_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(E)} F_\lambda$ est la décomposition de E en sous-espaces caractéristiques de u, et $(\pi_\lambda : E \to F_\lambda)_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)}$ les projections associées alors

$$d = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} \lambda \pi_\lambda \tag{14}$$

$$v = u - d = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} (u - \lambda \mathrm{Id}_E) \circ \pi_\lambda \tag{15}$$

et v est nilpotent d'indice $\max\{\mathrm{mult}(\lambda, \mu_u), \lambda \in \mathrm{Sp}_{\mathbb{K}}(u)\}$.

Bien entendu, cet énoncé (et les deux précédents) admettent des versions matricielles.

REMARQUE IV.4.5. Notons que la réciproque de ce résultat est vrai et plus facile à voir : s'il existe un couple $(d, v) \in \mathrm{End}_{\mathbb{K}}(E)$ qui vérifie les points $(i) - (iii)$ du Théorème IV.4.4, alors u est scindé.

DÉMONSTRATION. Pour prouver l'existence, on pose d et v comme dans l'énoncé (i.e. définis par les formules (14) et (15)). Notez que π_λ est un polynôme en u par le lemme des noyaux, donc $(v, d) \in \mathbb{K}[u]^2$, et v et d commutent, ce qui prouve (ii). Le point (iii) est claire par construction. Pour (i), d est diagonalisable car somme d'endomorphismes diagonalisables qui commutent deux à deux ($\pi_{\lambda_1} \circ \pi_{\lambda_2} = \pi_{\lambda_2} \circ \pi_{\lambda_1} = 0$). Enfin, v est bien nilpotent d'indice $m = \max\{\mathrm{mult}(\lambda, \mu_u), \lambda \in \mathrm{Sp}_{\mathbb{K}}(u)\}$. En effet, comme pour tout $(\lambda, \lambda_1, \lambda_2) \in \mathrm{Sp}_{\mathbb{K}}(u)^3$,

— $\pi_{\lambda_1} \circ \pi_{\lambda_2} = 0$ si $\lambda_1 \neq \lambda_2$,

— $\pi_\lambda^2 = \pi_\lambda$,

— $\pi_\lambda \in \mathbb{K}[u]$,

on a

$$v^m = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)} (u - \lambda \mathrm{Id}_E)^m \circ \pi_\lambda.$$

Comme pour tout $\lambda \in \mathrm{Sp}_{\mathbb{K}}(u)$, $(u - \lambda \mathrm{Id}_E)^m \circ \pi_\lambda$ est la restriction de $(u - \lambda \mathrm{Id}_E)^m$ à F_λ, et comme $m \geqslant \mathrm{mult}(\lambda, \mu_u)$, $(u - \lambda \mathrm{Id}_E)^m \circ \pi_\lambda = 0$. D'où $v^m = 0$.

Enfin, pour l'unicité, si (v', d') est un autre couple qui vérifie $(i) - (iii)$, on constate que v' commute à $(v' + d')$ (v' et d' commutent), donc commute avec u et donc avec tout polynôme en u. D'où v' commute avec v. De même d' commute avec d. Mais alors

$$d + v = d' + v'$$

implique $d - d' = v' - v$. Comme ces endomorphismes commutent, par le Lemme IV.4.3 $v' - v$ est nilpotent et par le Lemme IV.4.1 $d - d'$ est diagonalisable. Un endomorphisme à la fois diagonalisable et nilpotent est nul, d'où $d = d'$ et $v = v'$. □

Remarque IV.4.6. Attention à bien vérifier que les deux endomorphismes obtenus commutent. Étudiez l'exemple suivant :

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

Alors $A = D + N$, où

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, N = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix},$$

D est diagonale et N nilpotente. Cependant ce n'est pas la décomposition de Dunford de A : $DN \neq ND$. Il est clair que A est diagonalisable (son polynôme caractéristique est simplement scindé), donc sa décomposition de Dunford est simplement $A = A + 0$, et ici la matrice nulle joue le rôle de la partie nilpotente.

Exercice IV.4.7. Déterminer la décomposition de Dunford de la matrice

$$A = \begin{bmatrix} 5 & -1 & -1 \\ 0 & 6 & 3 \\ -2 & -2 & 1 \end{bmatrix}.$$

Solution IV.4.8. On calcul

$$\chi_A(T) = (T-3)^2(T-6).$$

La dimension de $\ker(A - 3\mathrm{Id}_3) = 1$, on a donc $E_3 \neq F_3$ (car F_3 est de dimension $\mathrm{mult}(3, \chi_A) = 2$) et A n'est pas diagonalisable. On cherche alors une décomposition de Dunford, et pour cela il nous faut les projecteurs π_λ sur les espaces caractéristiques F_λ. On sait que $\mu_A(T) = \chi_A(T)$ (car A n'est pas diagonalisable donc $\mu_A(T) \neq (T-3)(T-6)$). On calcul alors

$$\frac{1}{\chi_A(T)} = \frac{1}{9}\left(\frac{1}{T-6} - \frac{T}{(T-3)^2}\right).$$

De là on a

$$\mathrm{Id}_3 = \frac{1}{9}(A - 3\mathrm{Id}_3)^2 - \frac{1}{9}A(A - 6\mathrm{Id}_3),$$

d'où

$$\pi_6 = \frac{1}{9}(A - 3\mathrm{Id}_3)^2$$

et

$$\pi_3 = -\frac{1}{9}A(A - 6\mathrm{Id}_3).$$

On calcul alors (noter qu'ici $\pi_6 = \mathrm{Id}_3 - \pi_3$, ce qui simplifie les calculs)

$$D = 3\pi_3 + 6\pi_6 = 6\mathrm{Id}_3 - 3\pi_3 = 6\mathrm{Id}_3 - 2A + \frac{A^2}{3}$$

et donc

$$D = \begin{bmatrix} 5 & -1 & -1 \\ -2 & 4 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

est diagonalisable,

$$N = A - D = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 2 & 2 \\ -2 & -2 & -2 \end{bmatrix}$$

est nilpotente, et $A = D + N$ est la décomposition de Dunford de A.

Chapitre V

Applications

On va mettre en pratique les résultats des chapitres précédents. Dans ce chapitre on suppose que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Puissances et exponentielle de matrices

Soit $A \in \mathrm{M}_n(\mathbb{C})$.

1.1. Puissances. On est intéressé par le calcul de A^m pour $m \in \mathbb{N}^*$. Notons que le nombre d'opérations nécessaires pour calculer naïvement le produit AB de deux matrices carrées A et B de taille n est de n^3 produits et $n^2(n-1)$ sommes de complexes. Le calcul de A^m nécessiterait donc $(m-1)n^3$ produits et $(m-1)(n-1)n^2$ sommes. On va voir ici deux méthodes qui simplifient les calculs.

Méthode 1 : On peut utiliser la division euclidienne de T^m par un polynôme annulateur de A, par exemple $\chi_A(T)$ ou $\mu_A(T)$. Idéalement, on choisira un polynôme annulateur de petit degré pour simplifier la division euclidienne. Par exemple, on effectue la division euclidienne de T^m par $\mu_A(T)$: il existe un couple $(q, r) \in \mathbb{C}[T]$ tel que

$$T^m = q(T)\mu_A(T) + r(T)$$

avec $\deg(r) < \deg(\mu_A)$. Notez que pour calculer r, on peut évaluer en les valeurs propres ; pour tout $\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)$,

$$\lambda^m = q(\lambda)\mu_A(\lambda) + r(\lambda) = r(\lambda).$$

Si toutes les valeurs propres sont distinctes, on obtient autant d'équations ($\sharp\mathrm{Sp}_{\mathbb{C}}(A) = \deg(\mu_A)$ équations) que d'inconnues (nombre de coefficients inconnus de r est au plus $\deg(r) + 1 \leqslant \deg(\mu_A)$). On peut alors identifier r par résolution d'un système linéaire. Si μ_A dispose de racines multiples, on peut dériver l'équation $T^m = q(T)\mu_A(T) + r(T)$ par rapport à T et évaluer en ces valeurs propres. Par exemple, si λ est valeur propre double, on a

$$m\lambda^{m-1} = q'(\lambda)\mu_A(\lambda) + q(\lambda)\mu_A'(\lambda) + r'(\lambda) = r'(\lambda)$$

car $\mu_A(\lambda) = \mu_A'(\lambda) = 0$ comme λ est valeur propre double. On obtient de la même façon autant d'équations que d'inconnues, et on peut déterminer r à partir de la connaissance des puissances des valeurs propres.

Le calcul de puissance pour A donne alors :

$$A^m = q(A)\mu_A(A) + r(A) = r(A)$$

car μ_A est un polynôme annulateur. Si $m \geqslant \deg(\mu_A)$ les puissances de matrices à calculer dans $r(A)$ sont inférieures à m, et on obtient un calcul plus simple (modulo le calcul de μ_A et de ses racines les valeurs propres).

Méthode 2 : Supposons que l'on dispose de la décomposition de Dunford pour A : $A = D + N$ avec $DN = ND$, D diagonalisable et N nilpotente d'indice p. On commence alors par diagonaliser D : il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $P^{-1}DP = \Delta$, où $\Delta = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ est diagonale et les (λ_i) sont les valeurs propres de D (et A). On a alors, en posant $N' = P^{-1}NP$

$$\begin{array}{rcc} A^m & = & (D+N)^m \\ & = & (P(P^{-1}DP + P^{-1}NP)P^{-1})^m \\ & = & P(\Delta + N')^m P^{-1} \\ & = & P\sum_{i=0}^m \binom{m}{i}\Delta^i (N')^{m-i}. \end{array}$$

Ici, le binôme de Newton peut être utilisé car Δ et N' commutent comme D et N commutent. Dans ce binôme, les puissance de Δ sont facilement calculées :

$$\Delta^i = \mathrm{Diag}(\lambda_1^i, \ldots, \lambda_n^i).$$

Comme N est nilpotente d'indice p, N' l'est aussi, et seule les $p-1$ premières puissances de N' interviennent.

1.2. Exponentielle de matrices. On va définir $\exp(A)$, qui sera utilisée dans la section suivante pour résoudre des systèmes linéaires d'équations différentielles. La formule est la suivante :

$$\exp(A) = \sum_{k=0}^{+\infty} \frac{A^k}{k!} = \lim_{p \to +\infty} \sum_{k=0}^{p} \frac{A^k}{k!} \in \mathrm{M}_n(\mathbb{C}).$$

On admet le résultat suivant (notez que l'existence, et même le sens, de la limite ci-dessus n'est pas évident) :

Théorème V.1.1. *L'exponentielle de matrice est bien définie.*

On va voir comment calculer cette exponentielle en pratique, ainsi que quelques unes de ses propriétés.

Exemples V.1.2. Si $A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$,

$$\exp(A) = \mathrm{Diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

Si A est nilpotente d'indice m,

$$\exp(A) = \mathrm{Id}_n + A + \frac{A^2}{2!} + \ldots + \frac{A^{m-1}}{(m-1)!}.$$

Si A est triangulaire supérieure de la forme

$$A = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

alors $\exp(A)$ est triangulaire supérieure de la forme

$$\exp(A) = \begin{bmatrix} e^{\lambda_1} & * & * & * \\ 0 & e^{\lambda_2} & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & e^{\lambda_n} \end{bmatrix}.$$

Lemme V.1.3. *Pour tout* $P \in \mathrm{GL}_n(\mathbb{C})$,

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

Démonstration. Pour tout $k \in \mathbb{N}^*$,

$$(P^{-1}AP)^k = P^{-1}A^kP.$$

On en déduit que pour les sommes partielles :

$$\sum_{k=0}^{p} \frac{(P^{-1}AP)^k}{k!} = P^{-1}\sum_{k=0}^{p}\frac{A^k}{k!}P.$$

Le résultat suit par passage à la limite (on admet ce point). □

Remarque V.1.4. Ce lemme permet de définir l'exponentielle d'un endomorphisme à l'aide de sa représentation matricielle.

Corollaire V.1.5. *Soit* $A \in \mathrm{M}_n(\mathbb{C})$*, alors* $\exp(A)$ *est inversible, et*

$$\det(\exp(A)) = e^{\mathrm{tr}(A)}.$$

Démonstration. La matrice A est trigonalisable : il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que

$$P^{-1}AP = B := \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

Le lemme précédent donne

$$\exp(A) = P\exp(B)P^{-1} = P\begin{bmatrix} e^{\lambda_1} & * & * & * \\ 0 & e^{\lambda_2} & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & e^{\lambda_n} \end{bmatrix}P^{-1}.$$

On en déduit

$$\det(\exp(A)) = \Pi_{i=1}^n e^{\lambda_i} = \exp(\sum_{i=1}^n \lambda_i) \neq 0.$$

On en déduit que $\exp(A)$ est inversible. De plus, comme $\mathrm{tr}(A) = \mathrm{tr}(B) = \sum_{i=1}^n \lambda_i$, le résultat suit. □

On admet le résultat suivant :

Théorème V.1.6. *Soit* $(A, B) \in \mathrm{M}_n(\mathbb{C})^2$*. On suppose* A *et* B *commutent. Alors*

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

Remarque V.1.7. Attention, le résultat n'est pas vrai en général si A et B ne commutent pas. Ainsi, pour

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

A et B ne commutent pas et $\exp(A)\exp(B) \neq \exp(B)\exp(A)$.

Corollaire V.1.8. *Soit* $A \in \mathrm{M}_n(\mathbb{C})$*. Alors*

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

On a également le corollaire suivant qui permet de calculer en pratique une exponentielle de matrice à l'aide de la décomposition de Dunford :

Corollaire V.1.9. *Soit $A = D + N$ la décomposition de Dunford de A, où $DN = ND$, D diagonalisable et N nilpotente d'indice m. Alors*

$$\exp(A) = \exp(D)(\sum_{j=0}^{m-1} \frac{N^j}{j!}).$$

Dans la formule ci-dessus, on calcule $\exp(D)$ en déterminant $P \in \mathrm{GL}_n(\mathbb{C})$ telle que

$$P^{-1}DP = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n),$$

et on obtient

$$\exp(D) = P\mathrm{Diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})P^{-1}.$$

2. Suites récurrentes linéaires

Définition V.2.1. On appelle suite récurrente linéaire d'ordre n toute suite $(x_k)_{k\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ définie par une “relation de récurrence linéaire” de la forme :

$$\forall k \in \mathbb{N},\ x_{n+k} = a_{n-1}x_{n-1+k} + \ldots + a_0 x_k \tag{16}$$

où $(a_0, \ldots, a_{n-1}) \in \mathbb{K}^n$ est fixé.

Une telle suite est donc entièrement déterminée par la donnée de $(x_0, \ldots, x_{n-1}) \in \mathbb{K}^n$ et de la relation de récurrence linéaire.

Exemples V.2.2. Les suites récurrentes linéaires d'ordre 1 sont les suites géométriques. La suite de Fibonacci, définie par $f_0 = 0$, $f_1 = 1$, $f_{n+2} = f_{n+1} + f_n$ est une suite récurrente linéaire d'ordre 2.

On va se ramener à une présentation matricielle. On fixe $(x_k)_{k\in\mathbb{N}}$ une telle suite et on pose

$$\forall k \geqslant 1,\ X_k = \begin{bmatrix} x_k \\ x_{k+1} \\ \vdots \\ \vdots \\ x_{k+n-1} \end{bmatrix},\ X_0 = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \vdots \\ x_{n-1} \end{bmatrix},\ A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \\ a_0 & a_1 & \ldots & \ldots & a_{n-1} \end{bmatrix}.$$

On remarque alors que $(x_k)_{k\in\mathbb{N}}$ est déterminée par

$$\forall k \geqslant 0,\ X_{k+1} = AX_k$$

et donc

$$\forall k \geqslant 0,\ X_k = A^k X_0.$$

On peut alors déterminer $(x_k)_{k\in\mathbb{N}}$ via un calcul de puissances de matrice comme dans la Section 1.1 de ce chapitre.

Remarque V.2.3. On peut démontrer que l'espace $\mathcal{S}$ des suites définies par la relation (16) est un $\mathbb{K}$-espace vectoriel de dimension n. Si on pose

$$u : \begin{array}{ccc} \mathbb{K}^{\mathbb{N}} & \to & \mathbb{K}^{\mathbb{N}} \\ (x_k)_{k\in\mathbb{N}} & \mapsto & (x_{k+1})_{k\in\mathbb{N}} \end{array}$$

et

$$p(T) = T^n - a_{n-1}T^{n-1} - \ldots - a_1 T - a_0$$

on remarque que

$$\mathcal{S} = \ker(p(u)).$$

On peut alors déterminer une base de $\mathcal{S}$ on décomposant p en produit de facteurs deux à deux premiers entre eux. En effet, si $p = \Pi_{i=1}^{r}(T - r_i)^{m_i}$ est une telle décomposition, le lemme des noyaux [1] donne alors

$$\mathcal{S} = \bigoplus_{i=1}^{r} \ker(u - r_i \mathrm{Id})^{m_i}.$$

On peut alors vérifier que les suites

$$\left\{(k^j r_i^k)_{k\in\mathbb{N}}, j \in \{0, \ldots, m_i - 1\}\right\}$$

forment une base de $\ker(u - r_i\mathrm{Id})^{m_i}$, et on obtient une base de $\mathcal{S}$ par réunion.

Exercice V.2.4. Déterminer une base de l'espace des suites définies par une relation de récurrence linéaire d'ordre 2 sur $\mathbb{C}$.

3. Résolutions de systèmes différentiels linéaires

On pose

$$\mathcal{E} := \mathcal{C}^\infty(\mathbb{R}, \mathbb{K}).$$

C'est un espace vectoriel de dimension infinie sur $\mathbb{K}$.

Définition V.3.1. On appelle système différentiel linéaire à coefficients constants et à n inconnues un système de la forme

$$\left\{\begin{array}{lclcl} x_1'(t) & = & a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) & + & b_1(t) \\ x_2'(t) & = & a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) & + & b_2(t) \\ \vdots & & & & \\ x_n'(t) & = & a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) & + & b_n(t) \end{array}\right. \tag{17}$$

où les coefficients

$$(a_{ij})_{1\leqslant i,j\leqslant n} \in \mathbb{K}^{n^2}$$

sont des constantes, le "second membre"

$$(b_1(t), \ldots, b_n(t)) \in \mathcal{E}^n,$$

et les "inconnues"

$$(x_1(t), \ldots, x_n(t))$$

sont recherchées dans $\mathcal{E}$. Une "donnée initiale" pour le système (17) est la donnée de $t_0 \in \mathbb{R}$ et $x_i(t_0) = x_{i,0} \in \mathbb{K}$ pour $i \in [\![1, n]\!]$.

1. Notez ici que l'on utilise les points (1) et (2) du lemme des noyaux, qui se généralisent directement à la dimension infinie.

Exemple V.3.2. Pour $n = 1$, on obtient simplement l'équation différentielle du premier ordre :

$$x'(t) = ax(t) + b(t).$$

On rappelle que l'équation homogène associée est

$$x'(t) = ax(t), \tag{18}$$

de solutions

$$\mathrm{Sol}_H = \{t \mapsto \lambda e^{at}, \lambda \in \mathbb{K}\},$$

et si $f \in \mathcal{E}$ est une solution particulière de (18) (obtenue par exemple via la méthode de variation de la constante), alors l'ensemble des solutions de (18) est

$$\mathrm{Sol} = \{t \mapsto f(t) + \lambda e^{at}, \lambda \in \mathbb{K}\}.$$

La méthode de variation de la constante donne la formule générale suivante pour les solutions de (18) :

$$x(t) = \lambda e^{at} + e^{at} \int_{t_0}^{t} e^{-as} b(s) ds$$

et si on impose la condition initiale $x(t_0) = x_0$, alors l'unique solution est

$$e^{a(t-t_0)} x_0 + e^{at} \int_{t_0}^{t} e^{-as} b(s) ds.$$

On va voir que ces formules se généralisent en dimension supérieure, à l'aide de l'exponentielle de matrices.

On considère le système différentiel (17). On va le reformuler en termes matriciels. On pose

$$A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in \mathrm{M}_n(\mathbb{K}),$$

puis

$$B(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}, X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, X'(t) = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix},$$

vus comme des vecteurs de n fonctions lisses, et enfin la donnée initiale

$$X_0 = \begin{bmatrix} x_{1,0} \\ \vdots \\ x_{n,0} \end{bmatrix}.$$

On appelle alors A la "matrice du système", B le "second membre" et X "l'inconnue". Le système (17) est alors équivalent à l'équation matricielle

$$X'(t) = AX(t) + B(t), \tag{19}$$

avec condition initiale $X(0) = X_0$. On appelle système différentiel linéaire homogène associé à (19) le système

$$X'(t) = AX(t). \tag{20}$$

On utilisera des résultats généraux sur la dérivation de matrices à coefficients dans $\mathcal{E}$. On définit, pour $A(t) = [a_{ij}(t)] \in \mathrm{M}_{m,n}(\mathcal{E})$,

$$A'(t) := [a'_{ij}(t)] \in \mathrm{M}_{m,n}(\mathcal{E}).$$

Lemme V.3.3. *Soient* (X, Y) *deux vecteurs à valeurs dans* $\mathcal{E}$*, et* $\lambda \in \mathbb{K}$*. Alors*

$$(\lambda X(t) + Y(t))' = \lambda X'(t) + Y'(t).$$

Soient $A(t), B(t) \in \mathrm{M}_{p,q}(\mathcal{E})$ *et* $C(t) \in \mathrm{M}_{q,r}(\mathcal{E})$*. Alors*

$$(A + B)'(t) = A'(t) + B'(t)$$

et

$$(BC)'(t) = (B'C)(t) + (BC')(t).$$

Exercice V.3.4. Donner la preuve du lemme précédent.

On admet également le résultat suivant :

Proposition V.3.5. *Soit* $A \in \mathrm{M}_n(\mathbb{K})$*. Alors*

$$(\exp(tA))'(t) = A\exp(tA) = \exp(tA)A.$$

On peut alors démontrer :

Théorème V.3.6. *L'espace des solutions de (*20*) est*

$$\mathrm{Sol}_H = \{t \mapsto \exp((t - t_0)A)X_0,\ X_0 \in \mathbb{K}^n\}.$$

Remarque V.3.7. L'espace des solutions Sol_H est un $\mathbb{K}$-espace vectoriel de dimension n et de base donnée par les fonctions

$$t \mapsto \exp((t - t_0)A)e_i$$

où $(e_i)_{1\leqslant i\leqslant n}$ est la base canonique de $\mathbb{K}^n$.

Démonstration. Avec la Proposition V.3.5 on vérifie aisément que les fonctions $t \mapsto \exp((t - t_0)A)X_0$ sont des solutions de (20). Réciproquement, si $Y(t)$ est une solution de (20), on pose $X = \exp(-(t - t_0)A)Y(t)$ et on calcule

$$\begin{array}{rcl} X'(t) & = & (\exp(-(t - t_0)A))'Y(t) + \exp(-(t - t_0)A)Y'(t) \\ & = & -A\exp(-(t - t_0)A)Y(t) + \exp(-(t - t_0)A)AY(t) \\ & = & -A\exp(-(t - t_0)A)Y(t) + A\exp(-(t - t_0)A)Y(t) \\ & = & 0. \end{array}$$

Le vecteur $X(t)$ est donc constant, $X(t) = X(t_0) := X_0$, et on obtient

$$Y(t) = \exp((t - t_0)A)X_0.$$

□

De là on déduit exactement comme dans le cas $n = 1$:

Théorème V.3.8. *L'espace des solutions de (*19*) est*

$$\mathrm{Sol} = \{t \mapsto F(t) + \exp(tA)X_0,\ X_0 \in \mathbb{K}^n\},$$

où $F(t)$ *est une solution particulière de (*19*). La méthode de variation de la constante donne la formule générale suivante pour les solutions de (*19*) :*

$$X(t) = \exp(tA)X_0 + \exp(tA)\int_{t_0}^{t} \exp(-sA)B(s)ds$$

et si on impose la condition initiale $X(t_0) = X_0$*, alors l'unique solution est*

$$\exp((t - t_0)A)X_0 + \exp(tA)\int_{t_0}^{t} \exp(-sA)B(s)ds.$$

EXERCICE V.3.9. Résoudre le système différentiel suivant :

$$\left\{\begin{array}{rcl} X'(t) & = & AX(t)+B(t) \\ X(0) & = & X_0 \end{array}\right.$$

où

$$A = \begin{bmatrix} 5 & -1 & -1 \\ 0 & 6 & 3 \\ -2 & -2 & 1 \end{bmatrix}, B(t) = \begin{bmatrix} e^t \\ 0 \\ 0 \end{bmatrix}, X_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

REMARQUE V.3.10. Il existe d'autres méthodes pour résoudre un système de la forme (17). Ainsi, si on suppose A trigonalisable, il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP = C = [c_{ij}] \in \mathrm{M}_n(\mathbb{K})$ soit triangulaire supérieure. On se ramène alors au système

$$Z'(t) = CZ(t) + \tilde{B}(t)$$

en posant $Z(t) = P^{-1}X(t)$ et $\tilde{B}(t) = P^{-1}B(t)$. Ce système est équivalent aux équations suivantes :

$$\left\{\begin{array}{ccccccccccccc} z_1'(t) & = & c_{11}z_1(t) & + & c_{12}z_2(t) & + & \dots & + & \dots & + & c_{1n}z_n(t) & + & \tilde{b}_1(t) \\ z_2'(t) & = & 0 & + & c_{22}z_2(t) & + & \dots & + & \dots & + & c_{2n}z_n(t) & + & \tilde{b}_2(t) \\ \vdots & & & & & & & & & & & & \\ z_{n-1}'(t) & = & 0 & + & \dots & + & 0 & + & c_{n-1n-1}z_{n-1}(t) & + & c_{n-1n}z_n(t) & + & \tilde{b}_{n-1}(t) \\ z_n'(t) & = & 0 & + & \dots & + & \dots & + & 0 & + & c_{nn}z_n(t) & + & \tilde{b}_n(t) \end{array}\right.$$

que l'on peut résoudre successivement en partant de la dernière ligne.

Bibliographie

[LFA] Jacqueline Lelong–Ferrand et Jean–Marie Arnaudiès, *Cours de Mathématiques, Tome 1, Algèbre.* Dunod.

[MB] Michèle Benyounes, *Réduction des endomorphismes*, notes de cours

[MM] Roger Mansuy et Racher Mneimné, *Alèbre Linéaire - Réduction des endomorphismes.* Vuibert, 2012.

[RW1] Xavier Buff, Emmanuel Harlberstadt, François Moulin, Monique Ramis et Jacques Sauloy, *Mathématiques, tout-en-un pour la Licence 1*, sous la direction de J.-P. Ramis et A. Warusfel. Niveau L1. Dunod, 2007.

[RW2] Xavier Buff, Emmanuel Harlberstadt, François Moulin, Monique Ramis et Jacques Sauloy, *Mathématiques, tout-en-un pour la Licence 2*, sous la direction de J.-P. Ramis et A. Warusfel. Niveau L2. Dunod, 2007.