SORBONNE UNIVERSITÉ

Licence de physique intensive Sprint

Méthodes mathématiques II

Algèbre linéaire

Dan Israël (israel@lpthe.jussieu.fr)

František Kupka, *Etude pour Plans verticaux III* (1912), Centre Pompidou.

Ces notes de cours sont fondées sur le cours de « Méthodes mathématiques II » dispensé par l'auteur à Sorbonne Université dans le cadre de la licence de physique intensive Sprint lors de l'année universitaire 2020–2021. Cet enseignement a pour but d'approfondir les concepts de base d'algèbre linéaire qui ont été exposés en première année.

Dernière mise à jour

Le 28 février 2023

Table des matières

Cours n°1

Matrices et applications linéaires

L'objectif de ce chapitre, composé en grande partie de rappels, est de définir les objets de base utilisés dans les chapitres suivants ainsi que leurs propriétés principales.

1.1 Espaces vectoriels, sous-espaces vectoriels

Définition 1 (espace vectoriel). *Soit* E *un ensemble, non vide, muni d'une opération interne appelée addition,*

$$+ \quad \begin{cases} \mathsf{E} \times \mathsf{E} & \to \mathsf{E} \\ (\vec{u}, \vec{v}) & \mapsto \quad \vec{u} + \vec{v} \end{cases} \tag{1.1}$$

et d'une opération externe appelée multiplication par un scalaire,

$$* \quad \begin{cases} \mathsf{K} \times \mathsf{E} & \to \mathsf{E} \\ (\lambda, \vec{u}) & \mapsto \quad \lambda * \vec{u} \end{cases} \tag{1.2}$$

où K *désigne le corps des réels* $\mathbb{R}$ *ou le corps des complexes* $\mathbb{C}$. *Le triplet* $(\mathsf{E}, +, *)$ *forme un* K*-espace vectoriel si les propriétés suivantes sont satisfaites :*

1. *L'addition sur* E *est une opération :*
   a) *commutative, c.-à.-d. que pour tout* $(\vec{u}, \vec{v}) \in \mathsf{E}^2$, $u + v = v + u$
   b) *associative, c.-à.-d. que pour tout* $(\vec{u}, \vec{v}, \vec{w}) \in \mathsf{E}^3$, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
   c) *qui admet un élément neutre, noté* $\vec{0}_\mathsf{E}$ *et appelé vecteur nul, tel que pour tout* $\vec{u} \in \mathsf{E}$, $\vec{u} + \vec{0}_\mathsf{E} = \vec{0}_\mathsf{E} + \vec{u} = \vec{u}$
   d) *telle que tout élément* $\vec{v} \in \mathsf{E}$ *admet un opposé dans* E, *noté* $-\vec{v}$, *tel que* $\vec{v} + (-\vec{v}) = \vec{0}_\mathsf{E}$.
2. *La multiplication externe satisfait, pour tout* $\vec{v} \in \mathsf{E}$ *:*
   a) $1 * \vec{v} = \vec{v}$
   b) $\lambda * (\mu * \vec{v}) = (\lambda\mu) * \vec{v}$
   c) $(\lambda + \mu) * \vec{v} = \lambda * \vec{v} + \mu * \vec{v}$

3. *La multiplication est distributive sur l'addition, c.-à.-d. que*
   a) *Pour tout* $(\vec{u},\vec{v}) \in E^2$ *et* $\lambda \in K$, $\lambda * (\vec{u}+\vec{v}) = \lambda * \vec{u} + \mu * \vec{v}$
   b) *Pour tout* $u \in E$ *et* $(\lambda,\mu) \in K^2$, $(\lambda+\mu) * \vec{u} = \lambda * \vec{u} + \mu * \vec{u}$.

Exemple 1 Les corps $\mathbb{R}$ et $\mathbb{C}$ munis de l'addition et de la multiplication usuelles sont eux-mêmes respectivement un $\mathbb{R}$-espace vectoriel et un $\mathbb{C}$-espace vectoriel ; il en est de même pour leurs produit cartésiens $\mathbb{R}^n$ et $\mathbb{C}^n$ pour tout $n \in \mathbb{N}^*$ *(exercice : le vérifier)*. D'autres exemples plus abstraits d'espaces vectoriels sont donnés par :
— l'espace $K[X]$ des polynômes à coefficients à valeurs dans K (voir plus loin) ;
— l'espace des applications continues sur un intervalle $[a,b] \subset \mathbb{R}$, à valeurs dans K ;
— l'espace des solutions d'une équation différentielle, comme $f'' = af' + b$;
— l'espace des suites à valeurs dans K.

*
* *

Propriété 1 *Dans un* K*-espace vectoriel* $(E,+,*)$ *les propriétés suivantes sont satisfaites :*

1. *Pour tout* $\lambda \in K$, $\lambda * \vec{0}_E = \vec{0}_E$
2. *Pour tout* $\vec{v} \in E$, $0 * \vec{v} = \vec{0}_E$
3. *Si* $\lambda * \vec{v} = \vec{0}_E$ *alors* $\lambda = 0$ *ou* $\vec{v} = \vec{0}_E$
4. *Pour tout* $\vec{v} \in E$, $(-1) * \vec{v} = -\vec{v}$
5. *Pour tout* $(\vec{v}_1,\dots,\vec{v}_n) \in E^n$ *et pour tout* $(\lambda_1,\dots,\lambda_n) \in K^n$, $\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n \in E$.

Les démonstrations découlent directement de la définition 1. En effet,

1. Soit $\lambda \in K$ non nul. Pour tout $\vec{v} \in E$ il existe $\vec{w}$ tel que $\vec{v} = \lambda * \vec{w}$ (soit $w = (1/\lambda) * \vec{v}$). Par distributivité puis par définition du vecteur nul, $\vec{v} + \lambda * \vec{0}_E = \lambda * (\vec{w} + \vec{0}_E) = \lambda * \vec{w} = \vec{v}$. De même on a $\lambda * \vec{0}_E + \vec{v} = \lambda * (\vec{0}_E + \vec{w}) = \lambda * \vec{w} = \vec{v}$.
2. Soit $\vec{v} \in E$. Nous avons par distributivité $\vec{v} + 0 * \vec{v} = (1+0) * \vec{v} = 1 * \vec{v} = \vec{v}$. De même $0 * \vec{v} + \vec{v} = (0+1) * \vec{v} = 1 * \vec{v} = \vec{v}$.
3. Supposons $\lambda \neq 0$ et $\vec{v} \neq \vec{0}_E$. Nous avons d'une part $(1/\lambda) * (\lambda * \vec{v}) = (1/\lambda) * \vec{0}_E = \vec{0}_E$ d'après la propriété 1. D'autre part, $(1/\lambda) * (\lambda * \vec{v}) = (\lambda/\lambda) * \vec{v} = \vec{v}$; on en déduit que $\vec{v} = \vec{0}_E$, ce qui est contraire à l'hypothèse.
4. Nous avons $(-1) * \vec{v} + \vec{v} = (-1) * \vec{v} + 1 * \vec{v} = (-1+1) * \vec{v} = 0 * \vec{v} = 0_E$ et de même $\vec{v} + (-1) * \vec{v} = (1-1) * \vec{v} = 0 * \vec{v} = 0_E$.
5. La preuve se fait par récurrence. Soit $(v_1,v_2) \in E^2$ et $(\lambda_1,\lambda_2) \in K^2$. Nous avons premièrement $\vec{w}_1 = \lambda_1\vec{v}_1 \in E$ et $\vec{w}_2 = \lambda_2\vec{v}_2 \in E$ par définition de la multiplication externe. Nous avons ensuite $\vec{w}_1 + \vec{w}_2 \in E$ par définition de la loi de composition interne.

□

Définition 2 (sous-espace vectoriel). *Soit* $(\mathsf{E}, +, *)$ *un* K*-espace vectoriel et* $\mathsf{F} \subset \mathsf{E}$ *une partie de* E *(au sens des ensembles), non vide.* F *est un sous-espace vectoriel de* E *(donc en particulier un espace vectoriel en lui-même) si les conditions suivantes sont satisfaites) :*

1. *Pour tout* $(u, v) \in \mathsf{F}^2$*,* $\vec{u} + \vec{v} \in \mathsf{F}$
2. *Pour tout* $\lambda \in \mathsf{K}$ *et pour tout* $\vec{v} \in \mathsf{F}$*,* $\lambda * \vec{v} \in \mathsf{F}$*.*

Pour montrer que F est un espace vectoriel on remarque premièrement que F « hérite » des opérations $(+, *)$ de l'espace vectoriel E avec leurs propriétés (associativité, distributivité, etc.). On remarque ensuite que :

— Soit $\vec{v} \in \mathsf{F}$. Nous avons d'une part $0 * \vec{v} \in \mathsf{F}$ et d'autre part $0 * \vec{v} = \vec{0}_\mathsf{E}$ donc $\vec{0}_\mathsf{E} \in \mathsf{F}$ et le vecteur nul de E est aussi vecteur nul de F ;

— Soit $\vec{v} \in \mathsf{F}$. Nous avons d'une part $(-1) * \vec{v} \in \mathsf{F}$ et d'autre part $(-1) * \vec{v} = -\vec{v}$ donc $-\vec{v} \in \mathsf{F}$ et l'opposé de tout vecteur de F est dans F ;

— Soit $(\vec{v}_1, \vec{v}_2) \in \mathsf{F}^2$ et $(\lambda_1, \lambda_2) \in \mathsf{K}^2$. Nous avons $w_1 = \lambda_1 * \vec{v}_1 \in \mathsf{F}$ et $w_2 = \lambda_2 * \vec{v}_2 \in \mathsf{F}$ puis $\vec{w}_1 + \vec{w}_2 \in \mathsf{F}$. Le sous-espace F est donc stable par combinaison linéaire.

□

Exemple 2 Soit $\mathsf{K}[\mathsf{X}]$ l'espace vectoriel des polynômes à coefficients dans K, dont les vecteurs sont de la forme $\mathsf{P} = a_0 + a_1\mathsf{X} + a_2\mathsf{X}^2 + \cdots a_k\mathsf{X}^k + \cdots + a_n\mathsf{X}^n$ avec $a_k \in \mathsf{K}$. Le *degré* n d'un polynôme P est le plus petit entier naturel tel que, pour tout $r \in \mathbb{N}^*$, $a_{n+r} = 0$. On note que :

1. L'ensemble des polynômes de degré *exactement* égal à n n'est pas un sous-espace vectoriel de $\mathsf{K}[\mathsf{X}]$. Par exemple, si $\mathsf{P}_1 = \mathsf{X} + \mathsf{X}^n$ et $\mathsf{P}_2 = \mathsf{X} - \mathsf{X}^n$ sont de degré n, $\mathsf{P}_1 + \mathsf{P}_2 = 2\mathsf{X}$ est de degré 1.
2. L'ensemble des polynôme de degré inférieur ou égal à n, noté $\mathsf{K}_n[\mathsf{X}]$, est un sous-espace vectoriel de $\mathsf{K}[\mathsf{X}]$. Soit en effet $\mathsf{P}_1 = a_0 + \cdots + a_n\mathsf{X}^n$ et $\mathsf{P}_2 = b_0 + \cdots + b_n\mathsf{X}^n$. Nous avons $\mathsf{P}_1 + \mathsf{P}_2 = (a_0 + b_0) + \cdots + (a_n + b_n)\mathsf{X}^n \in \mathsf{K}_n[\mathsf{X}]$. Pour tout $\lambda \in \mathsf{K}$, $\lambda * \mathsf{P}_1 = (\lambda a_0) + \cdots + (\lambda a_n)\mathsf{X}^n \in \mathsf{K}_n[\mathsf{X}]$.

*
* *

Définition 3 (somme de sous-espaces vectoriels). *Soient* F *et* G *deux sous-espaces vectoriels de* $(\mathsf{E}, +, *)$*. La somme* $\mathsf{F} + \mathsf{G}$*, définie par*

$$\mathsf{F} + \mathsf{G} \overset{\text{déf.}}{=} \left\{ \vec{v} + \vec{w} \text{ t.q. } \vec{v} \in \mathsf{F}, w \in \mathsf{G} \right\} \tag{1.3}$$

est un sous-espace vectoriel de E*.*

Soient $(\vec{v}_1, \vec{v}_2) \in \mathsf{F}^2$ et $(\vec{w}_1, \vec{w}_2) \in \mathsf{G}^2$. Par définition, $\vec{v}_1 + \vec{w}_1 \in \mathsf{F} + \mathsf{G}$ et $\vec{v}_2 + \vec{w}_2 \in \mathsf{F} + \mathsf{G}$. Nous avons, par associativité de l'addition dans E, $(\vec{v}_1 + \vec{w}_1) + (\vec{v}_2 + \vec{w}_2) = (\vec{v}_1 + \vec{w}_1) + (\vec{v}_2 + \vec{w}_2) \in \mathsf{F} + \mathsf{G}$. Pour tout $\lambda \in \mathsf{K}$, par distributivité de la multiplication externe dans E, $\lambda * (\vec{v}_1 + \vec{w}_1) = \lambda * \vec{v}_1 + \lambda * \vec{w}_1 \in \mathsf{F} + \mathsf{G}$. □

Définition 4 (intersection de sous-espaces vectoriels). *Soient* F *et* G *deux sous-espaces vectoriels de* $(\mathsf{E}, +, *)$. *L'intersection* $\mathsf{F} \cap \mathsf{G}$, *définie par*

$$\mathsf{F} \cap \mathsf{G} \overset{\text{déf.}}{=} \left\{ \vec{v} \in \mathsf{E} \text{ t.q. } \vec{v} \in \mathsf{F} \text{ et } \vec{v} \in \mathsf{G} \right\} \tag{1.4}$$

est un sous-espace vectoriel de E.

Soit $(\vec{v}, \vec{w}) \in (\mathsf{F} \cap \mathsf{G})^2$. Comme F est un sous-espace vectoriel de E, $\vec{v} + \vec{w} \in \mathsf{F}$ et comme G est un sous-espace vectoriel de E, $\vec{v} + \vec{w} \in \mathsf{G}$; donc $\vec{v} + \vec{w} \in \mathsf{F} \cap \mathsf{G}$. De même, pour tout $\lambda \in \mathsf{K}$, comme F est un sous-espace vectoriel de E, $\lambda * \vec{v} \in \mathsf{F}$ et comme G est un sous-espace vectoriel de E, $\lambda * \vec{w} \in \mathsf{G}$; donc $\lambda * (\vec{v} + \vec{w}) = \lambda * \vec{v} + \lambda * \vec{w} \in \mathsf{F} \cap \mathsf{G}$ (en utilisant la distributivité). $\square$

Définition 5 (somme directe de sous-espaces vectoriels). *Soient* F *et* G *deux sous-espaces vectoriels de* $(\mathsf{E}, +, *)$. *On dit que* E *est la somme directe de* F *et* G, *ce que l'on note* $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$, *si :*

1. $\mathsf{E} = \mathsf{F} + \mathsf{G}$ *;*
2. $\mathsf{F} \cap \mathsf{G} = \{\vec{0}_{\mathsf{E}}\}$.

On dit aussi que F *et* G *sont supplémentaires dans* E.

Théorème 1 *Soit* $(\mathsf{E}, +, *)$ *et* F *et* G *deux sous-espaces vectoriels de* E. E *est la somme directe de* F *et* G, *c.-à.-d.* $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$, *si et seulement si tout vecteur* $\vec{v} \in \mathsf{E}$ *se décompose de manière unique comme* $\vec{v} = \vec{x} + \vec{y}$ *avec* $\vec{x} \in \mathsf{F}$ *et* $\vec{y} \in \mathsf{G}$.

Supposons $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$ et soit $\vec{v} \in \mathsf{E}$. Étant donné que $\mathsf{E} = \mathsf{F} + \mathsf{G}$, il existe $\vec{x} \in \mathsf{F}$ et $\vec{y} \in \mathsf{G}$ tels que $\vec{v} = \vec{x} + \vec{y}$. Supposons qu'on puisse trouver un autre couple $\vec{x}' \in \mathsf{F}$ et $\vec{y}' \in \mathsf{G}$ tel que $\vec{v} = \vec{x}' + \vec{y}'$. Nous avons alors $(\vec{x} - \vec{x}') + (\vec{y} - \vec{y}') = \vec{v} - \vec{v} = \vec{0}_{\mathsf{E}}$ donc $(\vec{x} - \vec{x}') = -(\vec{y} - \vec{y}')$, avec $(\vec{x} - \vec{x}') \in \mathsf{F}$ et $-(\vec{y} - \vec{y}') = (-1) * (\vec{y} - \vec{y}') \in \mathsf{G}$. On en déduit que $(\vec{x} - \vec{x}') \in \mathsf{F} \cap \mathsf{G}$ et $(\vec{y} - \vec{y}') \in \mathsf{F} \cap \mathsf{G}$. Par hypothèse $\mathsf{F} \cap \mathsf{G} = \{\vec{0}_{\mathsf{E}}\}$ d'où $\vec{x} = \vec{x}'$ et $\vec{y} = \vec{y}'$.

Réciproquement, supposons que tout vecteur $\vec{v} \in \mathsf{E}$ se décompose de manière unique comme $\vec{v} = \vec{x} + \vec{y}$ avec $\vec{x} \in \mathsf{F}$ et $\vec{y} \in \mathsf{G}$. Cela implique évidemment que $\mathsf{E} = \mathsf{F} + \mathsf{G}$. Supposons qu'il existe $\vec{x} \in \mathsf{F} \cap \mathsf{G}$ tel que $\vec{x} \neq \vec{0}_{\mathsf{E}}$. On a alors $\vec{x} + (-\vec{x}) = \vec{0}_{\mathsf{E}} = \vec{0}_{\mathsf{E}} + \vec{0}_{\mathsf{E}}$ ce qui contredit l'hypothèse de décomposition unique. On a donc $\mathsf{F} \cap \mathsf{G} = \{\vec{0}_{\mathsf{E}}\}$ d'où $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$. $\square$

Notation : à partir de maintenant le signe de la multiplication externe « $*$ » sera omis par souci de clarté.

1.2 Bases

Définition 6 (dépendance linéaire). *Soit* $(\mathsf{E}, +, *)$ *un* K*-espace vectoriel et une famille de* n *vecteurs de* E, $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$. *Considérons l'équation*

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n = 0\,. \tag{1.5}$$

où $\lambda_1 \ldots, \lambda_n \in \mathsf{K}$ *sont les inconnues.*

S'il existe une solution non-triviale de cette équation, c.-à.-d. tels que les λ_i *ne sont pas tous nuls, alors les vecteurs* $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n$ *sont dits linéairement dépendants et la famille de vecteurs appelée famille liée.*

Dans le cas contraire (c.-à.-d. si la seule solution est $\lambda_1 = \cdots = \lambda_n = 0$*) ces vecteurs sont dits linéairement indépendants, et la famille de vecteurs appelée famille libre.*

En effet, si la famille est liée il existe au moins un entier $1 \leqslant k \leqslant n$ tel que $\lambda_k \neq 0$ tout en satisfaisant la relation (1.5). On peut alors exprimer le vecteur $\vec{v}_k$ en termes des autres vecteurs :

$$\vec{v}_k = -\frac{1}{\lambda_k}\big(\lambda_1\vec{v}_1 + \cdots + \lambda_{k-1}\vec{v}_{k-1} + \lambda_{k+1}\vec{v}_{k+1} + \cdots + \lambda_n\vec{v}_n\big)\,. \tag{1.6}$$

Définition 7 (sous-espace engendré). *Soit* E *un* K*-espace vectoriel une famille de* n *vecteurs de* E*,* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$*, non tous égaux au vecteur nul. L'ensemble des combinaisons linéaires de ces vecteurs, noté* $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$*,*

$$\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n) \overset{\text{déf.}}{=} \{\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n, \lambda_k \in \mathsf{K}\} \subseteq \mathsf{E}\,, \tag{1.7}$$

est le sous-espace vectoriel de E *engendré par cette famille de vecteurs.*

Montrons que $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$ est bien un sous-espace vectoriel de E. Premièrement, il est évident, au sens des ensembles, que $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n) \subset \mathsf{E}$ car tous les vecteurs $\vec{v}_k$ sont contenus dans E. Deuxièmement, si $\vec{V} \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$ alors il existe $(\lambda_1, \ldots, \lambda_n) \in \mathsf{K}^n$ tels que $\vec{V} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$. On a alors pour tout $\Lambda \in \mathsf{K}$, $\Lambda\vec{V} = (\Lambda\lambda_1)\vec{v}_1 + \cdots + (\Lambda\lambda_n)\vec{v}_n \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$. De même si $\vec{W} \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$ alors il existe $(\mu_1, \ldots, \mu_n) \in \mathsf{K}^n$ tels que $\vec{W} = \mu_1\vec{v}_1 + \cdots + \mu_n\vec{v}_n$, et $\vec{V} + \vec{W} = (\lambda_1 + \mu_1)\vec{v}_1 + \cdots + (\lambda_n + \mu_n)\vec{v}_n \in \mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n)$. $\square$

Définition 8 (famille génératrice). *Soit* E *un espace vectoriel et une famille de* n *vecteurs de* E*,* $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$*. Cette famille est dite génératrice de* E *si*

$$\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_n) = \mathsf{E} \tag{1.8}$$

On peut de manière équivalente écrire

$$\mathsf{E} = \mathrm{Vect}(\vec{v}_1) + \mathrm{Vect}(\vec{v}_2) + \cdots + \mathrm{Vect}(\vec{v}_n)\,, \tag{1.9}$$

en utilisant la notion de somme de sous-espaces vectoriels, voir la définition 3.

Il existe naturellement une infinité de familles génératrices pour un espace vectoriel donné ; en outre une famille génératrice peut être libre ou liée.

Définition 9 (base d'un espace vectoriel). *Soit* E *un espace vectoriel. Une base de* E *est une famille* $\{\vec{e}_1, \ldots, \vec{e}_n \in \mathsf{E}\}$ *libre et génératrice.*

La propriété cruciale d'une base d'un espace vectoriel est qu'il est possible de décomposer de manière *unique* tout vecteur $\vec{u} \in E$ sur cette base. Premièrement, étant donné qu'une base est une famille génératrice de E,

$$\forall \vec{u} \in E\,,\ \exists u^1 \ldots, u^n \in K \ \text{t.q.}\ \vec{u} = u^1\vec{e}_1 + \cdots + u^n\vec{e}_n\,. \tag{1.10}$$

Deuxièmement, supposons l'existence de deux décompositions différentes d'unvecteur $\vec{u} \in E$ sur une certaine base de E. Nous avons alors

$$\left\{\begin{array}{lcl} \vec{u} & = & u^1\vec{e}_1 + \cdots + u^n\vec{e}_n \\ \vec{u} & = & v^1\vec{e}_1 + \cdots + v^n\vec{e}_n \end{array}\right. \implies 0 = (u^1 - v^1)\vec{e}_1 + \cdots + (u^n - v^n)\vec{e}_n\,. \tag{1.11}$$

La famille $\{\vec{e}_1, \ldots, \vec{e}_n\}$ étant libre par définition, nous avons alors, pour tout k, $u^k = v^k$, prouvant l'unicité de la décomposition d'un vecteur *sur une base donnée.* $\square$.

Le n-uplet de nombres $(v^1, v^2, \ldots, v^n) \in K^n$ constitue les *coordonnées* du vecteur $\vec{v}$ dans la base $\{\vec{e}_1, \ldots, \vec{e}_n \in E\}$.

En utilisant la notion de somme directe de sous-espaces vectoriels, voir la définition 5, si $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base de E, on peut écrire

$$E = \mathrm{Vect}(\vec{e}_1) \oplus \cdots \oplus \mathrm{Vect}(\vec{e}_n) = \bigoplus_{k=1}^{n} \mathrm{Vect}(\vec{e}_k)\,. \tag{1.12}$$

Définition 10 (dimension d'un espace vectoriel). *Soit E un K-espace vectoriel muni d'une base $\{\vec{e}_1, \ldots, \vec{e}_n\}$. L'entier n positif est appelé dimension de l'espace vectoriel E et on note* $\dim E = n$.

Propriété 2 *La dimension d'un espace vectoriel E est indépendante de la base choisie.*

Soient $\{\vec{e}_1, \ldots, \vec{e}_n\}$ et $\{\vec{f}_1, \ldots, \vec{f}_r\}$ deux bases de E. Supposons $r > n$. On peut décomposer $\vec{f}_1$ sur la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$, suivant :

$$\vec{f}_1 = {f_1}^1\vec{e}_1 + {f_1}^2\vec{e}_2 + \cdots + {f_1}^n\vec{e}_n\,, \tag{1.13}$$

où les coefficients ${f_1}^\ell$ ne sont pas tous nuls (autrement $\vec{f}_1 = \vec{0}_E$ et ce vecteur ne peut faire partie de la base). Supposons par exemple, sans perte de généralité, que ${f_1}^1 \neq 0$. On peut donc écrire

$$\vec{e}_1 = \frac{1}{{f_1}^1}\left(\vec{f}_1 - {f_1}^2\vec{e}_2 - {f_1}^3\vec{e}_3 \cdots - {f_1}^n\vec{e}_n\right)\,, \tag{1.14}$$

et une base de E est donnée par la famille de vecteurs $\{\vec{f}_1, \vec{e}_2, \vec{e}_3, \ldots, \vec{e}_n\}$. En conséquence, on peut décomposer $\vec{f}_2$ dans cette base :

$$\vec{f}_2 = {f_2}^1\vec{f}_1 + {f_2}^2\vec{e}_2 + {f_2}^3\vec{e}_3 + \cdots + {f_2}^n\vec{e}_n\,, \tag{1.15}$$

où l'un au moins des n coefficients ${f_2}^\ell$ est non-nul. Si uniquement ${f_2}^1$ est non-nul, alors $\vec{f}_1$ et $\vec{f}_2$ sont colinéaires, contredisant l'hypothèse ; donc l'un au moins des autres coefficients, par exemple ${f_2}^2$, est non-nul. On peut donc écrire :

$$\vec{e}_2 = \frac{1}{{f_2}^2}\left(\vec{f}_1 - {f_1}^2\vec{f}_2 - {f_1}^3\vec{e}_3 \cdots - {f_1}^n\vec{e}_n\right) . \tag{1.16}$$

et une base de E est donnée par la famille de vecteurs $\{\vec{f}_1, \vec{f}_2, \vec{e}_3, \ldots, \vec{e}_n\}$. En itérant ce processus n fois, on obtient que $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_n\}$ est une base de E, et donc que les vecteurs $\vec{f}_k$ avec $n < k \leqslant r$ s'expriment comme combinaison linéaire de $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_n\}$, contredisant l'hypothèse de départ. □

Propriété 3 *Un K-espace vectoriel de dimension 1 est isomorphe à K, vu comme un espace vectoriel.*

Soit E un K-espace vectoriel. Si $\dim E = 1$, alors toute base se compose d'un seul vecteur. Soit $\{\vec{e}\}$ une telle base ; on peut alors associer à tout vecteur $\vec{v}$ un unique élément de K suivant :

$$\vec{v} = v\vec{e} \quad \longleftrightarrow \quad v \in K . \tag{1.17}$$

On a donc une bijection entre E et K. Étant donné que cette bijection est linéaire, c.-à.-d. que l'image du vecteur $\lambda\vec{v} + \mu\vec{w}$ est $\lambda v + \mu w \in K$, on dit que cette bijection est un *isomorphisme* d'espaces vectoriels.

Définition 11 (rang d'une famille de vecteurs). *Soit E un K-espace vectoriel de dimension n et une famille de p vecteurs de E $\{\vec{v}_1, \ldots, \vec{v}_p\}$ non tous égaux au vecteur nul. Le rang de cette famille de vecteurs est la dimension du sous-espace vectoriels qu'ils engendrent :*

$$\mathrm{Rang}\,(\vec{v}_1, \ldots, \vec{v}_p) \overset{\text{déf.}}{=} \dim\left(\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_p)\right) \tag{1.18}$$

Si $p \leqslant n$, le rang r de la famille $\{\vec{v}_1, \ldots, \vec{v}_p\}$ satisfait évidemment

$$1 \leqslant r \leqslant p \leqslant n , \tag{1.19}$$

et $r = p$ si ces vecteurs forment une famille libre, car alors $\{\vec{v}_1, \ldots, \vec{v}_p\}$ est une base de $\mathrm{Vect}(\vec{v}_1, \ldots, \vec{v}_p)$.

Si le nombre p de vecteurs de la famille est strictement supérieur à n, dimension de E, cette famille est nécessairement liée et

$$1 \leqslant r \leqslant n < p . \tag{1.20}$$

Exemple 3 Soit $K_n[X]$ l'espace vectoriel des polynômes de degré au plus égal à n. Une base de cet espace vectoriel est donnée par les monômes $\{1, X, X^2, \ldots, X^n\}$ et la dimension de l'espace est égale à n. Une base de $K[X]$ est donnée par $\{1, X, \ldots, X^k, \ldots\}$ qui est une famille infinie mais dénombrable. Ce dernier espace vectoriel est donc de dimension infinie.

*
* *

Propriété 4 *Si* F *est un sous-espace vectoriel de* E *alors* $\dim \mathsf{F} \leqslant \dim \mathsf{E}$.

Soit $\{\vec{f}_1, \ldots, \vec{f}_r\}$ une base de F. Ces vecteurs sont par définition des vecteurs de E et, formant une base de F, ils sont nécessairement linéairement indépendants. Ils constituent donc une famille libre de vecteurs de E, mais non nécessairement génératrice.

Supposons qu'il existe un vecteur non-nul quelconque de E qui n'appartient pas à F que nous appelerons $\vec{h}_{r+1}$. Il est alors évident que $\{\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1}\}$ forme une famille libre, et que

$$\mathsf{F} \subset \mathrm{Vec}\left(\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1}\right) \subseteq \mathsf{E}\,. \tag{1.21}$$

Supposons à présent qu'il existe un vecteur non-nul quelconque de E qui n'appartient pas à $\mathrm{Vec}\,(\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1})$ que nous appelerons $\vec{h}_{r+2}$. Il est alors évident que $\{\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1}, , \vec{h}_{r+2}\}$ forme une famille libre, et que

$$\mathrm{Vec}\left(\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1}\right) \subset \mathrm{Vec}\left(\vec{f}_1, \ldots, \vec{f}_r, \vec{h}_{r+1}, \vec{h}_{r+2}\right) \subseteq \mathsf{E}\,. \tag{1.22}$$

Itérativement, si $\mathsf{F} \subsetneq \mathsf{E}$, nous construisons ainsi une base de E en complétant la base de F par $(n - r)$ vecteurs de $\mathsf{E} \backslash \mathsf{F}$, où n est la dimension de E. □

Propriété 5 *Soit* $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$ *un espace vectoriel. Si* $\{\vec{f}_1, \ldots, \vec{f}_k\}$ *est une base de* F *et* $\{\vec{g}_1, \ldots, \vec{g}_\ell\}$ *une base de* G *alors* $(\vec{f}_1, \ldots, \vec{f}_k, \vec{g}_1, \ldots, \vec{g}_\ell)$ *est une base de* E*, et en conséquence*

$$\dim \mathsf{E} = \dim \mathsf{F} + \dim \mathsf{G}\,. \tag{1.23}$$

Cette propriété découle directement de la définition 5 et et de la propriété 4. En effet, étant donné que $\mathsf{F} \cap \mathsf{G} = \{\vec{0}_{\mathsf{E}}\}$, les vecteurs $\{\vec{g}_1, \ldots, \vec{g}_\ell\}$ n'appartiennent pas à F et sont linéairement indépendants. Il complètent donc $\{\vec{f}_1, \ldots, \vec{f}_k\}$ pour former une base de $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$. □

1.3 Applications linéaires

Définition 12 (application linéaire). *Soient* E *et* F *deux* K*-espaces vectoriels et* $a : \mathsf{E} \to \mathsf{F}$ *une application. On dit que* a *est une application linéaire si :*

— $\forall \vec{u}, \vec{v} \in \mathsf{E},\ a(\vec{u} + \vec{v}) = a(\vec{u}) + a(\vec{v})$ *;*

— $\forall \vec{u} \in \mathsf{E},\ \forall \lambda \in \mathsf{K},\ a(\lambda \vec{u}) = \lambda a(\vec{u})$.

ou, d'une manière plus condensée, $a(\lambda \vec{u} + \mu \vec{v}) = \lambda a(\vec{u}) + \mu a(\vec{v})$.

Exemple 4 (applications linéaires sur $\mathsf{K}[X]$**).** Soit $P = a_0 + a_1 X + a_2 X^2 + \cdots \in \mathsf{K}[X]$. La dérivation et l'intégration sont des applications linéaires de $\mathsf{K}[X]$ dans lui-même :

$$\mathcal{D}\ : \left\{ \begin{array}{lcl} \mathsf{K}[X] & \to & \mathsf{K}[X] \\ a_0 + a_1 X + a_2 X^2 \cdots & \mapsto & a_1 + 2a_2 X \cdots \end{array} \right. \tag{1.24a}$$

$$\mathcal{I}\ : \left\{ \begin{array}{lcl} \mathsf{K}[X] & \to & \mathsf{K}[X] \\ a_0 + a_1 X + a_2 X^2 \cdots & \mapsto & a_0 X + \frac{a_1}{2} X^2 + \frac{a_2}{3} X^3 \cdots \end{array} \right. \tag{1.24b}$$

La preuve, évidente, est laissée en exercice.

Exemple 5 (projecteur) Soit un K-espace vectoriel E qui se décompose comme une somme directe de sous-espaces vectoriels, $\mathsf{E} = \mathsf{F} \oplus \mathsf{G}$. D'après le théorème 1, tout vecteur $\vec{v} = \mathsf{E}$ se décompose come $\vec{v} = \vec{x} + \vec{y}$ avec $\vec{x} \in \mathsf{F}$ et $\vec{y} \in \mathsf{G}$. Le projecteur sur F, noté $\mathcal{P}_\mathsf{F}$ (ou plus explicitement le projecteur sur F parallèlement à G), est donné par l'application linéaire :

$$\mathcal{P}_\mathsf{F} : \begin{cases} \mathsf{E} & \to \quad \mathsf{F} \\ \vec{v} = \vec{x} + \vec{y} & \mapsto \quad \mathcal{P}_\mathsf{F}(\vec{v}) = \vec{y} \end{cases} \tag{1.25}$$

Soient $\vec{v}_1$ et $\vec{v}_2$ deux vecteurs de E, qui se décomposent chacun comme $\vec{v}_i = \vec{x}_i + \vec{y}_i$, $i = 1, 2$, où $\vec{x}_i \in \mathsf{F}$ et $\vec{y}_i \in \mathsf{G}$. Pour tout $(\lambda_1, \lambda_2) \in \mathsf{K}^2$ nous avons la décomposition unique :

$$\vec{V} \overset{\text{déf.}}{=} \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 = \underbrace{\left(\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2\right)}_{\in\mathsf{F}} + \underbrace{\left(\lambda_1\vec{y}_1 + \lambda_2\vec{y}_2\right)}_{\in\mathsf{G}} \tag{1.26}$$

d'où

$$\mathcal{P}_\mathsf{F}(\vec{V}) = \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 = \lambda_1\mathcal{P}_\mathsf{F}(\vec{v}_1) + \lambda_2\mathcal{P}_\mathsf{F}(\vec{v}_2) \tag{1.27}$$

par identification. □

*
* *

Propriété 6 (image, noyau). *Soient* E *et* F *deux* K*-espaces vectoriels et* $a : \mathsf{E} \to \mathsf{F}$ *une application linéaire. Les deux sous-ensembles suivants :*

— *le noyau de* a*,* $\operatorname{Ker} a = \{\vec{v} \in \mathsf{E} \,|\, a(v) = \vec{0}_\mathsf{F}\} \subset \mathsf{E}$

— *l'image de* a*,* $\operatorname{Im} a = \{\vec{w} \in \mathsf{F} \,|\, \exists \vec{v} \in \mathsf{E} \,|\, a(\vec{v}) = \vec{w}\} \subset \mathsf{F}$.

sont respectivement un sous-espace vectoriel de E *et un sous-espace vectoriel de* F.

Montrons premièrement que $\operatorname{Ker} a$ est un sous-espace vectoriel de E. Si $\vec{v}$ et $\vec{v}'$ appartiennent à $\operatorname{Ker} a$ alors, par linéarité de a, pour tout $(\lambda, \mu) \in \mathsf{K}^2$,

$$a(\lambda\vec{v} + \mu\vec{v}') = \lambda a(\vec{v}) + \mu a(\vec{v}') = \vec{0}_\mathsf{F} \implies \lambda\vec{v} + \mu\vec{v}' \in \operatorname{Ker} a \tag{1.28}$$

Deuxièmement, si $\vec{w}, \vec{w}' \in \operatorname{Im} a$, alors il existe $\vec{v}, \vec{v}' \in \mathsf{E}$ tels que $\vec{w} = a(\vec{v})$ et $\vec{w}' = a(\vec{v}')$. On a alors

$$\lambda\vec{w} + \mu\vec{w}' = \lambda a(\vec{v}) + \mu a(\vec{v}') = a(\lambda\vec{v} + \mu\vec{v}') \in \operatorname{Im} a \tag{1.29}$$

par linéarité de l'application a.

Remarquons pour finir que le noyau et l'image d'une application linéaire ne sont jamais vides. En effet, comme pour tout vecteur $\vec{v} \in \mathsf{E}$, $0 * \vec{v} = \vec{0}_\mathsf{E}$, nous avons $a(\vec{0}_\mathsf{E}) = a(0 * \vec{v}) = 0 * a(\vec{v}) = \vec{0}_\mathsf{F}$ donc $a(\vec{0}_\mathsf{E}) = \vec{0}_\mathsf{F}$, soit $\vec{0}_\mathsf{E} \in \operatorname{Ker} a$ et $\vec{0}_\mathsf{F} \in \operatorname{Im} a$. □

Théorème 2 (caractérisation d'un projecteur). *Soit* E *un* K*-espace vectoriel et* $\mathcal{P}$ *une application linéaire de* E *dans* E *telle que, pour tout* $\vec{v} \in \mathsf{E}$*,*

$$\mathcal{P}^2(\vec{v}) \overset{\text{déf.}}{=} \left(\mathcal{P} \circ \mathcal{P}\right)(\vec{v}) = \mathcal{P}\left[\mathcal{P}(\vec{v})\right] = \mathcal{P}(\vec{v})\,. \tag{1.30}$$

alors

$$\mathsf{E} = \operatorname{Ker}\mathcal{P} \oplus \operatorname{Im}\mathcal{P}\,, \tag{1.31}$$

c.-à.-d. que le noyau et l'image de $\mathcal{P}$ *sont supplémentaires dans* E*, et* $\mathcal{P}$ *est le projecteur sur le sous-espace vectoriel* $\operatorname{Im}\mathcal{P} \subset \mathsf{E}$ *parallèlement à* $\operatorname{Ker}\mathcal{P}$*.*

Pour tout $\vec{v} \in \mathsf{E}$, écrivons

$$\vec{v} = \underbrace{\mathcal{P}(\vec{v})}_{\in \operatorname{Im}\mathcal{P}} + \big(\vec{v} - \mathcal{P}(\vec{v})\big)\,. \tag{1.32}$$

En utilisant la linéarité de $\mathcal{P}$ et la propriété (1.30) on obtient

$$\mathcal{P}\big(\vec{v} - \mathcal{P}(\vec{v})\big) = \mathcal{P}\big(\vec{v}\big) - \mathcal{P}^2(\vec{v}) = \mathcal{P}\big(\vec{v}\big) - \mathcal{P}\big(\vec{v}\big) = \vec{0}_{\mathsf{E}} \quad\Longrightarrow\quad \vec{v} - \mathcal{P}(\vec{v}) \in \operatorname{Ker}\mathcal{P}\,. \tag{1.33}$$

Cela montre que tout vecteur $\vec{v}$ se décompose comme somme d'un vecteur de $\operatorname{Im}\mathcal{P}$ et d'un vecteur de $\operatorname{Ker}\mathcal{P}$, c.-à.-d. que $\mathsf{E} = \operatorname{Ker}\mathcal{P} + \operatorname{Im}\mathcal{P}$.

Pour montrer que ces deux sous-espaces sont supplémentaires, il faut encore montrer que leur intersection ne contient que le vecteur nul. Soit $\vec{v} \in \operatorname{Im}\mathcal{P} \cap \operatorname{Ker}\mathcal{P}$. Le vecteur $\vec{v}$ appartenant à l'image de $\mathcal{P}$ existe alors $\vec{w} \in \mathsf{E}$ tel que $\mathcal{P}(\vec{w}) = \vec{v}$. Appliquons le projecteur à cette égalité. Nous trouvons d'une part

$$\mathcal{P}^2(\vec{w}) = \mathcal{P}(\vec{v}) = \vec{0}_{\mathsf{E}} \tag{1.34}$$

car $\vec{v}$ est dans le noyau de $\mathcal{P}$ et d'autre part, $\mathcal{P}$ étant un projecteur,

$$\mathcal{P}^2(\vec{w}) = \mathcal{P}(\vec{w}) = \vec{v}\,, \tag{1.35}$$

d'où $\vec{v} = \vec{0}_{\mathsf{E}}$, complétant la preuve. □

Définition 13 (endomorphisme). *Soit* E *un* K*-espace vectoriel. une application linéaire* a *de* E *dans* E *est appelée endomorphisme.*

Pour un endomorphisme, à la fois l'image et le noyau sont des sous-espaces vectoriels de E. Un projecteur est un exemple d'endomorphisme ; en raison de la propriété (1.30) cet endomorphisme est dit *idempotent*.

Définition 14 (rang d'une application linéaire). *Soient* E *et* F *des* K*-espaces vectoriels et* $\mathsf{a} : \mathsf{E} \to \mathsf{F}$ *une application linéaire. Le rang de* a *est défini comme la dimension de* $\operatorname{Im}\mathsf{a}$*, qui est un sous-espace vectoriel de* F *:*

$$\operatorname{Rang}\mathsf{f} \overset{\text{déf.}}{=} \dim\Big(\operatorname{Im}\mathsf{f}\Big) \tag{1.36}$$

Théorème 3 (théorème du rang). *Soient* E *et* F *des* K*-espaces vectoriels et* $\mathsf{a} : \mathsf{E} \to \mathsf{F}$ *une application linéaire. Nous avons :*

$$\operatorname{Rang}\mathsf{a} + \dim\big(\operatorname{Ker}\mathsf{a}\big) = \dim\big(\mathsf{E}\big)\,. \tag{1.37}$$

Notons que le membre de droite est la dimension de l'ensemble de départ de l'application a*.*

Soit $\{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de $\vec{E}$, où $n = \dim(E)$. Considérons premièrement le cas où $\operatorname{Ker} a = \vec{0}_E$ et montrons que la famille de vecteurs de F constituée de $\{a(\vec{e}_1), \ldots, a(\vec{e}_n)\}$ est une famille libre. Soit $(\lambda_1, \ldots, \lambda_n) \in K^n$. Nous avons par linéarité de a

$$\lambda_1 a(\vec{e}_1) + \cdots + \lambda_n a(\vec{e}_n) = \vec{0}_F \;\Leftrightarrow\; a(\lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n) = \vec{0}_F\,. \tag{1.38}$$

Comme, par hypothèse, $\operatorname{Ker} a = \vec{0}_E$, le membre de droite implique que $\lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n = \vec{0}_E$ dont la seule solution est $\lambda_1 = \cdots = \lambda_n = 0$.

Si on suppose maintenant que $\dim(\operatorname{Ker} a) = r > 0$, on choisit une base $\{\vec{n}_1, \ldots, \vec{n}_r\}$ pour le noyau de f, qui est un sous-espace vectoriel de E. Suivant la propriété 4, on peut alors compléter la base de $\operatorname{Ker} a$ par des vecteurs $\{\vec{c}_{r+1}, \ldots, \vec{c}_n\}$ pour former une base de E. Par construction, ces vecteurs n'appartiennent pas à $\operatorname{Ker} a$ et, par un raisonnement identique à celui développé autour de l'éqn. (1.38) on montre que les vecteurs $\{a(\vec{c}_{r+1}), \ldots, a(\vec{c}_n)\}$, qui appartiennent à l'image de a, sont linéairement indépendants.

Il reste à montrer que ces vecteurs forment une base de $\operatorname{Im} a$ (cela s'applique aussi au premier cas considéré). Soit donc $\vec{w} \in \operatorname{Im} a$. Par définition,

$$\exists \vec{v} \in \vec{E} \quad \text{t.q.} \quad a(\vec{v}) = \vec{w} \tag{1.39}$$

Décomposons le vecteur $\vec{v}$ dans la base $\{\vec{n}_1, \ldots, \vec{n}_r, \vec{c}_{r+1}, \ldots, \vec{c}_n\}$ comme

$$\vec{v} = v^1 \vec{n}_1 + \cdots + v^r \vec{n}_r + v^{r+1} \vec{c}_{r+1} + \cdots + v^n \vec{c}_n\,. \tag{1.40}$$

Nous avons alors, par linéarité

$$\begin{aligned}\vec{w} = a(\vec{v}) &= v^1 \cancel{a(\vec{n}_1)} + \cdots + v^r \cancel{a(\vec{n}_r)} + v^{r+1} a(\vec{c}_{r+1}) + \cdots + a(v^n \vec{c}_n) \\ &= v^{r+1} a(\vec{c}_{r+1}) + \cdots + v^n a(\vec{c}_n)\end{aligned} \tag{1.41}$$

démontrant que $\{a(\vec{c}_{r+1}), \ldots, a(\vec{c}_n)\}$ est bien une base de l'image de a, et en conséquence que $\dim\big(\operatorname{Im} a\big) = n - r$. $\square$

1.4 Matrice d'une application linéaire

Une matrice est, strictement parlant, un tableau bidimensionnel de nombres ; dans le cadre de ce cours, nous verrons uniquement les matrices comme représentant une application linéraire une fois un choix de bases effectué.

Définition 15 (matrice d'une application linéaire).

Soient E et F des K-espaces vectoriels de dimensions n et p, munis respectivement des bases $\mathcal{B}_E = \{\vec{e}_i, i = 1, \ldots, n\}$ et $\mathcal{B}_F = \{\vec{f}_j, j = 1, \ldots, p\}$. Soit une application linéaire $a :\ E \to F$. On peut décomposer l'image de chaque vecteur de la base $\mathcal{B}_E$ choisie pour E sur la base $\mathcal{B}_F$ choisie pour F :

$$a(\vec{e}_i) = \sum_{j=1}^{p} \vec{f}_j \, a^j{}_i \tag{1.42}$$

Le tableau de nombres de taille $p \times n$, de composantes $a^j{}_i \in K$, est la matrice associée à l'application linéaire a de $(E, \mathcal{B}_E)$ dans $(F, \mathcal{B}_F)$.

Propriété 7 *Soit* a *une application linéaire de* $(E, \mathcal{B}_E)$ *dans* $(F, \mathcal{B}_F)$*. La connaissance de la matrice associée à* a *permet de connaître l'image de tout vecteur de* E*.*

Cette propriété découle directement de la linéarité de a. Dans la base $\mathcal{B}_E$ un vecteur $\vec{v} \in E$ se décompose comme :

$$\vec{v} = \sum_{i=1}^{n} v^i \vec{e}_i \,. \tag{1.43}$$

Nous avons alors la décomposition :

$$\begin{aligned} a(\vec{v}) = a\left(\sum_{i=1}^{n} v^i \vec{e}_i\right) &= \sum_{i=1}^{n} v^i a(\vec{e}_i) = \sum_{i=1}^{n} v^i \sum_{j=1}^{p} \vec{f}_j \, a^j{}_i \\ &= \sum_{j=1}^{p} \left(\sum_{i=1}^{n} a^j{}_i v^i\right) \vec{f}_j \,, \end{aligned} \tag{1.44}$$

Soit, en présentant le résultat légèrement différemment :

$$a(\vec{v}) = \vec{w} \,, \quad \vec{v} = \sum_{i=1}^{n} v^i \vec{e}_i \,, \quad \vec{w} = \sum_{j=1}^{p} w^j \vec{f}_j \,, \quad w^j = \sum_{i=1}^{n} a^j{}_i v^i \tag{1.45}$$

Conventions et notations

1. Par convention, on représente les composantes $\{a^j{}_i, j = 1, \ldots, p, i = 1, \ldots n\}$ d'une matrice sous la forme d'un tableau de p lignes et n colonnes, où j représente le numéro de ligne et i le numéro de colonne.
2. on note en général l'objet matrice avec une lettre majuscule. Les composantes de la matrice A de peuvent se noter de deux manières équivalentes :
$$(A)^i{}_j \stackrel{\text{déf.}}{=} a^i{}_j \tag{1.46}$$
3. La convention adoptée dans ces notes (numéro de ligne en exposant, numéro de colonne en indice) n'est pas universelle, mais a été choisie en cohérence avec le cours du premier semestre. On peut tout à fait écrire $(A)_{ij} \stackrel{\text{déf.}}{=} a_{ij}$.
4. Soit a une application linéaire de de $(E, \mathcal{B}_E)$ dans $(F, \mathcal{B}_F)$ et A la matrice associée. La relation $a(\vec{e}_i) = \sum_{j=1}^{p} \vec{f}_j \, a^j{}_i$ indique que la i-ème colonne de la matrice donne la décomposition de l'image de $\vec{e}_i$, vecteur de base de E, dans la base $\mathcal{B}_F = \{\vec{f}_j, j = 1, \ldots, p\}$ de l'espace d'arrivée F :
$$A = \begin{array}{c} \begin{array}{ccccc} a(\vec{e}_1) & & a(\vec{e}_i) & & a(\vec{e}_n) \end{array} \\ \begin{pmatrix} a^1{}_1 & \cdots a^1{}_i & \cdots & a^1{}_n \\ a^2{}_1 & \cdots a^2{}_i & \cdots & a^2{}_n \\ \vdots & \vdots & & \vdots \\ a^p{}_1 & \cdots a^p{}_i & \cdots & a^p{}_n \end{pmatrix} \end{array} \tag{1.47}$$

Exemple 6 (dérivation de polynômes). Soit $K_n[X]$ l'espace vectoriel des polynômes de degré au plus n et $\mathcal{D} : K_n[X] \to K_n[X]$ l'application dérivée (voir l'exemple 4). L'action de cette application linéaire sur la base des monômes $\{1, X, \ldots, X^n\}$ est :

$$\mathcal{D}(1) = 0 \ , \quad \mathcal{D}(X^k) = kX^{k-1} \tag{1.48}$$

La matrice D associée est la matrice $n \times n$ de la forme :

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{1.49}$$

Nous remarquons que la première colonne ne contient que des zéros car l'image de 1 (dérivée d'une constante) est le polynôme nul. La dernière ligne ne contient également que des zéros, tout simplement car la dérivée de tout polynôme de degré au plus n est de degré au plus $(n-1)$.

*
* *

Définition 16 (matrice nulle). *La matrice* $p \times n$ *qui ne contient que des zéros est appelée matrice nulle, et notée* $\mathbb{O}_{p,n}$.

Propriété 8 (Espace vectoriel de matrices). *L'ensemble des matrices* $p \times n$ *à coefficients dans* K*, c.-à.-d. à* p *lignes et* n *colonnes, forme un espace vectoriel noté* $\mathcal{M}_K(p, n)$*, de dimension* np.

Soient A et B deux matrices de dimension $p \times n$. Si on associe ces deux matrices à des applications linéaires a et b de $(E, \mathcal{B}_E)$ dans $(F, \mathcal{B}_F)$, étant donné que, pour tout $\vec{v} \in E$ et pour tout $(\lambda, \lambda') \in K^2$, $(\lambda a + \lambda' b)(\vec{v}) = \lambda a(\vec{v}) + \lambda' b(\vec{v})$ la structure d'espace vectoriel de l'ensemble $\mathcal{M}_K(p, n)$ en découle directement.

Pour obtenir la dimension de cet espace vectoriel, il suffit d'en construire une base. Soit une famille de matrices $d_{k\ell}$, avec $k = 1, \ldots, p$ et $\ell = 1, \ldots, n$ telles que

$$\left(d_{k\ell}\right)^i{}_j = \begin{cases} 1 & \text{si} \quad i = k \text{ et} j = \ell \\ 0 & \text{sinon} \end{cases} \tag{1.50}$$

On peut alors décomposer toute matrice A, de composantes $a^i{}_j$, comme

$$A = \sum_{k=1}^{p} \sum_{\ell=1}^{n} d_{k\ell} a^k{}_\ell \, . \tag{1.51}$$

La famille $\{d_{k\ell}, k = 1, \dots, p; n = 1, \dots n\}$, qui contient np éléments, est donc une famille génératrice de $\mathcal{M}_K(p,n)$. Reste à montrer que cette famille est libre. Il est évident que

$$\sum_{k=1}^{p}\sum_{\ell=1}^{n} d_{k\ell}a^k{}_\ell = \mathbb{O}_{n,p} \quad\Longrightarrow\quad \forall k,\ell\,,\ a^k{}_\ell = 0\,, \tag{1.52}$$

car il faut annuler chaque composante de la matrice définie par le membre de gauche de l'égalité. □

Propriété 9 (composition d'applications linéaires). *Soient* E, F *et* G *des* K*-espaces vectoriels et soient* $a : E \to F$ *et* $b : F \to G$ *des applications linéaires. Alors l'application composée*

$$b\circ a \left\{ \begin{array}{ccc} E & \to & G \\ \vec{v} & \mapsto & (b\circ a)(\vec{v}) = b\big(a(\vec{v})\big) \end{array}\right. \tag{1.53}$$

est une application linéaire. Si de plus les espaces vectoriels E, F *et* G *sont munis respectivement des bases* $\mathcal{B}_E = \{\vec{e}_i, i = 1, \dots, n\}$, $\mathcal{B}_F = \{\vec{f}_i, i = 1, \dots, p\}$ *et* $\mathcal{B}_G = \{\vec{g}_k, k = 1, \dots, r\}$ *alors les composantes de la matrice associée à l'application* $(b\circ a)$ *sont données par*

$$(b\circ a)^k{}_i = \sum_{j=1}^{p} b^k{}_j\, a^j{}_i\,. \tag{1.54}$$

De la composition d'applications linéaires nous pouvons induire, de manière générale la notion de produit de matrices.

Définition 17 (produit matriciel). *Soient deux matrices* $A \in \mathcal{M}_K(p,n)$ *(donc* p *lignes et* n *colonnes) et* $B \in \mathcal{M}_K(q,p)$ *(donc* q *lignes et* p *colonnes). Le produit de la matrice* B *par la matrice* A*, dans cet ordre, donne une matrice* $B\cdot A \in \mathcal{M}_K(q,n)$ *(donc* q *lignes et* n *colonnes) de composantes :*

$$\big(B\cdot A\big)^i{}_j = \sum_{k=1}^{p} b^i{}_k\, a^k{}_j \tag{1.55}$$

Pour calculer le produit matriciel $B\cdot A$, on somme donc les composantes de la k-ième colonne B avec celles de la k-ième ligne de A.

Propriété 10 (associativité et distributivité du produit matriciel). *Le produit matricel est associatif, c.-à.-d. que pour toutes matrices* $A \in \mathcal{M}_K(p,n)$, $B \in \mathcal{M}_K(q,p)$ *et* $C \in \mathcal{M}_K(r,q)$,

$$(C\cdot B)\cdot A = C\cdot(B\cdot A) \tag{1.56}$$

Le produit matriciel est également distributif par rapport à l'addition, c.-à.-d. que, pour toutes matrices $A_1 \in \mathcal{M}_K(p,n)$, $A_2 \in \mathcal{M}_K(p,n)$ *et* $B \in \mathcal{M}_K(q,p)$,

$$B\cdot(A_1 + A_2) = B\cdot A_1 + B\cdot A_2\,. \tag{1.57}$$

La preuve de ces propriétés, qui découlent directement de la définition 17, est laissée en exercice.

Définition 18 (matrice transposée). *Soit une matrice* $A \in \mathcal{M}_K(p, n)$*. La matrice transposée de* A*, notée* A^T*, est une matrice à* n *lignes et* p *colonnes, donc un élément de* $\mathcal{M}_K(n, p)$*, dont les composantes sont données par :*

$$\left(A^T\right)^i{}_j = \left(A\right)^j{}_i\,, \tag{1.58}$$

c.-à.-d. en intervertissant le rôle des lignes et des colonnes par rapport à A*.*

Exemple 7 *Soit la matrice*

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\,. \tag{1.59}$$

Sa transposée est la matrice

$$M^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}\,. \tag{1.60}$$

*
* *

Propriété 11 (transposée d'un produit matriciel). *Soient deux matrices* $A \in \mathcal{M}_K(p, n)$ *et* $B \in \mathcal{M}_K(q, p)$*.*

$$\left(B \cdot A\right)^T = A^T \cdot B^T \tag{1.61}$$

La preuve de cette propriété est évidente en composantes. Nous avons

$$\begin{aligned} \left((B \cdot A)^T\right)^i{}_j = \left(B \cdot A\right)^j{}_i = \sum_{k=1}^{p} b^j{}_k a^k{}_i = \sum_{k=1}^{p} (A)^k{}_i (B)^j{}_k \\ = \sum_{k=1}^{p} (A^T)^i{}_k (B^T)^k{}_j = \left(A^T \cdot B^T\right)^i{}_j \end{aligned} \tag{1.62}$$

□

Définition 19 (matrice colonne). *Une matrice colonne est une matrice ne comprenant qu'une seule colonne, soit un élément de* $\mathcal{M}_{n,1}$ *pour un certain* $n > 1$*. Si* E *est un espace vectoriel muni d'une base* $\mathcal{B}_E = \{\vec{e}_i, i = 1, \ldots, n\}$*, on peut associer à un vecteur* $\vec{v} \in E$ *une matrice colonne* V *contenant les composantes de ce vecteur dans la base choisie :*

$$\vec{v} = \sum_{i=1}^{n} v^i \vec{e}_i \quad \longleftrightarrow \quad V = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{1.63}$$

Définition 20 (matrice ligne). *Une matrice ligne est une matrice ne comprenant qu'une seule ligne, soit un élément de $\mathcal{M}_{1,n}$ pour un certain $n > 1$*

Une matrice ligne est donc associée à une application linéaire d'un espace E de dimension n vers un espace vectoriel de dimension 1, qui est équivalent à une application de E vers K, donc une forme linéaire ω (cf. le cours du premier semestre). On peut ainsi considérer qu'une matrice ligne est associée à un vecteur de l'espace dual à E.

Définition 21 (matrice d'une famille de vecteurs). *Soit E un espace vectoriel de dimension p muni d'une base $\mathcal{B}_E = \{\vec{e}_i, i = 1, \dots, p\}$ et $\{\vec{v}_1, \dots, \vec{v}_n\}$ une famille de n vecteurs de E. La matrice associée à cette famille de vecteurs est la matrice $A \in \mathcal{M}_K(p, n)$ de composantes :*

$$\left(A\right)^i{}_j = \begin{pmatrix} (v_1)^1 & (v_2)^1 & \cdots & (v_k)^1 & \cdots & (v_n)^1 \\ (v_1)^2 & (v_2)^2 & \cdots & (v_k)^2 & \cdots & (v_n)^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ (v_1)^p & (v_2)^p & \cdots & (v_k)^p & \cdots & (v_n)^p \end{pmatrix} \tag{1.64}$$

c.-à.-d. que la k-ième colonne de A contient les composantes du vecteur $\vec{v}_k$ dans la base choisie, $\vec{v}_k = \sum_{i=1}^{p} (v_k)^i \vec{e}_i$.

Définition 22 (vecteurs lignes et colonnes d'une matrice). *Inversement, une matrice $A \in \mathcal{M}_K(p, n)$ arbitraire, peut être considérée indifféremment comme :*

— *la juxtaposition de n matrices colonnes $\mathbf{A}_j$, ou vecteurs colonnes, définis par :*

$$\left(\mathbf{A}_j\right)^i \stackrel{\text{déf.}}{=} A^i{}_j \tag{1.65}$$

— *la juxtaposition de p matrices lignes $\mathbf{A}^i$, ou vecteurs lignes, définis par :*

$$\left(\mathbf{A}^i\right)_j \stackrel{\text{déf.}}{=} A^i{}_j \tag{1.66}$$

Nous avons ainsi deux décompositions différentes d'une matrice A :

$$A = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_n \end{pmatrix} \tag{1.67a}$$

$$= \begin{pmatrix} \mathbf{A}^1 \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^p \end{pmatrix} . \tag{1.67b}$$

Ces deux notions sont naturellement échangées par transposition de la matrice A.

Définition 23 (rang d'une matrice). *Soit* $A \in n\mathcal{M}_K(p,n)$ *une matrice, que l'on considère comme la juxtaposition de ses* n *vecteurs colonnes :*

$$a^i{}_j = (\mathbf{A}_j)^i \tag{1.68}$$

Le rang de la matrice A *est le rang de la famille de vecteurs* $\{\vec{a}_1, \ldots, \vec{a}_n\}$ *dont les composantes, dans la base choisie, sont données par les vecteurs colonnes* $\mathbf{A}_1, \ldots, \mathbf{A}_n$.

Cette définition est cohérente avec la définition 14 du rang d'une application linéaire. En effet dans la décomposition (1.67a) de la matrice en vecteurs colonnes, $\mathbf{A}_k$ représente les composantes de la décomposition de $a(\vec{e}_k)$ dans la base choisie pour F. L'ensemble des vecteurs colonnes de la matrice correspond alors à l'ensemble des images des vecteurs de base de E, qui génèrent l'image de l'application linéaire. □

Rappelons que les formes linéaires sur E forment un espace vectoriel $E^\star$, espace dual de E. Si $\{\vec{e}_i, i = 1, \ldots, n\}$ est une base de E, alors la base de $E^\star$ appelée base duale est donnée par la famille de formes linéaires $\{e^j, j = 1, \ldots, n\}$ telles que

$$e^j(\vec{e}_i) = \delta^j{}_i \,. \tag{1.69}$$

Théorème 4 (rang des vecteurs lignes d'une matrice). *Soit* $A \in \mathcal{M}_K(p,n)$ *une matrice de rang* k*, c.-à.-d. que le rang de la famille de vecteurs* $\{\vec{a}_1, \ldots, \vec{a}_n\}$ *(dont les composantes, dans la base choisie, sont données par les* n *vecteurs colonnes* $\mathbf{A}_1, \ldots, \mathbf{A}_n$*), est égal à* k.

Le rang de la famile de vecteurs lignes de A*, c.-à.-d. le rang de la famille de formes linéaires* $\{\omega^1, \ldots, \omega^p\}$ *dont les composantes, dans la base duale à choisie, sont données par les* p *vecteurs lignes* $\mathbf{A}^1, \ldots, \mathbf{A}^p$ *est également égal à* k.

On considère, sans perte de généralité, que les k premiers vecteurs colonnes de A, soit les vecteurs colonnes $\mathbf{A}_1, \ldots, \mathbf{A}_k$, sont linéairement indépendants. On peut donc écrire tous les autres comme combinaison linéaire des k premiers :

$$\mathbf{A}_{k+r} = \sum_{j=1}^{k} \alpha_{j\,r} \mathbf{A}_j \,. \tag{1.70}$$

Considérons les éléments de la i-ème ligne de la matrice A. Nous obtenons alors les r derniers éléments de cette ligne comme combinaison linéaire des k premiers :

$$(\mathbf{A}_{k+r})^i = \sum_{j=1}^{k} \alpha_{j\,r} (\mathbf{A}_j)^i \quad \longleftrightarrow \quad a^i_{k+r} = \sum_{j=1}^{k} \alpha_{j\,r} a^i{}_j \,. \tag{1.71}$$

Le i-ème vecteur ligne $\mathbf{A}^i$ correspond à la forme linéaire ω^i dont le développement dans la base duale est donné par :

$$\omega^i = \sum_{j=1}^{n} a^i{}_j e^j \,. \tag{1.72}$$

Décomposons ce développement comme :

$$\omega^i = \sum_{j=1}^{k} a^i{}_j e^j + \sum_{r=1}^{n-k} \underbrace{a^i{}_{k+r}}_{=\sum_{j=1}^{k} \alpha_{j\,r} a^i{}_j} e^{k+r} = \sum_{j=1}^{k} a^i{}_j \Big(e^j + \sum_{r=1}^{n-k} \alpha_{j\,r} e^{k+r} \Big) \tag{1.73}$$

Nous avons donc montré que les formes linéaires $\{\omega^1, \ldots, \omega^p\}$ dont les composantes, dans la base choisie, sont données par les p vecteurs lignes $\mathbf{A}^1, \ldots, \mathbf{A}^p$, sont obtenues comme combinaisons linéaires des formes linéaires

$$\epsilon^j = e^j + \sum_{r=1}^{n-k} \alpha_{j\,r} e^{k+r} \,, \quad j = 1, \ldots, k \,, \tag{1.74}$$

donc le rang k' de la famille de vecteurs lignes $\mathbf{A}^1, \ldots, \mathbf{A}^p$ est inférieur ou égal à k. En reprenant le même raisonnement en intervertissant le rôle des vecteurs lignes et des vecteurs colonnes, on trouve que $k \leqslant k'$, d'où on déduit finalement que $k = k'$. □

1.5 Endomorphismes

On considère dans cette section plus spécifiquement les matrices associées aux endomorphismes, c.-à.-d. aux applications linéaires d'un K-espace vectoriel E dans lui-même. Ces matrices sont des *matrices carrées*, c.-à.-d. qu'elles possèdent le même nombre de ligne et de colonnes.

Remarque 1 (non-commutatitivité du produit matriciel). Soient $a : E \to E$ et $b : E \to E$ des endomorphismes. Les applications composées $(a \circ b)$ et $(b \circ a)$ sont toutes deux des applications linéaires de $E \to E$, c.-à.-d. des endomorphismes, et sont associées dans une base $\mathcal{B}_E$ à des matrices carrées, $A \cdot B$ et $B \cdot A$. Cependant, en général, ces deux matrices sont distinctes ; le produit matriciel de matrices carrées est dit *non-commutatif.*

Définition 24 (commutateur). *Soient* A *et* B *deux matrices carrées* $n \times n$*, c.-à.-d. des éléments de* $\mathcal{M}_K(n, n)$*. Le commutateur de ces deux matrices est défini comme :*

$$\big[A, B\big] \stackrel{\text{déf.}}{=} A \cdot B - B \cdot A \,. \tag{1.75}$$

Il est par construction antisymétrique, c.-à.-d. que $\big[A, B\big] = -\big[B, A\big]$.

Le commutateur de deux matrices est a priori distinct de la matrice nulle $\mathbb{O}_{n,n}$.

Exemple 8 (matrices de Pauli). En mécanique quantique il est fait grand usage des matrices de Pauli, qui sont trois matrices 2×2 à coefficients complexes définies par :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \,, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \,, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,. \tag{1.76}$$

On vérifie que ces matrices ne commutent pas entre elles. Nous avons par exemple

$$\sigma_1 \cdot \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{1.77}$$

et

$$\sigma_2 \cdot \sigma_1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \neq \sigma_1 \cdot \sigma_2 \tag{1.78}$$

Le commutateur de ces deux matrices est donc :

$$\left[\sigma_1,\, \sigma_2\right] = 2i\sigma_3\,. \tag{1.79}$$

On vérifie également les relations obtenues par permutation circulaire, c.-à.-d. $[\sigma_3,\, \sigma_1] = 2i\sigma_2$ et $[\sigma_2, \sigma_3] = 2i\sigma_1$, sont satisfaites.

* * *

Définition 25 (matrice identité) *La matrice identitié $n \times n$, notée $\mathbb{I}_n$, est la matrice associée à l'endomorphisme identité sur un espace vectoriel E de dimension n, c.-à.-d. l'application linéaire*

$$\mathcal{I} \;:\quad \begin{cases} E & \longrightarrow \quad E \\ \vec{v} & \longmapsto \quad \mathcal{I}(\vec{v}) = \vec{v} \end{cases}, \tag{1.80}$$

et ceci indépendamment de la base choisie sur E. Cette matrice est de la forme

$$\mathbb{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \longleftrightarrow \quad \left(\mathbb{I}_n\right)^i{}_j = \delta^i{}_j\,. \tag{1.81}$$

avec, rappelons-le,

$$\delta^i{}_j = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases}, \tag{1.82}$$

symbole de Kronecker.

Montrons que la matrice associée à l'application identité sur E est indépendante de la base choisie. Soit $\{\vec{e}_i, i = 1, \ldots, n\}$ une base quelconque de E. Nous avons

$$\forall i = 1, \ldots, n\,, \quad \mathcal{I}(\vec{e}_i) = \vec{e}_i = \sum_{j=1}^{n} \vec{e}_j\, \delta^j{}_i\,. \tag{1.83}$$

donc par identification $\left(\mathbb{I}_n\right)^i{}_j = \delta^i{}_j$. □

Propriété 12 *Pour toute matrice carrée $A \in \mathcal{M}_K(n,n)$, on vérifie que*

$$A \cdot \mathbb{I}_n = \mathbb{I}_n \cdot A = A \tag{1.84}$$

ainsi que

$$A \cdot \mathbb{O}_{n,n} = \mathbb{O}_{n,n} \cdot A = \mathbb{O}_{n,n} \,. \tag{1.85}$$

Pour démontrer la première propriété, on écrit simplement

$$(A \cdot \mathbb{I}_n)^i{}_j = \sum_{k=1}^{n} a^i{}_k \delta^k{}_j = a^i{}_j \ , \quad (\mathbb{I}_n \cdot A)^i{}_j = \sum_{k=1}^{n} \delta^i{}_k a^k{}_j = a^i{}_j \,. \tag{1.86}$$

La deuxième est évidente car tous les éléments de la matrice nulle sont nuls. □

Définition 26 (puissance et exponentielle d'un endomorphisme). *Soit a un endomorphisme associé à la matrice A dans une base de E. La puissance N-ième de a est définie comme l'endomorphisme*

$$a^N \stackrel{\text{déf.}}{=} \underbrace{(a \circ a \circ \cdots \circ a)}_{\text{N fois}} \tag{1.87}$$

et est représenté par la matrice

$$A^N \stackrel{\text{déf.}}{=} \underbrace{(A \cdot A \cdot \cdots \cdot A)}_{\text{N fois}} \tag{1.88}$$

Par convention, on pose $A^0 = \mathbb{I}_n$.

L'exponentielle de a (et donc de la matrice A) est définie par le développement en série entière de l'exponentielle,

$$\exp(a) = \sum_{k=0}^{\infty} \frac{a^k}{k!} \implies \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} \tag{1.89}$$

À noter que, pour des matrices, $\exp(A + B) \neq \exp(A) \cdot \exp(B)$ a priori.

Exemple 9 Soient A et B deux matrices carrées $n \times n$. Nous avons

$$(A+B)^2 = (A+B)\cdot(A+B) = A^2 + B^2 + A \cdot B + B \cdot A = A^2 + B^2 + 2A \cdot B - [A,B] \,. \tag{1.90}$$

De manière générale, la formule du binôme de Newton ne s'applique pas si les matrices ne commutent pas, c.-à.-d. si $[A, B] \neq \mathbb{O}_{n,n}$.

*
* *

Remarque 2 (matrice associée à un projecteur). *La matrice M associée à un projecteur satisfait l'équation $M^2 = M$ (voir le théorème 2).*

Définition 27 (matrice inverse). *Soit* $\mathsf{A} \in \mathcal{M}_{\mathsf{K}}(n,n)$ *une matrice carrée. La matrice inverse de* A*, si elle existe, est la matrice* $\mathsf{A}^{-1} \in \mathcal{M}_{\mathsf{K}}(n,n)$ *telle que :*

$$\mathsf{A} \cdot \mathsf{A}^{-1} = \mathsf{A}^{-1} \cdot \mathsf{A} = \mathbb{I}_n . \tag{1.91}$$

On notera que la matrice inverse d'une matrice carrée n'est pas nécessairement définie.

Propriété 13 (matrice associée à un isomorphisme). *Soit* $a : \mathsf{E} \to \mathsf{E}$ *un endomorphisme d'un espace vectoriel* E*. L'application* a *est un isomorphisme, c.-à.-d. un endomorphisme bijectif, si et seulement si la matrice* A *associée, dans une base quelconque, est inversible.*

Soit $\{\vec{e}_1, \ldots, \vec{e}_n\}$ une base de E. La matrice A est donnée par la décomposition de l'image des vecteurs de base dans la même base :

$$a(\vec{e}_i) = \sum_{j=1}^{n} \vec{e}_j a^j{}_i . \tag{1.92}$$

L'application étant bijective, il existe une application inverse, notée a^{-1}, telle que $a \circ a^{-1} = a^{-1} \circ a = \mathcal{I}$. Nous avons d'une part

$$\begin{aligned}(a^{-1} \circ a)(\vec{e}_i) = a^{-1}\Big(\sum_{j=1}^{n} \vec{e}_j a^j{}_i\Big) = \sum_{j=1}^{n} a^j{}_i \sum_{k=1}^{n} \vec{e}_k (a^{-1})^k{}_j = \sum_{k=1}^{n} \vec{e}_k \left(\sum_{j=1}^{n} (a^{-1})^k{}_j a^j{}_i\right) \\ = \sum_{k=1}^{n} \vec{e}_k \left(\mathsf{A}^{-1} \cdot \mathsf{A}\right)^k{}_i \end{aligned} \tag{1.93}$$

et d'autre part

$$(a^{-1} \circ a)(\vec{e}_i) = \mathcal{I}(\vec{e}_i) = \vec{e}_i = \mathbb{I}_n(\vec{e}_i) \tag{1.94}$$

donc nous avons montré que la matrice A^{-1} associée à l'application réciproque a^{-1} vérifie $\mathsf{A}^{-1} \cdot \mathsf{A} = \mathbb{I}_n$. On montre de même en appliquant d'abord a^{-1} puis a au vecteur $\vec{e}_i$ que $\mathsf{A} \cdot \mathsf{A}^{-1} = \mathbb{I}_n$ et donc que A^{-1} est la matrice inverse de la matrice A.

La réciproque est évidente. Supposons que la matrice A associée à l'application a admette un inverse A^{-1}. Alors l'application a admet une application réciproque a^{-1} dont la matrice associée dans la base choisie est donnée par A^{-1}. En effet, soient $\vec{v}$ et $\vec{w}$ deux vecteurs de E, et $\mathbf{V}$, $\mathbf{W}$ les vecteurs colonnes associés. Nous avons alors

$$a(\vec{v}) = \vec{w} \;\Leftrightarrow\; \mathsf{A} \cdot \mathbf{V} = \mathbf{W} \;\Leftrightarrow\; (\mathsf{A}^{-1} \cdot \mathsf{A}) \cdot \mathbf{V} = \mathsf{A}^{-1}\mathbf{W} \;\Leftrightarrow\; \mathbf{V} = \mathsf{A}^{-1}\mathbf{W} \;\Leftrightarrow \vec{v} = a^{-1}(\vec{w}) , \tag{1.95}$$

où l'application linéaire a^{-1} est entièrement déterminée par sa matrice A^{-1} dans la base choisie. □

Propriété 14 (inverse d'un produit de matrices). *Soient* A *et* B *deux matrices inversibles. Alors la matrice* $\mathsf{B} \cdot \mathsf{A}$ *est également inversible, et son inverse est donné par :*

$$(\mathsf{B} \cdot \mathsf{A})^{-1} = \mathsf{A}^{-1} \cdot \mathsf{B}^{-1} . \tag{1.96}$$

Nous avons en effet, par associativité

$$\big(A^{-1} \cdot B^{-1}\big) \cdot \big(B \cdot A\big) = A^{-1} \cdot (B^{-1} \cdot B) \cdot A = A^{-1} \cdot A = \mathbb{I}_n\,, \tag{1.97}$$

et de manière identique $(B \cdot A) \cdot (A^{-1} \cdot B^{-1}\big) = \mathbb{I}_n$. □

Propriété 15 (inverse de la transposée d'une matrice). *Soit A une matrice inversible. Alors la matrice transposée A^{T} est également inversible, et son inverse est donné par :*

$$\big(A^{\mathrm{T}}\big)^{-1} = \big(A^{-1}\big)^{\mathrm{T}} \tag{1.98}$$

Nous avons en effet d'après la propriété 11 :

$$A^{\mathrm{T}}\big(A^{-1}\big)^{\mathrm{T}} = \big(A^{-1} \cdot A\big)^{\mathrm{T}} = \mathbb{I}_n^{\mathrm{T}} = \mathbb{I}_n\,, \tag{1.99}$$

la transposée de la matrice identité étant la matrice identité elle-même (comme les éléments non-nuls sont tous diagonaux). On montre de même que $(A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = \mathbb{I}_n$. □

1.6 Changement de base

Si $a : E \to F$ est une application linéaire, on peut lui associer une matrice A une fois une base $\mathcal{B}_E = \{\vec{e}_i, i = 1, \dots, n\}$ de l'espace vectoriel de départ et une base $\mathcal{B}_F = \{\vec{f}_j, j = 1, \dots, p\}$ de l'espace vectoriel d'arrivée ont été choisies. Évidemment, le choix de base étant, dans les deux cas, arbitraire, la forme de la matrice dépend de la base choisie. Nous examinerons ici la relation entre les matrices associées à différent choix de bases.

Définition 28 (matrice de passage). *Soient $\mathcal{B}_E = \{\vec{e}_i, i = 1, \dots, n\}$ et $\widehat{\mathcal{B}}_E = \{\vec{\epsilon}_i, i = 1, \dots, n\}$ deux bases d'un même espace vectoriel E. Tout vecteur de $\widehat{\mathcal{B}}_E$ peut s'exprimer comme une combinaison linéaire des vecteurs de $\mathcal{B}_E$:*

$$\vec{\epsilon}_i = \sum_{j=1}^{n} \vec{e}_j P^j{}_i\,. \tag{1.100}$$

Les coefficients $\{P^j{}_i \in K, i = 1, \dots, n;\ j = 1, \dots, n\}$ apparaissant dans cette expression sont les composantes d'une matrice carrée P, appelée matrice de passage (ou matrice de changement de base) de $\mathcal{B}_E$ vers $\widehat{\mathcal{B}}_E$.

On peut ainsi voir le changement de base comme un endomorphisme $\mathcal{P}$ de l'espace vectoriel E, dont P est la matrice dans la base $\mathcal{B}_E$.

Propriété 16 (matrice de passage inverse). *Soit P, appelée la matrice de passage d'une base $\mathcal{B}_E$ vers une base $\widehat{\mathcal{B}}_E$. La matrice P est inversible, et la matrice de passage inverse P^{-1} est telle que :*

$$\vec{e}_i = \sum_{j=1}^{n} \vec{\epsilon}_j (P^{-1})^j{}_i\,. \tag{1.101}$$

En effet, étant donné que $\widehat{\mathcal{B}}_E$ est une base, tout vecteur de $\mathcal{B}_E$ peut s'exprimer comme une combinaison linéaire :

$$\vec{e}_i = \sum_{j=1}^{n} \vec{\epsilon}_j M^j{}_i \,. \tag{1.102}$$

En remplaçant

$$\vec{e}_i = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} \vec{e}_k P^k{}_j \right) M^j{}_i = \sum_{k=1}^{n} \vec{e}_k \left(\sum_{j=1}^{n} P^k{}_j M^j{}_i \right) \tag{1.103}$$

Les vecteurs de $\mathcal{B}_E$ étant linéairement indépendants, on en déduit que

$$\left(\sum_{j=1}^{n} P^k{}_j M^j{}_i \right) = \delta^k{}_j \implies P \cdot M = \mathbb{I}_n \,. \tag{1.104}$$

On montre de même que

$$\vec{\epsilon}_i = \sum_{j=1}^{n} \vec{e}_j P^j{}_i = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} \vec{\epsilon}_k M^k{}_j \right) P^j{}_i = \sum_{k=1}^{n} \vec{\epsilon}_k \left(\sum_{j=1}^{n} M^k{}_j P^j{}_i \right) \implies M \cdot P = \mathbb{I}_n \tag{1.105}$$

donc $M = P^{-1}$. □

Propriété 17 (changement de base pour la matrice d'une application linéaire). *Soit* $a : E \to F$ *est une application linéaire associée, une fois choisie une base* $\mathcal{B}_E = \{\vec{e}_i, i = 1, \ldots, n\}$ *de l'espace vectoriel de départ et une base* $\mathcal{B}_F = \{\vec{f}_j, j = 1, \ldots, p\}$ *de l'espace d'arrivée, à une matrice* $A \in \mathcal{M}_K(p, n)$.

On considère un changement de bases, tant pour l'espace de départ que pour l'espace d'arrivée. On choisit dans E *la base* $\widehat{\mathcal{B}}_E = \{\vec{\epsilon}_i, i = 1, \ldots, n\}$ *définie par une matrice de passage* P *de taille* $n \times n$ *:*

$$\vec{\epsilon}_i = \sum_{k=1}^{n} \vec{e}_k P^k{}_i \,. \tag{1.106}$$

On choisit dans F *la base* $\widehat{\mathcal{B}}_F = \{\vec{\phi}_j, j = 1, \ldots, p\}$ *définie par une matrice de passage* R *de taille* $p \times p$ *:*

$$\vec{\phi}_j = \sum_{\ell=1}^{p} \vec{f}_\ell R^\ell{}_j \,. \tag{1.107}$$

La matrice associée à l'application linéaire a *pour le choix de bases* $\widehat{\mathcal{B}}_E$ *et* $\widehat{\mathcal{B}}_F$ *est donnée par :*

$$\boxed{A' = R^{-1} A P} \tag{1.108}$$

Les deux matrices A *et* A' *étant deux représentations différentes d'une même application, elles sont dites équivalentes.*

La demonstration de cette propriété est simple. Par linéarité de l'application a nous avons premièrement :

$$a(\vec{\epsilon}_i) = a\left(\sum_{j=1}^{n} \vec{e}_j P^j{}_i\right) = \sum_{j=1}^{n} P^j{}_i a\left(\vec{e}_j\right) . \tag{1.109}$$

La matrice A étant associée à l'application pour le choix de bases $\mathcal{B}_E$ et $\mathcal{B}_F$ nous avons

$$a\left(\vec{e}_j\right) = \sum_{k=1}^{p} \vec{f}_k A^k{}_j . \tag{1.110}$$

Nous avons donc par substitution dans la relation précédente :[1]

$$a(\vec{\epsilon}_i) = \sum_{j=1}^{n} P^j{}_i \sum_{k=1}^{p} \vec{f}_k (A)^k{}_j = \sum_{k=1}^{p} \vec{f}_k \big(A \cdot P\big)^k{}_i \tag{1.111}$$

Il suffit alors d'exprimer $\vec{f}_k$ dans la base $\widehat{\mathcal{B}}_F$ en utilisant la matrice de passage inverse R^{-1} :

$$a(\vec{\epsilon}_i) = \sum_{k=1}^{p} \left(\sum_{\ell=1}^{p} \vec{\phi}_\ell \big(R^{-1}\big)^\ell{}_k\right) \big(A \cdot P\big)^k{}_j = \sum_{\ell=1}^{p} \vec{\phi}_\ell \big(R^{-1} \cdot A \cdot P\big)^\ell{}_k , \tag{1.112}$$

indiquant, par identification, que $A' = R^{-1} \cdot A \cdot P$. □

Propriété 18 (changement de base pour la matrice d'un endomorphisme). *Soit* $a : E \to E$ *un endomorphisme associé une fois choisie une base* $\mathcal{B}_E = \{\vec{e}_i, i = 1, \ldots, n\}$*, à une matrice* $A \in \mathcal{M}_K(p, n)$*.*

On considère une autre base de E*,* $\widehat{\mathcal{B}}_E = \{\vec{\epsilon}_i, i = 1, \ldots, n\}$ *définie par une matrice de passage* P *de taille* $n \times n$ *:*

$$\vec{\epsilon}_i = \sum_{k=1}^{n} \vec{e}_k P^k{}_i . \tag{1.113}$$

La matrice associée à l'application linéaire a *dans la base* $\widehat{\mathcal{B}}_E$ *est donnée par :*

$$\boxed{A' = P^{-1} A P} \tag{1.114}$$

Deux matrices A *et* A' *reliées par une telle relation sont dites semblables.*

Il s'agit bien évidemment d'un cas particulier du cas précédent.

1. Attention contrairement à ce que la notation laisse suggérer, $\vec{f}_k (A)^k{}_j$ est le produit du vecteur $\vec{f}_k$ par le scalaire $(A)^k{}_j$.

1.7 Trace d'un endomorphisme

Un endomorphisme a d'un espace vectoriel E est représenté dans une base $\mathcal{B}_E$ donnée par une matrice carrée A.

L'expression de cette matrice dépendant de la base choisie dans E, il est souhaitable de disposer de quantités indépendantes de ce choix, qui contiennent cependant une information « utile » concernant l'endomorphisme en question. Deux quantités de ce type que nous étudierons dans ce cours sont la trace d'un endomorphisme que nous définirons ici, et le déterminant qui sera l'objet du chapitre suivant.

Définition 29 (trace d'un endomorphisme). *Soit* $a : E \to E$ *un endomorphisme, représenté dans une base* $\mathcal{B}_E$ *par une matrice* $A \in \mathcal{M}_K(n,n)$. *La trace de* a*, notée* $\mathrm{Tr}(a)$*, est définie par :*

$$\mathrm{Tr}(a) \overset{\text{déf.}}{=} \sum_{i=1}^{n} A^i{}_i\,, \tag{1.115}$$

c.-à.-d. par la somme des éléments diagonaux de la matrice A. *Comme la définition le suggère, cette quantité est indépendante de la base choisie pour le calcul. On peut aussi écrire cette quantité comme* $\mathrm{Tr}\,A$.

Dans une autre base $\widehat{\mathcal{B}}_E$, définie par la matrice de passage P, l'endomorphisme a est représenté par la matrice $A' = P^{-1} \cdot A \cdot P$, d'après la la propriété 18. Nous voyons immédiatement que

$$\sum_{i=1}^{n} \left(A'\right)^i{}_i = \sum_{i=1}^{n} \left(P^{-1} \cdot A \cdot P\right)^i{}_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(P^{-1}\right)^i{}_j \sum_{k=1}^{n} A^j{}_k P^k{}_i = \sum_{k,j=1}^{n} \underbrace{\left(\sum_{i=1}^{n} P^k{}_i \left(P^{-1}\right)^i{}_j\right)}_{=\delta^k{}_j} A^j{}_k$$

$$= \sum_{j=1}^{n} A^j{}_j = \mathrm{Tr}(a)\,. \tag{1.116}$$

Propriété 19 (cyclicité de la trace). *Soient* a *et* b *deux endomorphismes sur un espace vectoriel* E.

$$\mathrm{Tr}(a \circ b) = \mathrm{Tr}(b \circ a)\,, \tag{1.117}$$

qui se généralise à toute permutation circulaire $(a \circ b \circ c \cdots)$ *d'endomorphismes.*

En termes matriciels, nous souhaitons démontrer que

$$\mathrm{Tr}(A \cdot B) = \mathrm{Tr}(B \cdot A)\,. \tag{1.118}$$

Nous avons

$$\mathrm{Tr}(A \cdot B) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} A^i{}_j B^j{}_i\right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} B^j{}_i A^i{}_j\right) = \mathrm{Tr}(B \cdot A) \tag{1.119}$$

□

Propriété 20 *D'autres propriétés utiles de la trace sont les suivantes :*

1. $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$
2. $\mathrm{Tr}(\lambda A) = \lambda \mathrm{Tr}(A)$
3. $\mathrm{Tr}\big(A^{T}\big) = \mathrm{Tr}(A)$

Leur demonstration, élémentaire, est laissée en exercice. Attention, la trace d'un produit de matrices $A \cdot B$ n'est pas égale au produit de leurs traces !

Exemple 10 (trace des matrices de Pauli). Reprenons les matrices de Pauli, matrices 2×2 étudiées à l'exemple 8 :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.120}$$

On voit immédiatement que ces trois matrices sont de trace nulle.

1.8 Bestiaire de matrices

Terminons cette section par un bestiaire de matrices carrées, donc associées à des endomorphismes, possédant des propriétés remarquables.

Définition 30 (matrices symétriques et antisymétriques). *Une matrice symétrique* S *est égale à sa transposée,*

$$S^{T} = S\,, \tag{1.121}$$

alors qu'une matrice antisymétrique A *lui est opposée,*

$$A^{T} = -A\,. \tag{1.122}$$

Propriété 21 Notons quelques propriétés élémentaires de ces matrices symétriques et antisymétriques :

1. Si A est symétrique, elle est entièrement caractérisée par les éléments diagonaux et les éléments au-dessus de la diagonale, car $A^{i}{}_{j} = A^{j}{}_{i}$; elle possède ainsi $n(n+1)/2$ composantes indépendantes.
2. Si A est antisymétrique, elle est entièrement caractérisée par les éléments au-dessus de la diagonale, $A^{i}{}_{j} = -A^{j}{}_{i}$ et les éléments diagonaux sont tous nuls, car $A^{i}{}_{i} = -A^{i}{}_{i} = 0$; elle possède ainsi $n(n-1)/2$ composantes indépendantes.
3. En conséquence, la trace d'une matrice antisymétrique est nulle (la réciproque est fausse !).
4. Toute matrice carrée se décompose en somme d'une matrice symétrique et d'une matrice antisymétrique :

$$M = \underbrace{\frac{A + A^{T}}{2}}_{S} + \underbrace{\frac{A - A^{T}}{2}}_{A} \tag{1.123}$$

5. Si S est une matrice symétrique, pour tout $N \in \mathbb{N}^\star$, S^N est une matrice symétrique.
6. Si A est une matrice antisymétrique, pour tout N pair, A^N est une matrice symétrique et pour tout N impair A^N est une matrice antisymétrique. On a en effet $(A^N)^\mathsf{T} = (-A)^N = (-1)^N A^N$.

Définition 31 (matrices diagonales). *Une matrice carrée* D *est diagonale si elle est de la forme :*

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{1.124}$$

ce qu'on peut noter

$$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\,. \tag{1.125}$$

C'est évidemment un cas particulier de matrice symétrique.

Propriété 22 Notons quelques propriétés immédiates mais néanmoins importantes des matrices diagonales :

1. Si D est une matrice diagonale,

$$D^N = \mathrm{diag}(\lambda_1^N, \lambda_2^N, \ldots, \lambda_n^N) \tag{1.126}$$

2. Si D est une matrice diagonale,

$$\exp(D) = \mathrm{diag}(\exp\lambda_1, \exp\lambda_2, \ldots, \exp\lambda_n)\,. \tag{1.127}$$

   Ce résultat se généralise directement aux fonctions développables en série entière de rayon de convergence infini (par exemple le cosinus ou le sinus).

3. Si D est une matrice diagonale,

$$\mathrm{Tr}\, D = \lambda_1 + \lambda_2 + \cdots + \lambda_n\,. \tag{1.128}$$

Définition 32 (endomorphisme nilpotent). *Soit* a *un endomorphisme d'un espace vectoriel* E*, associé dans une certaine base* $\mathcal{B}_E$ *à une matrice carrée* A*. Cet endomorphisme est dit nilpotent s'il existe un entier* $N > 1$ *tel que*

$$a^N = 0\,, \tag{1.129}$$

où 0 *désigne ici l'application nulle* $\vec{v} \to \vec{0}_E$*, ou de manière équivalente si*

$$A^N = \mathbb{O}_{n,n}\,. \tag{1.130}$$

Notons que la caractérisation matricielle, eqn. (1.130), est indépendante de la base choisie, car la matrice nulle $\mathbb{O}_{n,n}$ est invariante par changement de base : $P^{-1} \cdot \mathbb{O}_{n,n} \cdot P = \mathbb{O}_{n,n}$.

Exemple 11 Il est important de remarquer qu'une matrice peut être de carré nul sans être nulle pour autant. À partir des matrices de Pauli, voir l'éqn. (1.120), construisons les matrices :

$$\sigma_+ \stackrel{\text{déf.}}{=} \sigma_1 + i\sigma_2 = 2\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad \sigma_- \stackrel{\text{déf.}}{=} \sigma_1 - i\sigma_2 = 2\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{1.131}$$

On vérifie immédiatement que $\left(\sigma_\pm\right)^2 = \mathbb{O}_{2,2}$.

Cours n°2

Déterminant

L'objectif de ce chapitre est d'étudier le déterminant plus en détail que nous l'avions fait au premier semestre, et d'en déduire une méthode d'inversion de matrices.

Nous en profiterons pour donner quelques propriétés du groupe symétrique, ce qui constituera une première sensibilisation à la théorie de groupes, sur laquelle nous reviendrons à la fin du semestre si le temps le permet.

2.1 Généralités sur les groupes

Les groupes jouent un rôle fondamental en physique, étant donné que les *symétries* des systèmes physiques s'organisent dans cette structure mathématique ; la connaissance de ses propriétés permet de comprendre plus en profondeur l'origine de concepts comme la masse ou le spin des particules.

Nous commencerons par donner quelques notions générales avant de spécialiser la discussion au groupe symétrique, qui apparaît dans l'étude du déterminant.

Définition 1 (groupe). *Un groupe* $(\mathsf{G}, \cdot)$ *est un ensemble* G *muni d'une loi de composition interne :*

$$\cdot \quad \begin{cases} \mathsf{G} \times \mathsf{G} & \to \quad \mathsf{G} \\ (g_1, g_2) & \mapsto \quad g_1 \cdot g_2 \end{cases} \tag{2.1}$$

satisfaisant les propriétés suivantes :

1. *associativité,* $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$ *;*
2. *existence d'un élément neutre* $e \in \mathsf{G}$ *tel que, pour tout* $g \in \mathsf{G}$*,* $e \cdot g = g \cdot e = g$ *;*
3. *existence d'un inverse pour tout élément* $g \in \mathsf{G}$*, noté* g^{-1}*, tel que* $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Remarque 1 *Suivant les cas, la loi de composition interne peut désigner soit la multiplication « $\times$ » (de nombres ou matrices) soit la composition d'applications « $\circ$ », soit, de manière contre-intuivite, l'addition « $+$ ».*

Définition 2 (groupe abélien) *Un groupe* $(G, \cdot)$ *dont tous les éléments commutent, c.-à.-d. tel que*

$$\forall (g_1, g_2) \in G^2 \ , \quad g_1 \cdot g_2 = g_2 \cdot g_1 \tag{2.2}$$

est dit abélien ou commutatif. Dans le cas contraire il est dit non-abélien ou non-commutatif.

Exemple 1 Donnons quelques exemples simples de groupes :

1. les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ munis de l'addition usuelle sont des groupes ; le neutre est simplement le zéro usuel, et l'inverse de tout élément x, au sens de la définition précédente, est $(-x)$. Ces groupes sont commutatifs car l'addition l'est ;
2. Les ensembles $\mathbb{Q}^\star$, $\mathbb{R}^\star$ et $\mathbb{C}^\star$ munis de la multiplication usuelle sont des groupes ; le neutre est $e = 1$, et l'inverse de tout élément x, est évidemment $1/x$. Ces groupes sont commutatifs car la multiplication l'est.
3. l'ensemble $GL(n, K)$ des matrices $n \times n$ inversibles à coefficients dans K, muni de la multiplication matricielle $\cdot$, forme un groupe. En effet :
   a) le produit matriciel est associatif ;
   b) la matrice identité $\mathbb{I}_n$ est telle que, pour tout $M \in GL(n, K)$, $\mathbb{I}_n \cdot M = M \cdot \mathbb{I}_n$;
   c) l'inverse de $M \in GL(n, K)$ est évidemment la matrice inverse M^{-1}.

   Ce groupe est non-abélien (car la multiplication matricielle ne commute pas).

Exemple 2 (rotation dans le plan). En coordonnées complexes, une rotation dans le plan centrée en l'origine correspond à la transformation $z \to e^{i\theta} z$ où $\theta \in [0, 2\pi[$ est l'angle de rotation. On peut associer à chaque valeur de θ un élément du groupe des rotations que nous noterons $R(\theta)$. Si nous effectuons deux rotations successives dans le plan, nous avons $z \to e^{i\theta'}(e^{i\theta} z) = e^{i(\theta+\theta')} z = e^{i\theta}(e^{i\theta'} z)$. Les rotations dans le plan forment donc un groupe commutatif, et la loi de composition est donnée par :

$$R(\theta) \cdot R(\theta') = R(\theta') \cdot R(\theta) = R(\Theta) \ , \quad \Theta \equiv \theta + \theta' \mod 2\pi \ , \quad \Theta \in [0, 2\pi[\, . \tag{2.3}$$

*
* *

Propriété 1 *Tout groupe* $(G, \cdot)$ *satisfait les propriétés suivantes :*

1. *unicité de l'élément neutre ;*
2. *unicité de l'inverse* g^{-1} *de tout élément* $g \in G$ *;*
3. *le neutre est son propre inverse,* $e^{-1} = e$ *;*
4. *pour tout* $g \in G$*,* $(g^{-1})^{-1} = g$ *;*
5. *pour tout* $(g_1, g_2) \in G^2$*,* $(g_1 \cdot g_2)^{-1} = g_2^{-1} g_1^{-1}$*.*

Les demonstrations de ces propriétés sont évidentes.

1. Supposons l'existence d'un autre élément neutre e'. Comme e est un élément neutre, $e \cdot e' = e'$. Comme e' l'est aussi, $e \cdot e' = e$. Donc $e = e'$.

2. Supposons l'existence de deux inverses h et h' pour un même élément $g \in G$, c.-à.-d. que nous avons $g \cdot h = e$ et $g \cdot h' = e$, donc $g \cdot h = g \cdot h'$. Multiplions cette égalité à gauche par h on a $h \cdot (g \cdot h) = h \cdot (g \cdot h')$, soit, par associativité, $(h \cdot g) \cdot h = (h \cdot g) \cdot h'$ d'où, h étant inverse de g, $h = h'$;
3. Comme e est l'élément neutre, $e \cdot e = e$, donc par unicité (propriété 2) e est son propre inverse ;
4. Les relations $g \cdot g^{-1} = e$ et $g^{-1} \cdot g = e$ peuvent s'interpréter comme exprimant le fait que g est l'inverse de g^{-1} ;
5. Par associativité, $(g_2^{-1} \cdot g_1^{-1})(g_1 \cdot g_2) = g_2^{-1} \cdot (g_1^{-1} \cdot g_1) \cdot g_2 = g_2^{-1} \cdot g_2 = e$. On montre de même que $(g_1 \cdot g_2) \cdot (g_2^{-1} \cdot g_1^{-1}) = e$. ☐

Définition 3 (ordre d'un groupe). *L'ordre d'un groupe* G*, noté* $|G|$*, est le nombre de ses éléments (qui n'est pas nécessairement fini).*

Définition 4 (sous-groupe). *Soit* $(G, \cdot)$ *un groupe et* $H \subset G$ *une partie non-vide de* G*.* $(H, \cdot)$ *est un sous-groupe de* $(G, \cdot)$ *si et seulement si :*

1. *pour tout* $(h, h') \in H$*,* $h \cdot h' \in H$ *;*
2. $e \in H$ *;*
3. *pour tout* $h \in H$*,* $h^{-1} \in H$*.*

En d'autres termes, $H \subset G$ *est stable sous la loi de composition interne « · » du groupe* G *et* $(H, \cdot)$ *est lui-même un groupe.*

Exemple 3 Comme nous l'avons vu, l'ensemble $(\mathbb{Z}, +)$ des entiers relatifs muni de l'addition forme un groupe. Considérons $n\mathbb{Z}$, pour un certain entier positif n, défini comme l'ensemble des entiers multiples de n. Montrons que $(n\mathbb{Z}, +)$ forme un sous-groupe de $(\mathbb{Z}, +)$. Premièrement, $n\mathbb{Z}$ est de manière évidente une partie non-vide de $\mathbb{Z}$, qui contient zéro. Deuxièmement, si $k \in n\mathbb{Z}$ et $k' \in n\mathbb{Z}$, il existe $r \in \mathbb{Z}$ et $r' \in \mathbb{Z}$ tels que $k = rn$ et $k' = r'n$; en conséquence $k + k' = n(r + r') \in n\mathbb{Z}$. Troisièmement, pour $k \in n\mathbb{Z}$, $k = rn$ donc $-k = (-r)n \in n\mathbb{Z}$. ☐

Définition 5 (sous groupe engendré). *Soit* G *un groupe et* $S \subset G$ *une partie de* G *(qui n'est pas nécessairement un sous-groupe). Le sous-groupe engendré par* S*, noté* $\langle S \rangle$*, est le plus petit sous-groupe de* G *contenant tous les éléments de* S*.*

Définition 6 (sous-groupe cyclique). *Soit un élément* $g \in G$*. Le sous-groupe engendré par* g*, noté* $\langle g \rangle$*, contient les différentes puisssances de* g*, c.-à.-d.* $g^0 = e$*,* g*,* g^2*, etc. Le plus petit* $k \in \mathbb{N}^\star$*, s'il existe, tel que* $x^k = e$*, s'appelle l'ordre de* g *dans* G *et égal à l'ordre du sous-groupe* $\langle g \rangle$*. Le sous-groupe* $\langle g \rangle$ *est un exemple de groupe cyclique, c.-à.-d. généré par un élément unique.*

Remarque 2 *Si l'ordre du groupe* $|G|$ *est fini, alors l'ordre de tout élément* g *est nécessairement fini (autrement le groupe* G *contiendrait une infinité d'éléments).*

Exemple 4 (rotations discrètes). On considère à nouveau le groupe des rotations dans le plan, voir l'exemple 2. Pour tout $n \in \mathbb{N}^\star$, l'élément du groupe $r_n \overset{\text{déf.}}{=} R(2\pi/n)$, correspondant à la rotation discrète

$$z \overset{r_n}{\longmapsto} e^{2i\pi/n} z\,, \tag{2.4}$$

engendre le sous-groupe cyclique $\langle r_n \rangle = \{1, r_n, r_n^2, \ldots, r_n^{n-1}\}$ des rotations d'ordre n. L'ordre de ce sous-groupe est évidemment égal à n.

Exemple 5 (groupe $\mathbb{Z}/n\mathbb{Z}$). Pour tout $n \in \mathbb{N}^\star$, on définit le groupe $\mathbb{Z}/n\mathbb{Z}$ comme le groupe additif des entiers relatifs identifés modulo n. Cela signifie qu'un élément de $\mathbb{Z}/n\mathbb{Z}$ correspond à la classe des entiers de la forme $k + nr$ avec $r \in \mathbb{Z}$, et qu'on peut le désigner par un représentant k dans l'intervalle $[0, n-1]$. La loi de composition sur ce groupe est héritée de l'addition sur les entiers,

$$k + k' \equiv k'' \mod n\,, \tag{2.5}$$

et k'', représentant de la classe des entiers de la forme $k'' + n\mathbb{Z}$, est choisi dans le même intervalle que k et k'.

Par exemple, le groupe $\mathbb{Z}/3\mathbb{Z}$ contient trois éléments que l'on peut noter $\{0, 1, 2\}$. Les lois de composition sur ce groupe sont données par la table :

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

*
* *

Définition 7 (morphisme de groupes). *Soient* $(G, \cdot)$ *et* $(H, \circ)$ *des groupes et* $\varphi : G \to H$ *une application de* G *dans* H*. Cette application est un morphisme de groupes si :*

$$\forall g, g' \in G\,, \quad \varphi(g \cdot g') = \varphi(g) \circ \varphi(g')\,. \tag{2.6}$$

Lorsque cette application est bijective, elle est appelée isomorphisme de groupes.

Exemple 6 On considère l'application

$$\varphi : \left\{ \begin{array}{lcl} \langle r_n \rangle & \to & \mathbb{Z}/n\mathbb{Z} \\ r_n^k & \mapsto & k \end{array} \right. \tag{2.7}$$

c.-à.-d. qui associe à la rotation $z \mapsto e^{2i\pi k/n} z$, avec $k \in 0, \ldots, n-1$, l'élement k correspondant du groupe $\mathbb{Z}/n\mathbb{Z}$. Cette application est évidemment bijective, et correspond à un isomorphisme entre les deux groupes car $\varphi(r_n^k \cdot r_n^{k'}) = \varphi(r_n^{k+k'}) = k + k'$.

Des groupes isomorphes possèdent les mêmes propriétés. On voit que $\mathbb{Z}/n\mathbb{Z}$ est également un groupe cyclique d'ordre n, dont on peut prendre comme générateur $1 \in \mathbb{Z}/n\mathbb{Z}$.

2.2 Groupe symétrique

Définition 8 (permutation). *Une permutation* σ *de l'ensemble* $\{1, 2, \ldots, n\}$ *est une bijection de* $\{1, 2, \ldots, n\}$ *dans lui-même.*

Il est d'usage de noter une permutation σ de n éléments de la façon suivante (attention à ne pas comprendre cette notation comme une matrice !) :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}, \tag{2.8}$$

c.-à.-d. qu'on donne en-dessous de chaque entier $k \in \{1, \ldots, n\}$ son image $\sigma(k)$ par la permutation.

Étant donné que les permutations sont des applications de l'ensemble $\{1, 2, \ldots, n\}$ dans lui-même, il est possible de définir de définir la composition $\sigma' \circ \sigma$ de deux permutations, avec

$$(\sigma' \circ \sigma)(k) \overset{\text{déf.}}{=} \sigma'\big(\sigma(k)\big) . \tag{2.9}$$

Exemple 7 Soient les permutations de 4 éléments,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} , \quad \sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} . \tag{2.10}$$

Nous avons par exemple $(\sigma' \circ \sigma)(1) = \sigma'\big(\sigma(1)\big) = \sigma'(2) = 4$. Il est possible de calculer rapidement le résultat de la composition de deux permutations en adoptant la notation suivante, adaptée de (2.8) :

$$\sigma' \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 4 & 1 & 3 & 2 \end{pmatrix} . \tag{2.11}$$

On peut calculer de la même manière

$$\sigma \circ \sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \\ 1 & 3 & 4 & 2 \end{pmatrix} , \tag{2.12}$$

et on remarque en particulier que le produit de permutations n'est pas commutatif, c.-à.-d. que $\sigma \circ \sigma' \neq \sigma' \circ \sigma$ en général.

Définition 9 (groupe symétrique S_n). *Le groupe symétrique* S_n *est défini comme l'ensemble des permutations de* n *éléments muni du produit* $\circ$ *associé à la composition de permutations. L'ordre de ce groupe est* $|S_n| = n!$*, correspondant au nombre de permutations de* n *éléments.*

Montrons que S_n forme un groupe. L'associativité de la loi de composition provient de l'associativité de la composition d'applications. L'élément neutre du groupe correspond à la permutation identité ι, telle que $\iota(k) = k$ pour tout $k \in \{1, \ldots, n\}$. Enfin, l'existence d'un inverse est garantie car les permutations sont, par définition, des bijections de $\{1, \ldots, n\}$ dans lui-mm̂e. $\square$.

Exemple 8 La permutation inverse de

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \tag{2.13}$$

est donnée par

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \tag{2.14}$$

On vérifie immédiatement que

$$\sigma \circ \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \sigma^{-1} \circ \sigma = \iota\,. \tag{2.15}$$

*
* *

Le groupe symétrique S_n est un groupe non-commutatif (sauf pour $n \leqslant 2$), ce qui rend son étude un peu difficile. Il est alors utile d'introduire des classes particulières de permutations dont les actions sont plus simples à décrire.

Définition 10 (transposition). *Une transposition* $\tau_{ij} \in S_n$*, avec* $i, j \in \{1, 2, \ldots, n\}$ *et* $i \neq j$*, est une permutation dont l'action est donnée par :*

$$\tau_{ij}(i) = j\;, \quad \tau_{ij}(j) = i\;, \quad \tau_{ij}(k) = k \quad \text{si} \quad k \neq i \text{ et } k \neq j\,. \tag{2.16}$$

On remarque que

$$\tau_{ij} \circ \tau_{ij} = \iota\,, \tag{2.17}$$

donc toute transposition est égale à son propre inverse. Cela signifie également que toute transposition génère un sous-groupe cyclique de S_n d'ordre 2.

Passons à une classe un peu plus large de permutations particulières, généralisant la notion de transposition.

Définition 11 (cycle). *Un cycle est une permutation* $c \in S_n$ *agissant comme une permutation circulaire sur* $p \geqslant 2$ *éléments* $\{k_1, k_2, \ldots, k_p\} \subset \{1, 2, \ldots, n\}$*, non nécessairement successifs :*

$$c = \begin{pmatrix} 1 & 2 & \cdots & k_1 & \cdots & k_2 & \cdots & k_p & \cdots & n \\ 1 & 2 & \cdots & k_2 & \cdots & k_3 & \cdots & k_1 & \cdots & n \end{pmatrix} \tag{2.18}$$

L'entier p *est appelé longueur du cycle ; une transposition est ainsi un cycle de longueur 2. On peut noter un cycle de manière condensée comme* $c = (a_1, \ldots, a_p)$.

Notons qu'un cycle c de longueur p est un élément du groupe d'ordre p, car il est évident que c^p est la permutation identité.

Exemple 9 On considère la permutation :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 2 & 3 & 5 & 9 & 6 & 7 & 8 & 1 \end{pmatrix} \tag{2.19}$$

qui correspond à la permutation circulaire $1 \to 4$, $4 \to 5$, $5 \to 9$ et $9 \to 1$. Cette permutation est donc le cycle $(1, 4, 5, 9)$.

*
* *

Définition 12 (support d'une permutation). *Soit* $\sigma \in S_n$ *une permutation. Le support de* σ*, noté* $S(\sigma)$*, est le sous-ensemble de* $\{1, \ldots, n\}$ *sur lequel la permutation agit effectivement, c.-à.-d.*

$$S(\sigma) = \big\{i \in \{1, \ldots, n\} \text{ t.q. } \quad \sigma(i) \neq i\big\}. \tag{2.20}$$

En particulier, si σ *correspond au cycle* $(a_1, \ldots, a_p)$*,* $S(\sigma) = \{a_1, \ldots, a_p\}$*. Le support d'une permutation est stable sous l'action de* σ*, c.-à.-d. que* $\sigma(S) = S$*.*

Soit $i \in S$ tel que $\sigma(i) \neq i$. Comme toute permutation est injective $\sigma(\sigma(i)) \neq \sigma(i)$ donc $\sigma(i) \in S$ et $\sigma(S) \subset S$. Finalement, comme σ est une bijection σ et $S(\sigma)$ ont le même cardinal donc $\sigma(S) = S$. □

Propriété 2 *Deux permutations* σ_1 *et* σ_2 *dans* S_n *de supports disjoints, c.-à.-d. telles que* $S(\sigma_1) \cap S(\sigma_2) = \emptyset$*, commutent entre elles, c.-à.-d. que* $\sigma \circ \sigma' = \sigma' \circ \sigma$*.*

Montrons que pour tout $i \in \{1, 2, \ldots, n\}$, $\sigma_1(\sigma_2(i)) = \sigma_2(\sigma_1(i))$. Il faut distinguer plusieurs cas de figure :

— si $i \notin S_1 \cup S_2$, alors par définition $\sigma_1(i) = i$ et $\sigma_2(i) = i$ donc $\sigma_1(\sigma_2(i)) = \sigma_2(\sigma_1(i)) = i$.

— si $i \in S_1$ alors $i \notin S_2$ donc $(\sigma_1 \circ \sigma_2)(i) = \sigma_1(i)$ d'une part et $(\sigma_2 \circ \sigma_1)(i) = \sigma_2[\sigma_1(i)] = \sigma_1(i)$ car $\sigma_1(i) \notin S_2$ par hypothèse.

— si $i \in S_2$ le raisonnement est similaire.

□

Donnons maintenant un thèorème fondamental concernant le groupe des permutations, dont nous ne donnerons pas la démonstration, assez technique mais sans être très difficile.

Théorème 1 (décomposition d'une permutation en cycles). *Toute permutation* $\sigma \in S_n$ *se décompose en un produit de cycles à supports disjoints ; cette décomposition est unique, à l'ordre des cycles près.*

Notons que l'ordre des cycles dans la décomposition n'a pas d'importance car des cycles à supports disjoints commutent.

Exemple 10 Soit la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 3 & 1 & 7 & 6 & 5 & 2 & 8 & 10 & 9 \end{pmatrix} \tag{2.21}$$

On obtient un premier cycle en partant de 1 qui donne $\sigma(1) = 4$ puis 4 qui donne $\sigma(4) = 7$, etc. On examine ainsi l'*orbite* de 1 sous l'action de σ. Le premier cycle de la décomposition est donc $c = (1, 4, 7, 2, 3)$. Le premier élément restant non encore parcouru est 5, et on a $\sigma(5) = 6$ puis $\sigma(6) = 5$ donc le second cycle est $c' = (5, 6)$. Finalement en partant de 8 on obtient $\sigma(8) = 8$ puis, en partant de 9, $\sigma(9) = 10$ et $\sigma(10) = 9$, dont le troisième cycle est $c'' = (9, 10)$ Cette analyse donne la décomposition

$$\sigma = c \circ c' \circ c'' = (1, 4, 7, 2, 3) \circ (5, 6) \circ (9, 10)\,, \tag{2.22}$$

unique à l'ordre des facteurs près (qui n'a pas d'importance).

La connaissance de la décomposition d'une permutation en cycles permet de simplifier grandement les calculs. Considérons par exemple l'action de la permutation σ^{77}. Étant donné que les cycles à supports disjoints commutent, nous avons $\sigma^{77} = (1, 4, 7, 2, 3)^{77} \circ (5, 6)^{77} \circ (9, 10)^{77}$. Étant donné que $77 \equiv 1 \mod 2$ et $77 \equiv 2 \mod 5$, nous en déduisons que $\sigma^{77} = (1, 4, 7, 2, 3)^2 \circ (5, 6) \circ (9, 10)$.

On peut encore simplifier les choses en décomposant $c^2 = (1, 4, 7, 2, 3)^2$ comme un produit de cycles disjoints. Regardons tout d'abord l'orbite de 1. Nous avons $c^2(1) = 7$ puis $c^2(7) = 3$, $c^2(3) = 4$, $c^2(4) = 2$ et enfin $c^2(2) = 1$ Nous avons donc $(1, 4, 7, 2, 3)^2 = (1, 7, 3, 4, 2)$ et

$$\sigma^{77} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 1 & 4 & 2 & 6 & 5 & 3 & 8 & 10 & 9 \end{pmatrix} \tag{2.23}$$

De cette décomposition en cycles nous pouvons déduire une décomposition en « briques » encore plus élémentaires, les transpositions.

Corollaire 1 (décomposition d'une permutation en transpositions). *Toute permutation se décompose comme un produit de transpositions.*

Il suffit de constater tout cycle peut de décomposer comme produit de transpositions :

$$c = (k_1, k_2, k_3, \ldots, k_p) = (k_1, k_2) \circ (k_2 k_3) \circ \cdots (k_{p-1} k_p)\,, \tag{2.24}$$

et d'en déduire, qu'une fois la décomposition d'une permutation en cycles achevée suivant le théorème 1, on peut la décomposer plus finement en produit de transpositions. □

Attention, nous avons donné *une* méthode permettant de décomposer toute permutation en produit de transpositions, mais cependant cette décomposition n'est absolument pas unique ! On peut par également écrire $c = (k_1, k_2, k_3, \ldots, k_p) = (k_1, k_p) \circ (k_1 k_{p-1}) \circ \cdots (k_1 k_2)$.

Exemple 11 Soit

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 8 & 6 & 5 & 9 & 1 & 7 & 4 & 3 \end{pmatrix} \tag{2.25}$$

Commençons par une décomposition de σ en cycles. L'orbite de 1 donne $1 \to 2 \to 8 \to 4 \to 5 \to 9 \to 3 \to 6 \to 1$ donc on a un cycle $c = (1, 2, 8, 4, 5, 9, 3, 6)$ et il ne reste que 7 qui est envoyé sur lui-même. D'après les deux décompositions d'un cycle en transpositions données au-dessus, nous pouvons donc écrire immédiatement deux décompositions de σ en transpositions :

$$\begin{aligned} \sigma &= (1,2) \circ (2,8) \circ (9,4) \circ (4,5) \circ (5,9) \circ (9,3) \circ (3,6) \\ &= (1,6) \circ (1,3) \circ (1,9) \circ (1,5) \circ (1,4) \circ (1,8) \circ (1,2)\,. \end{aligned} \tag{2.26}$$

D'autres sont certainement possibles. Nous avons la séquence

$$(3,9) \circ \sigma = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 8 & 6 & 5 & 3 & 1 & 7 & 4 & 9 \end{smallmatrix}\right)\,,\ \ (8,4) \circ (3,9) \circ \sigma = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 6 & 5 & 3 & 1 & 7 & 8 & 9 \end{smallmatrix}\right)\,,\ \cdots \tag{2.27}$$

qui permet de remettre tous les éléments « dans l'ordre » en partant du dernier, et aboutit à

$$(1,2) \circ (3,1) \circ (4,3) \circ (5,3) \circ (6,1) \circ (8,4) \circ (3,9) \circ \sigma = \iota\,. \tag{2.28}$$

En utilisant le fait que les transpositions sont égales à leur propre inverse et la propriété 1.5, on trouve

$$\sigma = (3,9) \circ (8,4) \circ (6,1) \circ (5,3) \circ (4,3) \circ (3,1) \circ (1,2)\,. \tag{2.29}$$

Nous avons ainsi obtenu simplement trois décompositions distinctes d'une même permutation en produits de transpositions.

*
* *

Définition 13 (signature d'une permutation). *Soit $\sigma \in S_n$ une permutation. On définit tout d'abord le nombre d'inversions $\mathcal{I}(\sigma)$ que contient la permutation, défini comme*

$$\mathcal{I}(\sigma) \overset{\text{déf.}}{=} \Big\{\{i,j\} \text{ t.q. } \quad i < j \text{ et } \quad \sigma(i) > \sigma(j)\Big\} \in \mathbb{N}\,, \tag{2.30}$$

c.-à.-d. le nombre de paires d'éléments dont l'ordre a été inversé par la permutation. On définit ensuite la signature comme

$$\epsilon(\sigma) \overset{\text{déf.}}{=} (-1)^{\mathcal{I}(\sigma)} \in \{-1, 1\}\,. \tag{2.31}$$

Exemple 12 Reprenons l'exemple étudié précédemment,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 8 & 6 & 5 & 9 & 1 & 7 & 4 & 3 \end{pmatrix}\,. \tag{2.32}$$

Les paires inversées par la permutation sont $\{1,6\}$, $\{2,3\}$, $\{2,4\}$, $\{2,6\}$, $\{2,7\}$, $\{2,8\}$, $\{2,9\}$, $\{3,4\}$, $\{3,6\}$, $\{3,8\}$, $\{3,9\}$, $\{4,6\}$, $\{4,8\}$, $\{4,9\}$, $\{5,6\}$, $\{5,7\}$, $\{5,8\}$, $\{5,9\}$, $\{7,8\}$, $\{7,9\}$ et $\{8,9\}$. Soit 21 inversions, et la signature de σ est $\epsilon(\sigma) = -1$. Ce calcul direct n'est certainement pas des plus pratiques !

Propriété 3 (signature d'une transposition). *Soit une transposition* $\tau = (ij) \in S_n$. *Sa signature est donnée par*

$$\epsilon(\tau) = -1\,, \tag{2.33}$$

indépendamment des éléments échangés.

On considère sans perte de généralité que $i < j$. La transposition τ correspond à une permutation de la forme

$$\tau = \left(\begin{smallmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & 2 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{smallmatrix}\right)\,. \tag{2.34}$$

Examinns le nombre d'inversions qu'elle contient ; il s'agit de toutes les paires de la forme $\{i, k\}$ avec $k = i+1, \ldots, j$ et celles de la forme $\{k, j\}$ avec $k = i+1, \cdots, j-1$. Le nombre total d'inversions est donc

$$\mathcal{I}(\tau) = (j-i) + (j-i-1) = 2(j-i) - 1 \in 2\mathbb{N}+1 \tag{2.35}$$

qui est toujours un nombre impair ; on en déduit que $\epsilon(\tau) = -1$. $\square$

Lemme 1 *Soit* $\sigma \in S_n$ *une permutation. Sa signature est donnée par*

$$\epsilon(\sigma) = \prod_{i \neq j} \frac{\sigma(j) - \sigma(i)}{j - i}\,, \tag{2.36}$$

où tout couple une (i, j) *n'intervient qu'une fois ; on peut par exemple restreindre la somme à* $i < j$.

Remarquons pour commencer que $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$, donc le choix de restreindre la somme à $i < j$ n'influence pas le résultat. Il suffit ensuite de remarquer que $\sigma(j) - \sigma(i) > 0$ si la paire $\{i, j\}$ ne subit pas d'inversion et négatif sinon. Le signe de cette expression est bien celui de ϵ_σ. Enfin concernant la valeur absolue de cette expression, on note que tant le numérateur que le dénominateur contiennent, exactement une fois, chaque facteur $|i - j|$ avec $i < j$, étant donné que σ est une bijection ; la valeur absolue est donc égale à 1 après simplification. $\square$.

Théorème 2 *Soient* σ *et* σ' *deux permutations quelconques* n *éléments. La signature de leur produit satisfait*

$$\epsilon(\sigma \circ \sigma') = \epsilon(\sigma) \circ \epsilon(\sigma')\,. \tag{2.37}$$

Autrement dit, la signature ϵ *est un morphisme entre le groupe* S_n *et le groupe* $(\{+1, -1\}, \times)$ *(c.-à.-d. le groupe multiplicatif contenant les deux éléments* 1 *et* -1*).*

Ce théorème se déduit directement du lemme 1. Partons de

$$\epsilon(\sigma) = \prod_{i \neq j} \frac{\sigma(j) - \sigma(i)}{j - i} \tag{2.38}$$

et, σ' étant une bijection, pour tout couple $\{i, j\}$ il existe un unique couple d'antécédents $\{k, \ell\}$, tels que $\sigma'(k) = i$ et $\sigma'(\ell) = j$. Nous avons donc

$$\epsilon(\sigma) = \prod_{k \neq \ell} \frac{\sigma(\sigma'(\ell)) - \sigma(\sigma'(k))}{\sigma'(\ell) - \sigma'(k)}. \tag{2.39}$$

On a donc

$$\epsilon(\sigma)\epsilon(\sigma') = \prod_{k \neq \ell} \frac{\sigma(\sigma'(\ell)) - \sigma(\sigma'(k))}{\sigma'(\ell) - \sigma'(k)} \prod_{k \neq \ell} \frac{\sigma'(\ell) - \sigma'(k)}{\ell - k} = \prod_{k \neq \ell} \frac{\sigma(\sigma'(\ell)) - \sigma(\sigma'(k))}{\ell - k} = \epsilon(\sigma \circ \sigma') \tag{2.40}$$

□

Corollaire 2 (propriétés de la signature d'une permutation) *Le théorème 2 permet d'obtenir des propriétés importantes de la signature d'une permutation :*

1. $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$ *;*
2. *si* σ *se décompose en produit de* r *transpositions, alors* $\epsilon(\sigma) = (-1)^r$ *;*
3. *si* c *est un cycle de longueur* p*, alors* $\epsilon(c) = (-1)^{p-1}$*.*

Les démonstrations de ces propriétés sont évidentes :

1. on utilise $\epsilon(\sigma)\epsilon(\sigma^{-1}) = \epsilon(\sigma \circ \epsilon^{-1}) = \epsilon(\iota) = 1$;
2. si $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_r$ on a $\epsilon(\sigma) = \epsilon(\tau_1)\epsilon(\tau_2) \cdots \epsilon(\tau_r) = (-1)^r$ d'après la propriété 3 ;
3. un cycle de longueur p se décompose par exemple comme $c = (k_1, k_2) \circ (k_2, k_3) \circ (k_{p-1}, k_p)$ soit comme produit de $(p-1)$ transpositions, d'où le résultat.

□

Nous pouvons retenir de ce qui précède que :

$\epsilon(\sigma) = +1$ *si la permutation* σ *se décompose en un nombre pair de transpositions, et* $\epsilon(\sigma) = -1$ *si* σ *se décompose en un nombre impair de transpositions.*

Remarquons cependant que la définition même de la signature ne fait pas appel à une décomposition en transpositions proprement dite et, cette décomposition n'étant pas unique (voir l'exemple 11), nous remarquons que *la signature d'une permutation* σ *est indépendante de la décomposition en transpositions choisie.*

La compréhension acquise du groupe symétrique et de ses propriété nous permettra d'étudier en détail le déterminant à la section suivante, et d'en déduire nombre de ses propriétés.

2.3 Déterminant d'une matrice

Le déterminant peut être considéré de différentes points de vue soit comme le déterminant d'une famille de vecteurs, soit comme le déterminant d'une matrice, soit comme le déterminant d'un endomorphisme ; nous commencerons par le premier.

2.3.1 Déterminant d'une famille de vecteurs

Définition 14 (déterminant d'une famille de vecteurs dans une base). *Soit* E *un espace vectoriel de dimension* n *muni d'une base* $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$*. Le déterminant* $\det_{\mathcal{B}}$ *associé à cette base est l'unique application*

$$\det_{\mathcal{B}} : \begin{cases} E^n & \to \quad K \\ (\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n) & \mapsto \quad \det_{\mathcal{B}}(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n) \end{cases} \tag{2.41}$$

multilinéaire, c.-à.-d. que pour tout $k \in \{1, \ldots, n\}$

$$\det_{\mathcal{B}}(\vec{v}_1, \ldots, \lambda\vec{v}^k + \mu\vec{w}^k, \ldots, \vec{v}_n) = \lambda \det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{v}^k, \ldots, \vec{v}_n) + \mu \det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{w}^k, \ldots, \vec{v}_n) \tag{2.42}$$

totalement antisymétrique,

$$\forall \sigma \in S_n \ , \quad \det_{\mathcal{B}}(\vec{v}_{\sigma(1)}, \vec{v}_{\sigma(2)}, \ldots, \vec{v}_{\sigma(n)}) = \epsilon(\sigma) \det_{\mathcal{B}}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) \tag{2.43}$$

et telle que $\det(\vec{e}_1, \ldots, \vec{e}_n) = 1$*.*

L'unicité du déterminant peut se comprendre en faisant appel au cours du premier semestre. Le déterminant est en effet rien d'autre qu'une n-forme, et nous avions vu que l'espace vectoriel des n-formes est de dimension 1. La condition sur le déterminant de la base fixe alors l'indétermination restante.

Propriété 4 (déterminant d'une famille liée). *Soit* E *un espace vectoriel de dimension* n *muni d'une base* $\mathcal{B}$*, et* $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ *une famille de vecteurs non nuls. Si cette famille est liée alors le déterminant* $\det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{v}_n)$ *est nul. En particulier,*

$$\det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{v}_k, \ldots, \vec{v}_k, \ldots, \vec{v}_n) = 0\,. \tag{2.44}$$

Cette propriété est une conséquence directe de l'antisymétrie. En effet si la famille est liée, il existe $r \in \{0, \ldots, n\}$ tel que $\vec{v}_r = \mu_1\vec{v}_1 + \cdots + \mu_n\vec{v}_n$, où la somme de droite n'inclut pas $\vec{v}_r$ et où les coefficients $\mu_1, \ldots, \mu_n$ ne sont pas tous nuls. Par multilinéarité on peut développer ce déterminant, et pour chacun des termes obtenus

$$\mu_\ell \det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{v}_r, \ldots, \vec{v}_r, \ldots, \vec{v}_n) = -\mu_\ell \det_{\mathcal{B}}(\vec{v}_1, \ldots, \vec{v}_r, \ldots, \vec{v}_r, \ldots, \vec{v}_n) = 0\,, \tag{2.45}$$

par antisymétrie du déterminant sous l'échange des deux arguments égaux. □

Propriété 5 (déterminant d'une famille de vecteurs en composantes). *Soit* E *un espace vectoriel de dimension* n *muni d'une base* $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$*, et* $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ *une famille de* n *vecteurs, qui se décomposent dans la base choisie comme :*

$$\vec{v}_k = \sum_{\ell=1}^{n} v_k^\ell \vec{e}_\ell\,. \tag{2.46}$$

Le déterminant de la famille de vecteurs est alors donné par :

$$\boxed{\det_{\mathcal{B}}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) v_1^{\sigma(1)} v_2^{\sigma(2)} \cdots v_n^{\sigma(n)}} \tag{2.47}$$

En utlisant la multilinéarité du déterminant,

$$\begin{aligned}\det_{\mathcal{B}}(\vec{v}_1,\ldots,\vec{v}_n) &= \det_{\mathcal{B}}\left(\sum_{\ell_1=1}^{n} v_1^{\ell_1}\vec{e}_{\ell_1},\ldots,\sum_{\ell_n=1}^{n} v_n^{\ell_n}\vec{e}_{\ell_n}\right)\\ &= \sum_{\ell_1,\ell_2,\ldots,\ell_n=1}^{n} v_1^{\ell_1}v_2^{\ell_2}\cdots v_n^{\ell_n}\det_{\mathcal{B}}\left(\vec{e}_{\ell_1},\vec{e}_{\ell_2},\ldots,\vec{e}_{\ell_n}\right)\,. \end{aligned} \tag{2.48}$$

D'après la propriété 4, les seuls termes non-nuls dans la somme multiple sont les termes pour lesquels les n vecteurs $\{\vec{e}_{\ell_1},\vec{e}_{\ell_n},\ldots,\vec{e}_{\ell_n}\}$ sont tous distincts. Par définition, étant donné que S_n contient toutes les bijections de S_n dans lui-même, il existe une unique permutation σ telle que

$$\sigma(1,2,\ldots,n) = (\ell_1,\ell_2,\ldots,\ell_n)\,. \tag{2.49}$$

Sommer sur tous les termes non-nuls de la somme multiple revient alors à sommer sur toutes les permutations de n éléments. Nous avons donc :

$$\det_{\mathcal{B}}(\vec{v}_1,\ldots,\vec{v}_n) = \sum_{\sigma\in S_n} v_1^{\sigma(1)}v_2^{\sigma(2)}\cdots v_n^{\sigma(n)}\det_{\mathcal{B}}\left(\vec{e}_{\sigma(1)},\vec{e}_{\sigma(2)},\ldots,\vec{e}_{\sigma(n)}\right) \tag{2.50}$$

puis, en utilisant l'antisymétrie (cf. eqn. (2.42)), on trouve :

$$\det_{\mathcal{B}}(\vec{v}_1,\ldots,\vec{v}_n) = \sum_{\sigma\in S_n} \epsilon(\sigma) v_1^{\sigma(1)}v_2^{\sigma(2)}\cdots v_n^{\sigma(n)}\underbrace{\det_{\mathcal{B}}\left(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right)}_{=1} \tag{2.51}$$

d'où le résultat. □

Remarque 3 (symbole de Levi–Civita). *Le déterminant peut également s'écrire en utilisant le symbole de Levi–Civita étudié au premier semestre :*

$$\begin{aligned}&\epsilon_{i_1\cdots i_n} = +1 \text{ si } (i_1 i_2\cdots i_n) \text{ est une permutation paire de } (12\cdots n)\\ &\epsilon_{i_1\cdots i_n} = -1 \text{ si } (i_1 i_2\cdots i_n) \text{ est une permutation impaire de } (12\cdots n)\\ &\epsilon_{i_1\cdots i_n} = 0 \text{ si un de ses indices apparaît au moins deux fois.}\end{aligned} \tag{2.52}$$

Nous avons alors

$$\det_{\mathcal{B}}(\vec{v}_1,\ldots,\vec{v}_n) = \sum_{i_1,\ldots,i_n} \epsilon_{i_1\cdots i_n} v_1^{i_1}\cdots v_n^{i_n}\,. \tag{2.53}$$

*
* *

On considère une famille de vecteurs $(\vec{v}_1,\ldots,\vec{v}_p)$, avec $p \leqslant n$, et on cherche à determiner s'ils sont linéairement indépendants. Lorsque $p = n$, il suffit de vérifier que $\det(\vec{v}_1,\ldots,\vec{v}_p) \neq 0$ comme nous venons de le montrer. Lorsque $p < n$ le critère prend une forme un peu différente.

Propriété 6 (critère d'indépendance linéaire). *Soit* E *un espace vectoriel de dimension* n, *et* $\{\vec{v}_1, \ldots, \vec{v}_p\}$ *une famille de vecteurs avec* $p < n$. *Cette famille est libre (c.-à.-d. que les vecteurs sont linéairement indépendants) si et seulement si l'un (au moins) des déterminants* $p \times p$ *formés avec une sélection* p *parmi leur* n *composantes (identique pour tous le vecteurs), dans une base quelconque, est non nul.*

Soit $\mathcal{B} = \{\vec{e}_i, i = 1, \ldots, n\}$ une base de E. On suppose sans perte de généralité que le déterminant formé par les p premières composantes de tous les vecteurs est non nul. Ces composantes ${v_i}^j$, pour $i = 1, \ldots, p$ (numéro du vecteur) et $j = i, \ldots, p$ (composante du vecteur) peuvent être comprises comme les composantes de p vecteurs « tronqués » $\vec{w}_i$ dans le sous-espace vectoriel à p dimensions $\mathsf{F} =$ Vect $(\vec{e}_1, \ldots, \vec{e}_p)$, avec $\vec{w}_i = \sum_{j=1}^{p} {v_i}^j \vec{e}_j$. Comme, par hypothèse, $\det(\vec{w}_1, \ldots, \vec{w}_p) \neq 0$, ce sont des vecteurs linéairement indépendants de F, c.-à.-d. que

$$\sum_{i=1}^{p} \lambda_i \vec{w}_i = \vec{0}_\mathsf{F} \implies \lambda_1 = \cdots = \lambda_p = 0\,. \tag{2.54}$$

Si on considère maintenant les vecteurs « complets » $\vec{v}_i$, pour tester l'indépendance linéaire on considère le système

$$\sum_{i=1}^{p} \lambda_i \vec{v}_i = \vec{0}_\mathsf{E}\,, \tag{2.55}$$

qui contient les mêmes équations que le système (2.54), plus $(n-p)$ équations homogènes supplémentaires qui sont compatibles avec la solution $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$ imposée par les p premières équations. Donc que la famille de vecteurs $\{\vec{v}_1, \ldots, \vec{v}_p\}$ est libre.

Pour la réciproque nous devons anticiper sur la suite du cours et utiliser les résultats de la sous-section 2.3.2. Supposons que la famille $(\vec{v}_1, \ldots, \vec{v}_p)$ est libre. Elle génère alors un sous-espace vectoriel Vect$\{\vec{v}_1, \ldots, \vec{v}_p\} \subset \mathsf{E}$ de dimension p. On peut alors compléter ces vecteurs par $(n-p)$ vecteurs de base $\{\vec{e}_{k_{p+1}}, \ldots, \vec{e}_{k_n}\}$ afin de former une nouvelle base de E, $\hat{\mathcal{B}} = \{\vec{v}_1, \ldots, \vec{v}_p, \vec{e}_{k_{p+1}}, \ldots, \vec{e}_{k_n}\}$. Renumérotons les vecteurs $\{\vec{e}_1, \ldots, \vec{e}_n\}$ de la base $\mathcal{B}$ suivant :

$$\{1, 2, \ldots, n\} = \{k_1, \ldots, k_p\} \cup \{k_{p+1}, \ldots k_n\} \tag{2.56}$$

et définissions une base $\tilde{\mathcal{B}} = \{\vec{f}_i, i = 1, \ldots, n\}$ à partir de la base $\mathcal{B}$ par un ce réarrangement de l'ordre des vecteurs :

$$\forall i \in \{1, \ldots, n\}\,, \quad \vec{f}_i = \vec{e}_{k_i}\,. \tag{2.57}$$

Écrivons alors la matrice P comprenant les composantes des vecteurs de la base $\hat{\mathcal{B}}$ dans cette base $\tilde{\mathcal{B}}$ réorganisée, c.-à.-d. la matrice de passage de $\tilde{\mathcal{B}}$ vers $\hat{\mathcal{B}}$:

$$P = \begin{pmatrix} p^{k_1}_1 & \cdots & \cdots & p^{k_1}_p & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ p^{k_p}_1 & \cdots & \cdots & p^{k_p}_p & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & & & \vdots & \ddots & \ddots & 0 \\ p^{k_n}_1 & \cdots & \cdots & p^{k_n}_p & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{2.58}$$

Par construction son déterminant est non-nul car il s'agit d'une matrice de passage. Cette matrice étant triangulaire par blocs, en utilisant la propriété 8, on trouve

$$\det P = \begin{vmatrix} p^{k_1}_1 & \cdots & \cdots & p^{k_1}_p \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ p^{k_p}_1 & \cdots & \cdots & p^{k_p}_p \end{vmatrix} \times \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{vmatrix} = \begin{vmatrix} p^{k_1}_1 & \cdots & \cdots & p^{k_1}_p \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ p^{k_p}_1 & \cdots & \cdots & p^{k_p}_p \end{vmatrix} \times 1 \neq 0 \tag{2.59}$$

par hypothèse. □

2.3.2 Déterminant d'une matrice carrée

Comme nous l'avions vu au cours 1, une matrice A de taille $n \times p$ peut être considérée comme une juxtaposition de p vecteurs colonnes $\mathbf{A}_i$, chacun contenant les composantes d'un vecteur $\vec{a}_i$ dans une certaine base $\mathcal{B}$ d'un espace vectoriel de dimension n. Pour une matrice de taille $n \times n$, nous obtenons une famille de n vecteurs colonnes dont nous pouvons calculer le déterminant.

Remarque 4 *Comme un vecteur colonne est défini par rapport à une base donnée, nous écrirons le déterminant sans l'indice $\mathcal{B}$ faisant référence à la base choisie.*

Définition 15 (déterminant d'une matrice carrée). *Soit $A \in \mathcal{M}_K(n,n)$ une matrice carrée. Son déterminant est égal à celui de la famille de ses vecteurs colonnes,*

$$\det A \overset{\text{déf.}}{=} \det\left(\mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_n}\right) \tag{2.60}$$

et s'exprime donc en termes des composantes $(A)^i{}_j = a^i{}_j$, d'après la propriété 5, comme

$$\boxed{\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a^{\sigma(1)}_1 a^{\sigma(2)}_2 \cdots a^{\sigma(n)}_n} \tag{2.61}$$

Le déterminant de A s'écrit aussi comme

$$\det A = |\mathbf{A}_1\, \mathbf{A}_2 \cdots \mathbf{A}_n| = \begin{vmatrix} a^1{}_1 & a^1{}_2 & \cdots & a^1{}_n \\ \vdots & \vdots & & \vdots \\ a^n{}_1 & a^n{}_2 & \cdots & a^n{}_n \end{vmatrix} . \tag{2.62}$$

Exemple 13 (déterminants en petite dimension). Considérons les déterminants associés aux matrices en dimension 1,2 et 3 :

1. en dimension 1, une matrice A se réduit à un coefficient $a \in K$ et $\det A = a$;
2. en dimension 2, $S_2 = \{\iota, (1,2)\}$ et nous avons

$$\det A = \begin{vmatrix} a^1{}_1 & a^1{}_2 \\ a^2{}_1 & a^2{}_2 \end{vmatrix} = \epsilon\,(\iota) a^1{}_1 a^2{}_2 + \epsilon\,\big((1,2)\big) a^2{}_1 a^1{}_2 = a^1{}_1 a^2{}_2 - a^2{}_1 a^1{}_2\,; \tag{2.63}$$

3. en dimension 3, $S_3 = \{\iota, (1,2,3), (1,3,2), (1,2), (1,3), (2,3)\}$ et on a donc

$$\begin{aligned}\det A &= \begin{vmatrix} a^1{}_1 & a^1{}_2 & a^1{}_3 \\ a^2{}_1 & a^2{}_2 & a^2{}_3 \\ a^3{}_1 & a^3{}_2 & a^3{}_3 \end{vmatrix} \\ &= a^1{}_1 a^2{}_2 a^3{}_3 + a^1{}_2 a^2{}_3 a^3{}_1 + a^1{}_3 a^2{}_1 a^3{}_2 - a^1{}_2 a^2{}_1 a^3{}_3 - a^1{}_3 a^2{}_2 a^3{}_1 - a^1{}_1 a^2{}_3 a^3{}_2 \end{aligned} \tag{2.64}$$

   où nous avons fait figurer tout d'abord les deux permutations circulaires vers la droite et la gauche (permutations paires) puis les 3 transpositions (permutations impaires). Il existe différentes méthodes graphiques pour retrouver ce résultat.

Propriété 7 *Le déterminant de matrices carrées satisfait les propriétés suivantes :*

1. *Pour toute matrice* $A \in\in \mathcal{M}_K(n,n)$ *et pour tout* $\lambda \in K$*,* $\det(\lambda A) = \lambda^n \det A$ *;*
2. *Le déterminant d'une matrice est égal à celui de sa transposée,* $\det A^T = \det A$ *;*
3. *Pour toutes matrices* $A, B \in\in \mathcal{M}_K(n,n)$*,*

$$\boxed{\det(A \cdot B) = \det A\ \det B} \tag{2.65}$$

Les preuves de ces propriétés sont les suivantes :

1. Nous avons par multilinéarité

$$\det(\lambda A) = \big|\lambda \mathbf{A_1}, \lambda \mathbf{A_2}, \ldots, \lambda \mathbf{A_n}\big| = \lambda \big|\mathbf{A_1}, \lambda \mathbf{A_2}, \ldots, \lambda \mathbf{A_n}\big| = \cdots = \lambda^{\mathbf{n}} \big|\mathbf{A_1}\ \mathbf{A_2} \cdots \mathbf{A_n}\big| \tag{2.66}$$

2. On écrit, par définition de la transposée

$$\det A^T = \sum_{\sigma \in S_n} \epsilon(\sigma) (A^T)^{\sigma(1)}{}_1 \cdots (A^T)^{\sigma(n)}{}_n = \sum_{\sigma \in S_n} \epsilon(\sigma) a^1{}_{\sigma(1)} \cdots a^k{}_{\sigma(k)} \cdots a^n{}_{\sigma(n)}\,. \tag{2.67}$$

   Comme toute permutation est une bijection, l'ensemble $\{\sigma(1), \ldots, \sigma(n)\}$ est identique à $\{1, 2, \ldots, n\}$. On peut donc poser $\ell = \sigma(k)$, ou de manière équivalente $k = \sigma^{-1}(\ell)$, et écrire le produit de composantes de la matrice, en réordonnant les termes, comme

$$a^1{}_{\sigma(1)} \cdots a^k{}_{\sigma(k)} \cdots a^n{}_{\sigma(n)} = a^{\sigma^{-1}(1)}{}_1 \cdots a^{\sigma^{-1}(\ell)}{}_\ell \cdots a^{\sigma^{-1}(n)}{}_n\,.$$

   Nous avons donc

$$\det A^T = \sum_{\sigma \in S_n} \epsilon(\sigma) a^{\sigma^{-1}(1)}{}_1 \cdots a^{\sigma^{-1}(\ell)}{}_\ell \cdots a^{\sigma^{-1}(n)}{}_n \tag{2.68}$$

On peut finalement troquer la somme sur les $\sigma \in S_n$ par la somme sur les $\rho = \sigma^{-1} \in S_n$, qui décrit également une fois le groupe, et on trouve

$$\det A^{\mathrm{T}} = \sum_{\rho \in S_n} \epsilon(\rho^{-1}) a^{\rho(1)}{}_1 \cdots a^{\rho(n)}{}_n = \sum_{\rho \in S_n} \epsilon(\rho) a^{\rho(1)}{}_1 \cdots a^{\rho(n)}{}_n = \det A\,, \tag{2.69}$$

où nous avons utilisé la propriété $\epsilon(\rho^{-1}) = \epsilon(\rho)$.

3. Les colonnes de la matrice $C = A \cdot B$ ont pour composantes

$$(\mathbf{C}_j)^i = \sum_{k=1}^{n} a^i{}_k b^k{}_j = \sum_{k=1}^{n} b^k{}_j (\mathbf{A}_k)^i\,, \tag{2.70}$$

donc les vecteurs colonnes de C sont des combinaisons linéaires des vecteurs colonnes de A, $\mathbf{C}_j = \sum_{k=1}^{n} b^k{}_j \mathbf{A}_k$. Nous avons donc, en effectuant un tel développement pour chaque colonne et en utilisant la multilinéarité du déterminant,

$$\det C = |\mathbf{C}_1 \cdots \mathbf{C}_n| = \sum_{k_1, k_2, \ldots, k_n = 1}^{n} b^{k_1}{}_1 \cdots b^{k_n}{}_n |\mathbf{A}_{k_1} \cdots \mathbf{A}_{k_n}| \tag{2.71}$$

Dans chaque terme $|\mathbf{A}_{k_1} \cdots \mathbf{A}_{k_n}|$, l'antisymétrie du déterminant indique que seuls les termes pour lesquels $(k_1, k_2, \ldots, k_n)$ est une permutation de $(1, 2, \ldots, n)$ contribuent. On peut donc comme précédemment remplacer la somme multiple par une somme sur toutes les permutations de n éléments et on obtient :

$$\det C = \sum_{\sigma \in S_n} b^{\sigma(1)}{}_1 \cdots b^{\sigma(n)}{}_n |\mathbf{A}_{(\sigma(1)} \cdots \mathbf{A}_{\sigma(n)}| = \left(\sum_{\sigma \in S_n} b^{\sigma(1)}{}_1 \cdots b^{\sigma(n)}{}_n \epsilon(\sigma) \right) |\mathbf{A}_1 \cdots \mathbf{A}_n| \tag{2.72}$$

où nous avons utilisé l'antisymétrie de A dans la dernière étape, donnant le résultat attendu. □

Remarque 5 (déterminant d'une famille de matrices ligne) *Étant donné que le déterminant d'une matrice* A *est égal au déterminant de sa matrice transposée* A^{T}*, on peut aussi définir le déterminant par rapport à da décomposition en matrices ligne* $\mathbf{A}^i$*, c.-à.-d. comme*

$$\det A = \begin{vmatrix} \mathbf{A}^1 \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^n \end{vmatrix}. \tag{2.73}$$

Remarque 6 *Attention, en général* $\det(A + B) \neq \det A + \det B$. *Il ne faut pas confondre les propriétés du déterminant avec celles de la trace.*

Propriété 8 déterminant d'une matrice triangulaire). *Soit* $M \in \mathcal{M}_{K}(n, n)$ *une matrice triangulaire supérieure par blocs, c.-à.-d. de la forme*

$$M = \begin{pmatrix} A & C \\ \mathbb{O} & B \end{pmatrix}, \tag{2.74}$$

où A *est une matrice carrée* $p \times p$*,* B *une matrice carrée* $q \times q$*,* C *une matrice* $p \times q$ *et* $\mathbb{O}$ *la matrice nulle* $q \times p$*, avec* $p + q = n$*. Le déterminant de* M *satisfait :*

$$\det M = \det A \det B\,. \tag{2.75}$$

Reprenons la définition du déterminant d'une matrice,

$$\det M = \sum_{\sigma \in S_n} \epsilon(\sigma) m^{\sigma(1)}_{1} m^{\sigma(2)}_{2} \cdots m^{\sigma(n)}_{n}\,. \tag{2.76}$$

Comme M est une matrice de la forme (2.74), il est clair qu'un terme de la somme ne peut être éventuellement non nul que si

$$\sigma(1), \sigma(2), \ldots \sigma(p) \in \{0, 1, \ldots, p\}, \tag{2.77}$$

pour ne pas qu'y contribuent des éléments de la matrice nulle $\mathbb{O}$. Cela implique évidemment que, pour les entiers restants,

$$\sigma(p+1), \sigma(p+2), \ldots \sigma(n) \in \{p+1, p+2, \ldots, n\}, \tag{2.78}$$

Il existe donc des permutations $\pi \in S_p$ et $\rho \in S_q$ telles que :

$$\forall j \in \{1, \ldots, p\}\,, \quad \pi(j) = \sigma(j) \tag{2.79a}$$

$$\forall k \in \{1, \ldots, q\}\,, \quad p + \rho(k) = \sigma(p+k) \tag{2.79b}$$

$$\tag{2.79c}$$

On peut alors écrire le déterminant de M, en ne gardant que les termes éventuellement non nuls, comme

$$\det M = \sum_{\pi \in S_p, \rho \in S_q} \epsilon(\sigma) a^{\pi(1)}_{1} \cdots a^{\pi(p)}_{p}\, b^{\rho(1)}_{1} \cdots b^{\rho(q)}_{q}\,, \tag{2.80}$$

en termes des composantes de A et B. Les relations (2.79) signifiant que les permutations subistant dans la somme correspondent à la composition de deux permutations à supports disjoints, π et ρ, nous avons $\epsilon(\sigma) = \epsilon(\pi)\epsilon(\rho)$, d'où

$$\begin{aligned} \det M &= \sum_{\pi \in S_p \epsilon} a^{\pi(1)}_{1} \cdots a^{\pi(p)}_{p} \sum_{\rho \in S_q} b^{\rho(1)}_{1} \cdots b^{\rho(q)}_{q} \\ &= \det A \det B\,. \end{aligned} \tag{2.81}$$

□

Notons qu'on obtiendrait la même propriété avec une matrice triangulaire inférieure par blocs, c.-à.-d. de la forme

$$M = \begin{pmatrix} A & \mathbb{O} \\ C & B \end{pmatrix} , \tag{2.82}$$

en utilisant le fait que $\det M = \det M^{\mathrm{T}}$, nous ramenant au cas précédent.

Par récurrence sur le nombre de blocs, on peut aussi facilement généraliser la propriété démontrée à

$$\det \begin{pmatrix} A_1 & \star & \star & \star \\ 0 & A_2 & \star & \star \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & A_s \end{pmatrix} = \det A_1 \det A_2 \cdots \det A_s \,. \tag{2.83}$$

Corollaire 3 *Soit $A \in \mathcal{M}_K(n, n)$ une matrice triangulaire supérieure, c.-à.-d. de la forme*

$$A = \begin{pmatrix} a^1{}_1 & a^1{}_2 & \cdots & a^1{}_n \\ 0 & a^2{}_2 & \cdots & a^2{}_n \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a^n{}_n \end{pmatrix} \tag{2.84}$$

Le déterminant de A est alors donné par le produit de ses éléments diagonaux, soit

$$\det A = a^1{}_1 a^2{}_2 \cdots a^n{}_n \,. \tag{2.85}$$

Ce résultat est une conséquence directe de l'équation (2.83).

Remarquons que dans le cas d'une matrice diagonale, c.-à.-d. une matrice carrée de la forme

$$A = \begin{pmatrix} a^1{}_1 & 0 & \cdots & 0 \\ 0 & a^2{}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a^n{}_n \end{pmatrix} \tag{2.86}$$

on trouve évidemment

$$\det A = a^1{}_1 a^2{}_2 \cdots a^n{}_n \,, \tag{2.87}$$

s'agissant d'un cas particulier du précédent. On pourrait dans ce dernier cas en faire une demonstration beaucoup plus directe que nous laissons en exercice.

2.4 Développement d'un déterminant

Une méthode extrêmement puissante pour évaluer un déterminant consiste, au lieu de calculer par force brute la somme sur toutes les permutations de S_n, à effectuer un développement en lignes ou colonnes pour se ramener au calcul de déterminants plus petits, puis par itération à éviter tout calcul direct de déterminant.

Définition 16 (mineur d'une matrice). *Soit $A \in \mathcal{M}_K(n,n)$ une matrice $n \times n$. On note $\hat{A}^i{}_j$ la sous-matrice $(n-1) \times (n-1)$ obtenue en supprimant de A la i-ème ligne et la j-ième colonne. Le déterminant de la sous-matrice $\hat{A}^i{}_j$ est appelé mineur d'ordre $(n-1)$ de la matrice A, pour le couple (i,j).*

On introduit également le cofacteur $\Delta^i{}_j$, égal au déterminant du mineur $\hat{A}^i{}_j$ à un signe près :

$$\boxed{\Delta^i{}_j \overset{\text{déf.}}{=} (-1)^{i+j} \det \hat{A}^i{}_j} \tag{2.88}$$

Propriété 9 (développement selon une ligne ou selon une colonne)
Soit $A \in \mathcal{M}_K(n,n)$ une matrice $n \times n$. Le déterminant de A peut se développer selon la i-ème ligne comme :

$$\det A = \sum_{j=1}^{n} a^i{}_j \Delta^i{}_j \ , \quad i \textit{ fixé} \tag{2.89}$$

ou bien selon la j-ème colonne comme :

$$\det A = \sum_{i=1}^{n} a^i{}_j \Delta^i{}_j \ , \quad j \textit{ fixé} \tag{2.90}$$

Pour démontrer la formule (2.89), développons le vecteur colonne $\mathbf{A}_j$ comme

$$\mathbf{A}_j = \begin{pmatrix} a^1{}_j \\ a^2{}_j \\ a^3{}_j \\ \vdots \\ a^n{}_j \end{pmatrix} = \begin{pmatrix} a^1{}_j \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a^2{}_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ a^n{}_j \end{pmatrix} . \tag{2.91}$$

Par linéarité du déterminant par rapport à la j-ième colonne, on peut développer $\det A$ comme une somme de determinants de matrices où, dans la j-ème colonne, apparaît un des termes du développement (2.91).

Plus explicitement N^i_j la matrice obtenue à partir de la matrice A en remplaçant a j-ième colonne par le i-ème terme du développement (2.91), c.-à.-d.

$$N^i{}_j = \begin{pmatrix} a^1{}_1 & \cdots & a^1{}_{j-1} & 0 & a^1{}_{j+1} & \cdots & a^1{}_n \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & a^i{}_j & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \leftarrow i\text{-ème ligne} \\ \\ \\ \\ \end{matrix} \tag{2.92}$$

Nous avons alors, par linéarité du déterminant de n vecteurs colonne par rapport au j-ième vecteur,

$$\det A = \det N^1{}_j + \det N^2{}_j + \cdots + \det N^n{}_j \,. \tag{2.93}$$

Il reste à calculer le déterminant d'un des termes de la somme, c.-à.-d. le déterminant d'une matrice $N^i{}_j$. Si nous voyons $\det N^i{}_j$ comme le déterminant d'une famille de vecteurs colonne, il est possible par une série de $j-1$ transpositions, déplacer la j-ème colonne vers la gauche pour la mettre en première position :

$$\det N^i{}_j = (-1)^{j-1} \begin{vmatrix} 0 & a^1{}_1 & \cdots & a^1{}_{j-1} & a^1{}_{j+1} & \cdots & a^1{}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a^i{}_j & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \tag{2.94}$$

où le signe $(-)^{j-1}$ est la signature de la permutation effectuée.

Il est ensuite possible, en permutant les vecteurs lignes de la matrice, d'amener l'élément $a^i{}_j$ en haut à gauche, par une succession de $(i-1)$ transpositions, afin d'obtenir :

$$\det N^i{}_j = (-1)^{j-1}(-1)^{i-1} \begin{vmatrix} a^i{}_j & a^i{}_1 & \cdots & a^i{}_{j-1} & a^i{}_{j+1} & \cdots & a^i{}_n \\ 0 & & & & & & \\ \vdots & & & \hat{A}^i{}_j & & & \\ 0 & & & & & & \end{vmatrix} \tag{2.95}$$

faisant apparaître la matrice $\hat{A}^i{}_j$ obtenue, rappelons-le, en ôtant de la matrice A la i-ème ligne et la j-ème colonne. Étant donné que les éléments de la première colonne, sauf le premier, sont nuls, on peut utiliser la propriété 8 concernant le déterminant de matrices triangulaires par blocs et on obtient

$$\det N^i{}_j = (-1)^{j-1}(-1)^{i-1} a^i{}_j \det \hat{A}^i{}_j = a^i{}_j \, \Delta^i{}_j \,, \tag{2.96}$$

puis, en considérant tous les termes dans la somme (2.93), le résultat attendu. La formule (2.90) concernant le développement en lignes se démontre de la même manière, soit en prenant la transposée, soit en développant le i-ème vecteur ligne d'une manière analogue à (2.91). $\square$

Le développement suivant une ligne ou suivant une colonne permet de simplifier grandement le calcul d'un déterminant de grande taille, car il permet de passer du déterminant d'une matrice $n \times n$ (donc avec $n!$ termes dans la somme sur les permutations) à une somme de n déterminants de matrices de taille $(n-1) \times (n-1)$, qui peuvent eux-mêmes ce calculer de cette manière en itérant le processus.

Exemple 14 Soit le déterminant

$$\det M = \begin{vmatrix} 1 & 2 & 3 & 1 \\ -1 & 0 & 1 & 2 \\ 2 & -1 & 1 & 1 \\ 1 & 0 & -1 & 1 \end{vmatrix} \tag{2.97}$$

Il est évidemment judicieux de développer suivant la seconde colonne qui comporte deux zéros. On trouve

$$\begin{aligned} \det M &= 2 \times \Delta^1{}_2 + 0 \times \Delta^2{}_2 + (-1) \times \Delta^3{}_2 + 0 \times \Delta^4{}_2 \\ &= 2 \times (-1)^{2+1} \begin{vmatrix} -1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix} + (-1) \times (-1)^{3+2} \begin{vmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 1 & -1 & 1 \end{vmatrix} \\ &= -2(-1 + 1 - 4 - 2 - 1 - 2) + (1 + 6 + 1 - 1 + 2 + 3) = 18 + 12 \\ &= 30 \end{aligned} \tag{2.98}$$

*
* *

Comme dans cet exemple il est utile, le cas échéant, de choisir une ligne ou une colonne comportant des zéros pour diminuer le nombre de termes à calculer. Lorsque la matrice ne comporte pas ou pas suffisamment de zéros pour procéder de la sorte, il est possible d'en faire apparaître d'avantage en combinant les lignes ou les colonnes dans le déterminant.

Propriété 10 (combinaison de colonnes ou de lignes).
Soit $A \in \mathcal{M}_K(n, n)$ une matrice, qui se décompose en termes de ses vecteurs colonnes comme $A = (\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$. La matrice obtenue en ajoutant à un vecteur colonne $\mathbf{A}_j$ une combinaison linéaire arbitraire des autres colonnes, ce que l'on notera

$$\mathbf{A}_j \leftarrow \mathbf{A}_j + \sum_{k \neq j} \lambda_k \mathbf{A}_k \,, \tag{2.99}$$

possède le même déterminant que A, c.-à.-d. que

$$\det A = |\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_j, \cdots \mathbf{A}_n| = \left| \mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_j + \sum_{k \neq j} \lambda_k \mathbf{A}_k, \cdots \mathbf{A}_n \right| . \tag{2.100}$$

En raisonnant avec la transposée de la matrice, on arrive à la même conclusion si on ajoute a à un vecteur ligne $\mathbf{A}^i$ une combinaison linéaire arbitraire des autres colonnes, ce que l'on notera $\mathbf{A}^i \leftarrow \mathbf{A}^i + \sum_{\ell \neq j} \lambda_\ell \mathbf{A}^\ell$:

$$\det A = \begin{vmatrix} \mathbf{A}^1 \\ \vdots \\ \mathbf{A}^i \\ \vdots \\ \mathbf{A}^n \end{vmatrix} = \begin{vmatrix} \mathbf{A}^1 \\ \vdots \\ \mathbf{A}^i + \sum_{\ell \neq j} \lambda_\ell \mathbf{A}^\ell \\ \vdots \\ \mathbf{A}^n \end{vmatrix} \tag{2.101}$$

Considérons le déterminant

$$\mathbf{d} = \left| \mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_j + \sum_{k \neq j} \lambda_k \mathbf{A}_k, \cdots \mathbf{A}_n \right| \tag{2.102}$$

Par linéarité du déterminant par rapport à chacun des vecteurs colonnes, nous avons

$$\mathbf{d} = \det A + \sum_{k \neq j} \lambda_k \left| \mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_k, \cdots \mathbf{A}_n \right|. \tag{2.103}$$

Chacun des termes de la somme contient deux fois le même vecteur colonne, et s'annule donc en vertu de la propriété 4. Nous avons donc $\mathbf{d} = \det A$. La preuve de la propriété (2.101) en découle directement en prenant la transposée de la matrice. □

Remarque 7 Il faut faire attention à utiliser ces manipulations à bon escient. Toute manipulation doit pouvoir se décomposer en étapes consistant chacune à une seule substition dans une colonne $\mathbf{A}_j \leftarrow \mathbf{A}_j + \sum_{k \neq j} \lambda_k \mathbf{A}_k$ ou bien une seule substitution de ligne $\mathbf{A}^i \leftarrow \mathbf{A}^i + \sum_{\ell \neq j} \lambda_\ell \mathbf{A}^\ell$.

On pourrait si nous n'y prenons pas garde arriver à des conclusions absurdes. Par exemple, le remplacement simultané $\mathbf{A}_1 \leftarrow \mathbf{A}_1 - \mathbf{A}_2$, $\mathbf{A}_2 \leftarrow \mathbf{A}_2 - \mathbf{A}_1$, s'il était autorisé, permettrait d'obtenir que tout déterminant est nul (car on obtiendrait un déterminant avec deux colonnes identiques) !

Exemple 15 Soit le déterminant :

$$\det M = \begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 1 & 6 & 2 \\ 2 & -1 & -3 & 1 \\ 1 & 2 & 9 & -3 \end{vmatrix} \tag{2.104}$$

On peut remarquer que la troisème colonne contient uniquement des multiples de 3. On peut donc faire la manipulation suivante :

$$\det A = \begin{vmatrix} 1 & 2 & 3 & 4 \\ -3 & -3 & 0 & -6 \\ 3 & 1 & 0 & 5 \\ -2 & -4 & 0 & -15 \end{vmatrix} \begin{matrix} \\ \mathbf{A}^2 \leftarrow \mathbf{A}^2 - 2\mathbf{A}^1 \\ \mathbf{A}^3 \leftarrow \mathbf{A}^3 + \mathbf{A}^1 \\ \mathbf{A}^4 \leftarrow \mathbf{A}^4 - 3\mathbf{A}^1 \end{matrix} \tag{2.105}$$

et ce déterminant ce calcule simplement par un développement suivant la troisième colonne :

$$\det A = m^1{}_3 \Delta^1{}_3 = 3(-1)^{1+3} \begin{vmatrix} -3 & -3 & -6 \\ 3 & 1 & 5 \\ -2 & -4 & -15 \end{vmatrix} \tag{2.106}$$

À cette étape on peut soit faire le calcul direct soit faire apparaître encore des zéros par les remplacements sur les colonnes, $\mathbf{A}_2 \leftarrow \mathbf{A}_2 - \mathbf{A}_1$ et $\mathbf{A}_3 \leftarrow \mathbf{A}_3 - 2\mathbf{A}_1$, donnant :

$$\begin{aligned} \det A = 3 \begin{vmatrix} -3 & 0 & 0 \\ 3 & -2 & -1 \\ -2 & -2 & -11 \end{vmatrix} &= 3 \times (-3) \times (-1)^{1+1} \begin{vmatrix} -2 & -1 \\ -2 & -11 \end{vmatrix} \\ &= -9(22-2) = -180 \end{aligned} \tag{2.107}$$

Les combinaisons de lignes et de colonnes permettent non seulement de faire apparaître des zéros mais, pour des matrices dépendant de paramètres, d'obtenir le déterminant sous une forme plus factorisée en faisant apparaître un même facteur sur une ligne ou une colonne qu'on peut alors mettre en facteur en dehors du déterminant.

Exemple 16
Calculons le déterminant suivant, en effectuant tout d'abord $\mathbf{A}^1 \leftarrow \mathbf{A}^1 + \mathbf{A}^2 + \mathbf{A}^3$:

$$\begin{aligned} d = \begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} &= \begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} \\ &= (a+b+c) \begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} \end{aligned} \tag{2.108}$$

On peut ensuite faire la substitution $\mathbf{A}_2 \leftarrow \mathbf{A}_2 - \mathbf{A}_1$ et $\mathbf{A}_2 \leftarrow \mathbf{A}_2 - \mathbf{A}_1$:

$$\begin{aligned} d &= (a+b+c) \begin{vmatrix} 1 & 0 & 0 \\ 2b & -b-c-a & 0 \\ 2c & 0 & -c-a-b \end{vmatrix} \\ &= (a+b+c)^3 \begin{vmatrix} 1 & 0 & 0 \\ 2b & -1 & 0 \\ 2c & 0 & -1 \end{vmatrix} = (a+b+c)^3 \times 1 \times (-1)^{1+1} \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} \\ &= (a+b+c)^3 \end{aligned} \tag{2.109}$$

2.5 Inversion de matrice

Une application importante du déterminant est l'inversion de matrices. Le déterminant permet premièrement de savoir si une matrice est inversible, et deuxièmement de calculer cet inverse.

Propriété 11 *Le déterminant de la matrice identité est* $\det(\mathbb{I}_n) = 1$.

Nous avons immédiatement

$$\det \mathbb{I}_n = \sum_{\sigma \in S_n} \epsilon(\sigma) (\mathbb{I}_n)^{\sigma(1)}_1 (\mathbb{I}_n)^{\sigma(2)}_2 \cdots (\mathbb{I}_n)^{\sigma(n)}_n = \sum_{\sigma \in S_n} \epsilon(\sigma) \delta^{\sigma(1)}_1 \delta^{\sigma(2)}_2 \cdots \delta^{\sigma(n)}_n = 1\,, \tag{2.110}$$

car seule la permutation identité (avec $1 = \sigma(1)$, $2 = \sigma(2)$, etc.) contribue.

Lemme 2 *Si la matrice* $A \in \mathcal{M}_K(n,n)$ *est inversible, alors le déterminant de son inverse* A^{-1} *satisfait*

$$\det\left(A^{-1}\right) \det A = 1\,. \tag{2.111}$$

Il suffit d'utiliser la propriété 7.1 au produit $A^{-1} \cdot A$. Nous avons d'une part

$$\det(A^{-1} \cdot A) = \det(A^{-1}) \det A\,, \tag{2.112}$$

et d'autre part

$$\det(A^{-1} \cdot A) = \det(\mathbb{I}_n) = 1\,, \tag{2.113}$$

d'après la propriété 11. □

Théorème 3 (condition d'inversibilité).
Une matrice $A \in \mathcal{M}_K(n,n)$ *est inversible si et seulement si* $\det A \neq 0$.

Soit $A \in \mathcal{M}_K(n,n)$ une matrice telle que $\det A \neq 0$. Cela indique que ses n vecteurs colonnes sont linéairement indépendants, comme $|\mathbf{A}_1 \cdots \mathbf{A}_n| \neq 0$. On peut considérer chaque vecteur colonne $\mathbf{A}_i$ comme formé des composantes de l'images $a(\vec{e}_i)$ du vecteur de base $\vec{e}_i$ de la base $\mathcal{B}$ choisie pour E. Si les n vecteurs $\{a(\vec{e}_i), i = 1, \ldots, n\}$ sont linéairement indépendants, il forment une base de E et l'application a associée à la matrice A est bijective. D'après la propriété 13 du cours 1, la matrice A est alors inversible.

La réciproque de ce théorème est évidente, en utilisant le lemme 2 qui implique que, si A^{-1} existe, $\det A \neq 0$. □

*
* *

Le déterminant permet non seulement de savoir si une matrice est inversible, mais aussi de calculer efficacement son inverse proprement dit.

Définition 17 (comatrice). *Soit* $A \in \mathcal{M}_K(n,n)$ *une matrice carrée* $n \times n$. *La comatrice associée à* A, *notée* $\mathrm{Com}\,A$, *est la matrice* $n \times n$ *formée par les cofacteurs de* A *:*

$$(\mathrm{Com}\,A)^i{}_j \overset{\text{déf.}}{=} \Delta^i{}_j = (-1)^{i+j} \det \hat{A}^i{}_j\,. \tag{2.114}$$

Théorème 4 *Soit* $A \in \mathcal{M}_K(n,n)$ *une matrice carrée* $n \times n$. *On a*

$$A \cdot (\mathrm{Com}\,A)^{\mathrm{T}} = (\mathrm{Com}\,A)^{\mathrm{T}} \cdot A = \mathbb{I} \det A \tag{2.115}$$

impliquant, si A *est inversible (c.-à.-d. si* $\det A \neq 0$*), que sa matrice inverse est donnée par :*

$$\boxed{A^{-1} = \frac{1}{\det A} (\mathrm{Com}\,A)^{\mathrm{T}}} \tag{2.116}$$

Considérons l'élément de la i-ème ligne et de la j-ième colonne du produit matriciel $A \cdot (\mathrm{Com}\,A)^{\mathrm{T}}$. Il faut distinguer deux cas de figure :

— Si $j = i$, nous obtenons, en prenant en compte la transposition de la comatrice :

$$\left(A \cdot (\mathrm{Com}\,A)^{T}\right)^{i}{}_{i} = \sum_{k=1}^{n} a^{i}{}_{k}\,(\mathrm{Com}\,A)^{i}{}_{k} = \sum_{k=1}^{n} a^{i}{}_{k}\Delta^{i}{}_{k}\,. \tag{2.117}$$

Cette expression n'est rien d'autre que le développement par rapport à la i-ème ligne du déterminant de A, voir l'éqn. (2.89), donc cette quantité est égale à $\det A$.

— si $j \neq i$, la somme

$$\left(A \cdot (\mathrm{Com}\,A)^{T}\right)^{i}{}_{j} = \sum_{k=1}^{n} a^{i}{}_{k}\,(\mathrm{Com}\,A)^{j}{}_{k} = \sum_{k=1}^{n} a^{i}{}_{k}\Delta^{j}{}_{k}\,. \tag{2.118}$$

peut s'interpréter, en comparant avec l'éqn. (2.89), comme le développement suivant la j-ième ligne du déterminant d'une matrice obtenue à partir de A en remplaçant les éléments de la j-ième ligne par ceux de la i-ème ligne. La matrice en question ayant deux lignes égales, son déterminant est nul.

En combinant ces deux cas de figure, nous avons donc montré que :

$$\left(A \cdot (\mathrm{Com}\,A)^{T}\right)^{i}{}_{j} = \delta^{i}{}_{j} \det A\,. \tag{2.119}$$

On obtient de la même manière, en considérant le développement (2.91) par rapport aux colonnes, que

$$\left((\mathrm{Com}\,A)^{T} \cdot A\right)^{i}{}_{j} = \delta^{i}{}_{j} \det A\,. \tag{2.120}$$

Lorsque $\det A \neq 0$, on peut évidemment diviser ces deux égalités par $\det A$, et on obtient que $(\mathrm{Com}\,A)^{T}$ est bien la matrice inverse de A. □

Exemple 17 Soit une matrice 2×2 inversible quelconque,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\,, \quad \det A = ad - bc \neq 0\,. \tag{2.121}$$

Sa comatrice est donnée par

$$\mathrm{Com}\,A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \tag{2.122}$$

et on en déduit son inverse :

$$A^{-1} = \frac{1}{\det A}(\mathrm{Com}\,A)^{T} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\,. \tag{2.123}$$

2.6 Déterminant d'un endomorphisme

Rappelons qu'une matrice carrée peut être interprétée comme l'ensemble des composantes des images d'une base $\mathcal{B}_E = \{\vec{e}_i, i = 1, \dots n\}$ par un endomorphisme $a : E \to E$, c.-à.-d.

$$\vec{e}_i \;\mapsto\; a(\vec{e}_i) = \sum_{j=1}^{n} \vec{e}_j \, a^i{}_j \,. \tag{2.124}$$

La matrice A, de composantes $(A)^i{}_j = a^i{}_j$, obtenue de la sorte dépend évidemment de la base choisie sur E.

Rappelons d'après le chapitre 1 que dans une autre base $\widehat{\mathcal{B}}_E = \{\vec{\epsilon}_i, i = 1, \dots, n\}$, reliée à la première par la matrice de passage P dont les composantes sont données par :

$$\vec{\epsilon}_i = \sum_{j=1}^{n} \vec{e}_j P^j{}_i \,, \tag{2.125}$$

la matrice A' associée à l'endomorphisme a s'exprime comme (voir éqn. (1.114)) :

$$A' = P^{-1} \cdot A \cdot P \,, \tag{2.126}$$

c.-à.-d. que A et A' sont des matrices semblables.

Propriété 12 *Deux matrices semblables ont même déterminant.*

Soient $A, A' \in \mathcal{M}_K(n, n)$ deux matrices semblables. Il existe alors une matrice inversible $P \in cM_K(n, n)$ telle que $A' = P^{-1} \cdot A \cdot P$. Calculons le déterminant de A'. Nous avons :

$$\det A' = \det \left(P^{-1} \cdot A \cdot P\right) = \det P^{-1} \times \det A \times \det P = \frac{1}{\det P} \times \det A \times \det P = \det A \,. \tag{2.127}$$

□

Définition 18 (déterminant d'un endomorphisme). *Soit* $a : E \to E$ *un endomorphisme d'un espace vectoriel* E *de dimension* n*, et soit* $\mathcal{B}_E$ *une base quelconque de* E *et* A *la matrice associèe à l'endomorphisme* a *dans cette base. Le déterminant de l'endorphisme* a*, noté* $\det a$*, est alors défini comme :*

$$\det a \overset{\text{déf.}}{=} \det A \,, \tag{2.128}$$

indépendamment de la base choisie.

L'indépendance par rapport au choix de la base découle directement de la propriété 12.

Propriété 13 *Soient* $a : E \to E$ *et* $b : E \to E$ *deux endomorphismes d'un espace vectoriel* E *de dimension* n*. Nous avons*

$$\det(a \circ b) = \det a \times \det b \tag{2.129}$$

Propriété 14 *Soit* $a : E \to E$ *et* $b : E \to E$ *un endomorphisme d'un espace vectoriel* E *de dimension* n*. Cet endomorphisme est bijectif si et seulement si* $\det a \neq 0$*, et alors*

$$\det(a^{-1}) = \frac{1}{\det a} \,. \tag{2.130}$$

Ces propriétés sont immédiates en passant à l'écriture matricielle, dans une base arbitraire de E. □

Cours n°3

Systèmes linéaires

L'objectif de ce bref chapitre est d'appliquer les outils d'algèbre linéaire, en particulier le déterminant, pour résoudre des systèmes linéaires d'équations algébriques. De tels systèmes joueront un rôle important dans les chapitres suivants concernant la réduction des endomorphismes.

3.1 Systèmes linéaires

Définition 1 (système d'équations linéaires algébriques). *Un système d'équations linéaires algébriques, défini pour une famille de paramètres* $\{a^i{}_j \in K; i = 1, \ldots, p, j = 1, \ldots n\}$ *et* $\{b^i \in K; i = 1, \ldots, p\}$ *est donné par le jeu de* p *équations linéaires :*

$$\left\{\begin{array}{rcl} a^1_1 x^1 + a^1{}_2 x^2 + \cdots + a^1{}_n x^n & = & b^1 \\ a^2_1 x^1 + a^2{}_2 x^2 + \cdots + a^2{}_n x^n & = & b^2 \\ \vdots & & \vdots \\ a^p_1 x^1 + a^p{}_2 x^2 + \cdots + a^p{}_n x^n & = & b^p \end{array}\right. \tag{3.1}$$

dont on cherche l'ensemble des solutions $\mathcal{S}$*, éventuellement vide, pour les* n *inconnues du système* $\{x^i \in K; i = 1, \ldots, n\}$.

Définition 2 (systèmes compatibles et incompatibles). *Un système d'équations linéaires est dit incompatible si l'ensemble des solutions* $\mathcal{S}$ *est l'ensemble vide, et dit compatible si* $\mathcal{S}$ *est non vide.*

Remarque 1 (interprétation géométrique). *Chacune des équations algébriques composant un système linéaire,*

$$a^i_1 x^1 + a^i{}_2 x^2 + \cdots + a^i{}_n x^n = b^i, \tag{3.2}$$

peut être interprétée géométriquement comme l'équation d'un hyperplan affine $\mathcal{H} \subset K^n$*. Lorsque* $b^i = 0$*, l'équation (3.2) définit un hyperplan vectoriel* $\mathcal{H}$ *qui est un sous-espace vectoriel de* K^n *de dimension* $(n-1)$*.*

En effet lorsque $b^i = 0$, si $(x^1, \dots, x^n) \in \mathcal{H}$ et $(y^1, \dots, y^n) \in \mathcal{H}$ alors toute combinaison linéaire de ces deux vecteurs de K^n appartient à $\mathcal{H}$:

$$\begin{aligned} & a^i{}_1(\lambda x^1 + \mu y^1) + a^i{}_2(\lambda x^2 + \mu y^2) + \dots + a^i{}_n(\lambda x^n + \mu y^n) \\ & \qquad = \lambda \underbrace{\left(a^i{}_1 x^1 + a^i{}_2 x^2 + \dots + a^i{}_n x^n\right)}_{=0} + \mu \underbrace{\left(a^i{}_1 y^1 + a^i{}_2 y^2 + \dots + a^i{}_n y^n\right)}_{=0} = 0 \end{aligned} \tag{3.3}$$

□

* * *

En conséquence, le système défini par le jeu d'équations (3.11) peut être interprété géométriquement comme le lieu géométrique dans K^n associé à l'intersection, éventuellement vide, de p hyperplans affines.

Exemple 1 (équation de droite dans l'espace). Soit le système d'équations linéaires :

$$\begin{cases} x + y + z & = & 1 \\ z & = & 2 \end{cases} \tag{3.4}$$

La première équation définit l'unique plan de $\mathbb{R}^3$ passant par les points de coordonnées $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$. La seconde équation définit la plan parallèle au plan (Ox, Oy) passant par le point de coordonnées $(0, 0, 2)$. Leur intersection, non vide, définit une droite représentée sur la figure 3.1.

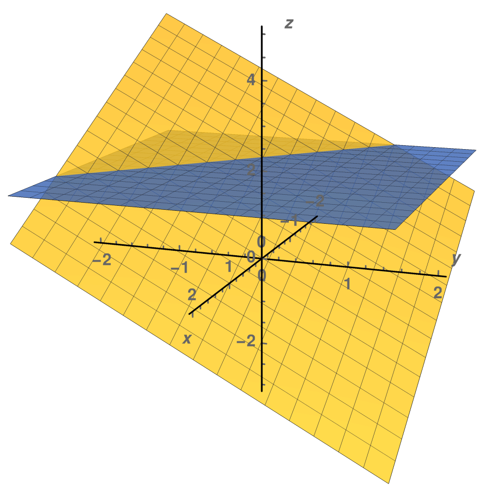

FIGURE 3.1 – Intersection de deux plans.

3.2 Points de vue vectoriel et matriciel

Le système (3.11) peut s'interpréter du point de vue vectoriel en définissant les vecteurs colonne suivants :

$$\mathbf{B} \overset{\text{déf.}}{=} \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^p \end{pmatrix} \quad , \qquad \mathbf{A}_i \overset{\text{déf.}}{=} \begin{pmatrix} {a^1}_i \\ {a^2}_i \\ \vdots \\ {a^p}_i \end{pmatrix} \quad , \quad i \in \{1, \ldots, n\} \tag{3.5}$$

de telle sorte que (3.11) s'écrive :

$$x^1 \mathbf{A}_1 + x^2 \mathbf{A}_2 + \cdots + x^n \mathbf{A}_n = \mathbf{B} \tag{3.6}$$

Propriété 1 *Soit le système linéaire d'équations algébriques défini par l'équation (3.11) et exprimé en termes des vecteurs colonnes (3.5) par l'équation (3.6). On considère que les vecteurs colonnes* $\{\mathbf{A}_i, i = 1, \ldots, n\}$ *et* $\mathbf{B}$ *sont associés au développement dans une base* $\mathcal{B}$ *d'un espace vectoriel* E *de dimension* p *des vecteurs* $\{\vec{a}_i,\ i = 1, \ldots, n\}$ *et* $\vec{b}$ *respectivement.*

1. *Si* $\vec{b} \notin \mathrm{Vect}\,(\vec{a}_1, \ldots, \vec{a}_n)$ *alors le système n'admet pas de solution ;*
2. *Si* $\vec{b} \in \mathrm{Vect}\,(\vec{a}_1, \ldots, \vec{a}_n)$ *alors le système admet des solutions et ces solutions dépendent de* $n - r$ *paramètres, où* $r = \mathrm{Rang}\,(\vec{a}_1, \ldots, \vec{a}_n)$.

La résolution du système (3.6) revient à déterminer l'ensemble :

$$\mathcal{S} = \Big\{ (x^1, \ldots, x^n) \in K^p \quad \text{t.q.} \qquad x^1 \vec{a}_1 + x^2 \vec{a}_2 + \cdots + x^n \vec{a}_n = \vec{b} \Big\} \,. \tag{3.7}$$

c.-à.-d. l'ensemble des combinaisons linéaires des vecteurs $\{\vec{a}_i,\ i = 1, \ldots, n\}$ égales à $\vec{b}$.

Notons $F = \mathrm{Vect}\,(\vec{a}_1, \ldots, \vec{a}_n)$, sous-espace vectoriel de E. Dans le premier cas $\vec{b} \notin F$ et il n'existe aucune combinaison linéaire des vecteurs $\{\vec{a}_i,\ i = 1, \ldots, n\}$ égale à $\vec{b}$, donc $\mathcal{S} = \emptyset$.

Dans le second cas de figure, soit $r = \mathrm{Rang}\,(\vec{a}_1, \ldots, \vec{a}_n)$. Le rang r d'une famille de n vecteurs d'un espace de dimension p satisfait à la fois $r \leqslant n$ et $r \leqslant p$, voir le théorème 4 du cours 1. Parmi les n vecteurs $\{\vec{a}_i,\ i = 1, \ldots, n\}$, r sont linéairement indépendants ; on peut considérer, sans perte de généralité, que ce sont les r premiers. Nous avons alors

$$F = \mathrm{Vect}\,(\vec{a}_1, \ldots, \vec{a}_n) = \mathrm{Vect}\,(\vec{a}_1, \ldots, \vec{a}_r) \,. \tag{3.8}$$

En conséquence,

$$\forall (x^{r+1}, \ldots, x^n) \in K^{n-r} \ , \quad \vec{b} \in F \implies \vec{b} - x^{r+1} \vec{a}_{r+1} - \cdots - x^n \vec{a}_n \in F \,. \tag{3.9}$$

Il existe donc, pour tout $(x^{r+1}, \ldots, x^n) \in K^{n-r}$, des coefficients $(x^1, \ldots, x^r) \in K^r$ tel que

$$\vec{b} - x^{r+1} \vec{a}_{r+1} - \cdots - x^n \vec{a}_n = x^1 \vec{a}_1 + \cdots + x^r \vec{a}_r \,, \tag{3.10}$$

et les $(n - r)$ coefficients $\{x^{r+1}, \ldots, x^n\}$ paramètrent l'ensemble $\mathcal{S}$ des solutions du problème.

Propriété 2 (systèmes homogènes). *Un système linéaire d'équations algébrique est dit homogène si le second membre est nul, c.-à.-d. si il est de la forme*

$$\left\{\begin{array}{rcl} a^1_1 x^1 + a^1{}_2 x^2 + \cdots + a^1{}_n x^n & = & 0 \\ \vdots & & \vdots \\ a^p_1 x^1 + a^p{}_2 x^2 + \cdots + a^p{}_n x^n & = & 0 \end{array}\right. \tag{3.11}$$

L'ensemble $\mathcal{S}$ des solutions d'un tel système n'est jamais vide, car $x^1 = \ldots = x^n = 0$, c.-à.-d. le vecteur nul, est toujours solution. De manière générale, si $r = \mathrm{Rang}\,(\vec{a}_1, \ldots, \vec{a}_n)$, $\mathcal{S}$ est un espace vectoriel de dimension $n - r$, sous-espace vectoriel de $\mathbb{R}^n$.

Nous avons les deux cas de figure suivants :

1. si $\mathrm{Rang}\,(\vec{a}_1, \ldots, \vec{a}_n) = n$ alors ces vecteurs forment une famille libre et $a^1_1 x^1 + a^1{}_2 x^2 + \cdots + a^1{}_n x^n = 0$ a pour seule solution $x^1 = \cdots = x^n = 0$ et $\mathcal{S} = \{\vec{0}\}$;
2. si $r = \mathrm{Rang}\,(\vec{a}_1, \ldots, \vec{a}_n) < n$ remarquons premièrement que :

$$\left\{\begin{array}{l} x^1\vec{a}_1 + \cdots + x^n\vec{a}_n = 0 \\ y^1\vec{a}_1 + \cdots + y^n\vec{a}_n = 0 \end{array}\right. \implies (\lambda x^1 + \mu y^1)\vec{a}_1 + \cdots + (\lambda x^n + y^n)\vec{a}_n = 0\,. \tag{3.12}$$

   Donc si $(x^1, \ldots, x^n)$ et $(y^1, \ldots, y^n)$ sont des solutions, alors $(\lambda x^1 + \mu y^1, \ldots, \lambda x^n + \mu y^n)$ l'est aussi, établissant la nature de sous-espace vectoriel de $\mathcal{S}$.
   En reprenant la discussion précédente, pour tout $(x^{r+1}, \ldots, x^n) \in K^{n-r}$ il existe des coefficients $(x^1, \ldots, x^r) \in K^r$ tel que $x^1\vec{a}_1 + \cdots + x^n\vec{a}_n = -x^{r+1}\vec{a}_{r+1} - \cdots - x^n\vec{a}_n$, et les vecteurs de $\mathcal{S}$ dépendent de $(n - r)$ paramètres indépendants $\{x^{r+1}, \ldots, x^n\}$, donnant la dimension de cet espace vectoriel. $\square$

*
* *

Le système (3.11) peut aussi s'interpréter de manière matricielle. Introduisons une matrice $A \in \mathcal{M}_K(p, n)$ de composantes

$$(A)^i{}_j = a^i_j \tag{3.13}$$

associée à une application linéaire $a : E \to F$, où $\dim E = n$ et $\dim F = p$ et le vecteur colonne

$$\mathbf{X} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}, \tag{3.14}$$

donnant les composantes du vecteur $\vec{x}$ dans la base de E choisie. Le système prend la forme

$$A \cdot \mathbf{X} = \mathbf{B} \;\leftrightarrow\; a(\vec{x}) = \vec{b}\,. \tag{3.15}$$

L'ensemble des solutions $\mathcal{S}$ est alors donné par l'ensemble des vecteurs $\vec{x}$ dont l'image par l'application linéaire a est égale au vecteur $\vec{b}$. *Attention, ce n'est pas un espace vectoriel pour autant, sauf si $\vec{b} = \vec{0}_F$!*

On peut alors reprendre l'analyse de l'espace $\mathcal{S}$ des solutions de ce point de vue et nous obtenons les deux cas de figure suivants :

1. si $\vec{b} \notin \mathrm{Im}\,(a)$ alors $\mathcal{S} = \emptyset$;
2. si $\vec{b} \in \mathrm{Im}\,(a)$ alors, par définition, il existe $\vec{x} \in E$ tel que $a(\vec{x}) = \vec{b}$. Pour tout vecteur $\vec{k} \in \mathrm{Ker}\,(a) \subset E$, $a(\vec{k}) = \vec{0}_F$ donc si $\vec{x}$ est solution alors $\vec{x} + \vec{k}$ l'est aussi. D'après le théorème du rang, si r est le rang de l'application a, alors

$$r + \dim \mathrm{Ker}\,(a) = n\,, \tag{3.16}$$

et on retrouve que l'espace des solutions dépend bien de $(n - r)$ paramètres, correspondant à la décomposition d'un vecteur $\vec{k}$ arbitraire du noyau de a dans la base de E choisie.

Définition 3 (rang d'un système). *Soit un système linéaire d'équations algébriques défini par l'équation (3.11). Le rang r du système est défini indifféremment comme le rang de la famille de vecteurs $(\vec{a}_1, \ldots, \vec{a}_n)$ associés aux vecteurs colonnes définis par l'éqn. (3.5) ou comme le rang de l'application linéaire a associée à la matrice définie par l'équation (3.13).*

Propriété 3 (rang d'une famille de vecteurs, déterminant principal). *Soit une famille de vecteurs $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ d'un espace vectoriel E de dimension p, avec $n \leqslant p$. On considère une sous-famille de q de ces vecteurs, avec $q \leqslant p$. En utilisant la propriété 6 précédemment démontrée, nous savons que s'il existe au moins un déterminant $q \times q$ non-nul, formé avec q composantes des vecteurs de cette sous-famille, alors cette sous-famille est libre.*

Le rang r de la famille de vecteurs est donné par la plus grande valeur de q pour laquelle une telle sous-famille existe, et le déterminant non-nul associé, qui n'est pas nécessairement unique, et dit déterminant principal de la famille de vecteurs.

Exemple 2 Considérons le système linéaire

$$\left\{\begin{array}{lll} x + y & & = 1 \\ x & + z & = 1 \\ & y - z & = 0 \end{array}\right. \tag{3.17}$$

On vérifie tout d'abord que le vecteur colonne $\mathbf{B} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ est dans l'image de l'application linéaire associée à la matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ car il s'agit de l'image du premier vecteur de base (première colonne).

Nous avons $\begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{vmatrix} = 0$ donc le rang du système est inférieur à 3. On peut trouver des sous-déterminants 2×2 non nuls, par exemple $\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$, qui sont donc des déterminants principaux et le rang du système est égal à 2. Ainsi l'espace des solutions, s'il existe, dépend d'un paramètre réel. On peut prendre par exemple comme paramètre z et on obtient par substitution la solution

$$\mathcal{S} = \{(1-z, z, z)\,,\ \in \mathbb{R}\}\,. \tag{3.18}$$

*
* *

3.3 Systèmes de Cramer

Définition 4 (système de Cramer). *Soit un système linéaire de n équations algébriques à n inconnues,*

$$\left\{\begin{array}{lcl} a^1_1 x^1 + a^1{}_2 x^2 + \cdots + a^1{}_n x^n & = & b^1 \\ a^2_1 x^1 + a^2{}_2 x^2 + \cdots + a^2{}_n x^n & = & b^2 \\ \vdots & & \vdots \\ a^n_1 x^1 + a^n{}_2 x^2 + \cdots + a^n{}_n x^n & = & b^n \end{array}\right. \tag{3.19}$$

Ce système est dit de Cramer si la matrice A associée, de composantes $\left(A\right)^i{}_j = a^i_j$, est inversible.

Un système de Cramer est compatible et admet évidemment une solution unique, donnée en notation matricielle par :

$$A \cdot \mathbf{X} = \mathbf{B} \quad \Longrightarrow \quad \mathbf{X} = A^{-1} \cdot \mathbf{B}\,. \tag{3.20}$$

La résolution du problème implique donc l'inversion d'une matrice $n \times n$ et sa multiplication par le vecteur colonne $\mathbf{B}$. Il existe cependant une méthode plus efficace, dite de Cramer.

Propriété 4 (solution d'un système de Cramer). *Soit un système de Cramer de n équations à n inconnues, donné sous forme matricielle par $A \cdot \mathbf{X} = \mathbf{B}$. On définit la matrice :*

$$\hat{A}_i \overset{\text{déf.}}{=} \left(\mathbf{A}_1\, \mathbf{A}_2\, \cdots\, \mathbf{A}_{i-1}\, \mathbf{B}\, \mathbf{A}_{i+1}\, \cdots\, \mathbf{A}_n\right), \tag{3.21}$$

obtenue en remplaçant dans la matrice A le i-ème vecteur colonne $\mathbf{A}_i$ par le vecteur colonne $\mathbf{B}$. La solution du problème est alors donnée par

$$x_i = \frac{\det \hat{A}_i}{\det A}\,, \quad i = 1, \ldots, n\,. \tag{3.22}$$

En termes des vecteurs colonnes, voir l'éqn. (3.6), le système s'écrit

$$x^1 \mathbf{A}_1 + x^2 \mathbf{A}_2 + \cdots + x^n \mathbf{A}_n = \mathbf{B} \tag{3.23}$$

Le déterminant de la matrice $\hat{A}_i$ peut alors s'exprimer de la manière suivante :

$$\begin{aligned} \det \hat{A}_i &= \det\left(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{B}, \ldots,\ \mathbf{A}_n\right) \\ &= \det\left(\mathbf{A}_1, \mathbf{A}_2, \ldots, x^1\mathbf{A}_1 + x^2\mathbf{A}_2 + \cdots + x^n\mathbf{A}_n, \ldots,\ \mathbf{A}_n\right) \\ &= x^1 \det\left((\mathbf{A}_1)\mathbf{A}_2, \ldots, (\mathbf{A}_1), \ldots,\ \mathbf{A}_n\right) + x^2 \det\left(\mathbf{A}_1, (\mathbf{A}_2), \ldots, (\mathbf{A}_2), \ldots,\ \mathbf{A}_n\right) + \cdots \end{aligned} \tag{3.24}$$

Le seul terme non nul dans cette somme est celui qui implique le vecteur colonne $\mathbf{A}_i$, étant donné que les autres donnent des déterminants contenant deux vecteurs colonnes identiques. On en déduit que

$$\det \hat{A}_i = x^i \det\left(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_i, \ldots, \mathbf{A}_n\right) \implies x^i = \frac{\det \hat{A}_i}{\det A} \tag{3.25}$$

$\square$

Exemple 3 Soit le système linéaire

$$\begin{cases} 2x + 3y + z &= 9 \\ x + 2y + 3z &= 6 \\ 3x + y + 2z &= 8 \end{cases} \tag{3.26}$$

Le système est de Cramer car

$$\begin{vmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix} = 18\,. \tag{3.27}$$

La solution du système est donnée par

$$x = \frac{1}{18}\begin{vmatrix} 9 & 3 & 1 \\ 6 & 2 & 3 \\ 8 & 1 & 2 \end{vmatrix} = \frac{35}{18} \tag{3.28a}$$

$$y = \frac{1}{18}\begin{vmatrix} 2 & 9 & 1 \\ 1 & 6 & 3 \\ 3 & 8 & 2 \end{vmatrix} = \frac{29}{18} \tag{3.28b}$$

$$z = \frac{1}{18}\begin{vmatrix} 2 & 3 & 9 \\ 1 & 2 & 6 \\ 3 & 1 & 8 \end{vmatrix} = \frac{5}{18} \tag{3.28c}$$

3.4 Systèmes échelonnés

Rappelons dans cette dernière sous-section quelques outils standard permettant la manipulation et la résolution de système linéaires.

Propriété 5 (systèmes équivalents). *Soit un système linéaire d'équations algébriques :*

$$\begin{array}{c} L_1 \\ L_2 \\ \vdots \\ L_p \end{array} \begin{cases} a^1_1 x^1 + a^1{}_2 x^2 + \cdots + a^1{}_n x^n &= b^1 \\ a^2_1 x^1 + a^2{}_2 x^2 + \cdots + a^2{}_n x^n &= b^2 \\ \qquad\qquad \vdots & \quad \vdots \\ a^p_1 x^1 + a^p{}_2 x^2 + \cdots + a^p{}_n x^n &= b^p \end{cases} \tag{3.29}$$

Ce système est équivalent à tout système obtenu à partir de celui-ci par une combinaison des trois opérations suivantes :

1. *multiplication d'une équation par une constante non-nulle* $\lambda \in K^\star$, $L_i \leftarrow \lambda L_i$;
2. *permutation de l'ordre des lignes,* $L_i \leftarrow L_{\sigma(i)}$, *avec* $\sigma \in S_p$;
3. *ajout d'un multiple d'une autre ligne,* $L_i \leftarrow L_i + \lambda L_j$, *avec* $i \neq j$.

c.-à.-d. qu'ils possèdent le même ensemble $\mathcal{S}$ *de solutions.*

Ces propriétés, dont la démonstration est immédiate, permettent de présenter le système sous une forme « canonique » où le rang de système, et la présence éventuelle de solutions, sont évidents.

Définition 5 (système échelonné). *Soit un système linéaire d'équations algébriques. Ce système est dit échelonné si :*

1. *Pour tout* $i \in \{1, \ldots, p\}$, *il existe* $k_i \geqslant 0$ *tel que parmi les coefficients de la* i*-ème ligne,* $\{a^i{}_1, \ldots, a^i{}_n\}$, *les* k_i *premiers coefficients sont nuls ;*
2. *le nombre* k_i *de premiers coefficients nuls est une fonction croissante de* i, *c.-à.-d. que la* i*-ème ligne contient plus de premiers coefficients nuls que la* $(i-1)$*-ème ligne, et moins que la* $(i+1)$*-ème ligne, le cas échéant.*

Un système se met sous forme échelonnée en utilisant les propriétés 5, qui permettent en particulier d'éliminer des variables par combinaisons linéaires des équations du système. L'algorithme permettant de mettre un système sous forme échelonnée de manière systématique à l'aide de ces opérations élémentaires est appelé *pivot de Gauß*. En pratique l'algorithme se présente de la manière suivante pour un système générique de p équations :

1. par une permutation des équations du système, on place en premier une équation pour laquelle le coefficient de la variable x^1 est le plus simple possible, idéalement égal à 1 ;
2. en remplaçant chacune des autres équations par une combinaison linéaire avec la première, on fait disparaître x^1 de celles-ci ;
3. parmi les $(p-1)$ équations restantes, on cherche celle pour laquelle le coefficient de x^2 et le plus simple, on la place en seconde position dans le système et on l'utilise pour éliminer x^2 de toutes les autres ;
4. on itère le processus jusqu'à avoir mis le système sous forme échelonnée.

Remarque 2 *Pour mettre un système sous forme échelonnée, en particulier pour que les coefficients nuls apparaissent en premier, on peut réétiqueter les inconnues par toute permutation, c.-à.-d. remplacer* $\{x^1, x^2, \ldots, x^n\} \mapsto \{x^{\sigma(1)}, x^{\sigma 2}, \ldots, x^{\sigma(n)}\}$ *avec* $\sigma \in S_n$ *simultanément dans toutes les équations du système.*

Exemple 4 Considérons le système

$$\left\{ \begin{array}{rcl} x + 3y + z & = & 2 \\ x - y + 2z & = & 1 \\ -x + 2y - 6z & = & 3 \end{array} \right. \tag{3.30}$$

Pour mettre ce système sous forme échelonnée on fait d'abord $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 + L_1$ pour éliminer x, ce qui donne

$$\left\{\begin{array}{rcl} x+3y+z &=& 2 \\ -4y+z &=& -1 \\ 5y-5z &=& 1 \end{array}\right. \tag{3.31}$$

Pour éliminer y de la dernière équation on fait ensuite $L_3 \leftarrow 4L_3 + 5L_2$ ce qui donne

$$\left\{\begin{array}{rcl} x+3y+z &=& 2 \\ -4y+z &=& -1 \\ -15z &=& -1 \end{array}\right. \tag{3.32}$$

et on voit immédiatement que le système possède une solution unique, en partant de la dernière équation qui donne $z = 1/15$ puis en substiuant dans la seconde, etc.

Exemple 5 Considérons le système

$$\left\{\begin{array}{rcl} x+3y+z &=& 2 \\ x-y+2z &=& 1 \\ 2x+2y+3z &=& 3 \end{array}\right. \tag{3.33}$$

Pour mettre ce système sous forme échelonnée par les opérations $L_3 \leftarrow L_3 - L_1 - L_2$ puis $L_2 \leftarrow L_2 - L_1$ ce qui donne

$$\left\{\begin{array}{rcl} x+3y+z &=& 2 \\ -4y+z &=& -1 \\ 0 &=& 0 \end{array}\right. \tag{3.34}$$

Ainsi, le système admet une infinité de solutions, que l'on peut donner par exemple sous la forme $\mathcal{S} = \{(3-7y, y, 4y-1),\ y \in \mathbb{R}\}$.

*
* *

Propriété 6 (rang d'un système échelonné). *Le rang d'un système échelonné est égal au nombre de lignes pour lesquelles le membre de gauche de l'équation ne contient pas que des coefficients nuls.*

Il suffit pour démontrer cette propriété de réaliser que les vecteurs ligne correspondants sont nécessairement linéairement indépendants. □

Propriété 7 (solutions d'un système échelonné). Une fois un système d'équations linéaires mis sous une forme échelonnée, on peut immédiatement en déduire quel sont les solutions de ce système. Distinguons plusieurs cas de figure, en prenant des exemples de systèmes à 3 inconnues.

1. Le système possède trois équations, et se présente sous forme échelonné comme

$$\left\{\begin{array}{rcl} x+3y+z &=& 2 \\ -4y+z &=& -1 \\ -15z &=& -1 \end{array}\right. \tag{3.35}$$

   Il s'agit d'un système de Cramer, qui possède une solution unique.

2. Le système possède quatre ou plus équations, et se présente sous forme échelonné comme

$$\left\{\begin{array}{rcl} x+3y+z & = & 2 \\ -4y+z & = & -1 \\ -15z & = & -1 \\ 0 & = & 0 \end{array}\right. \tag{3.36}$$

   Il possède une solution unique, la dernière équation évidemment vraie n'y changeant rien.

3. Le système possède quatre ou plus équations, et se présente sous forme échelonnée comme

$$\left\{\begin{array}{rcl} x+3y+z & = & 2 \\ -4y+z & = & -1 \\ -15z & = & -1 \\ 0 & = & -1 \end{array}\right. \tag{3.37}$$

   Il ne possède pas de solution, la dernière équation étant évidemment fausse. Un tel système est dit *incompatible*.

4. Le système possède trois équations, et se présente sous forme échelonnée comme

$$\left\{\begin{array}{rcl} x+3y+z & = & 2 \\ -4y+z & = & -1 \\ 0 & = & 0 \end{array}\right. \tag{3.38}$$

   Il possède une famille de solutions paramétrée par un entier, par exemple x.

5. Le système possède trois équations, et se présente sous forme échelonnée comme

$$\left\{\begin{array}{rcl} x+3y+z & = & 2 \\ -4y+z & = & -1 \\ 0 & = & 1 \end{array}\right. \tag{3.39}$$

   Ce système est incompatible en raison de la dernière équation.

6. Le système possède trois équations, et se présente sous forme échelonnée comme

$$\left\{\begin{array}{rcl} x+3y+z & = & 2 \\ -1 & = & -1 \\ 0 & = & 0 \end{array}\right. \tag{3.40}$$

   Il possède une famille de solutions paramétrées par deux entiers, par exemple x et y.

Remarque 3 (inversion de matrice). *En partant de l'expression matricielle d'un système de Cramer, voir l'éqn. (3.20), on remarque qu'il est possible d'utiliser les techniques de résolution de systèmes (soit la formule de Cramer, soit la mise sous forme échelonnée par exemple) pour inverser une matrice, en gardant le vecteur-colonne* **B** *du membre de droite générique. Suivant les cas cette méthode peut être la plus rapide.*

Cours n°4

Réduction des endomorphismes I : diagonalisation

Ce chapitre et le suivant constituent le cœur de cet enseignement d'algèbre linéaire, la réduction des endomorphismes sur un espace vectoriel.

La première motivation pour mener ce programme tient à la simplification des opérations sur les matrices associées aux endomorphismes une fois une base $\mathcal{B}$ sur l'espace vectoriel E choisie, en particulier le produit d'endomorphismes, le calcul de polynômes d'endomorphismes et plus généralement de fonctions d'endomorphismes comme l'exponentielle. Il existe une classe particulière de matrices carrées pour lesquelles ces calculs sont particulièrement simples, les matrices diagonales, voir la propriété 22 du chapitre 1. Dès lors se posent les questions suivantes :

1. pour un endomorphisme $a : E \to E$ donné, existe-t-il un base $\mathcal{B}$ telle que la matrice A associée est diagonale ?

2. Lorsqu'une telle base n'existe pas, quelle est la forme la plus simple que peut prendre la matrice A par un choix de base judicieux ?

Dans ce chapitre, nous répondrons au premier point, le second sera l'objet du chapitre suivant.

La seconde motivation est liée à l'interprétation des vecteurs de la base $\mathcal{B}$, si elle existe, telle que la matrice A associée à l'endomorphisme a est diagonale. De tels vecteurs, par définition, satisfont à une équation du type $a(v) = \lambda v$ pour un certain paramètre $\lambda \in K$, et ne se « mélangent » donc pas avec les autres vecteurs sous l'action de l'endomorphisme. Ces vecteurs, appelés *vecteurs propres* de a – qu'ils forment une base ou non – jouent donc un rôle particulier vis à vis de cette appplication. Dans un contexte physique de tels vecteurs peuvent par exemple correspondre à différents modes de vibration d'un système (chaîne d'oscillateurs, membrane vibrante,...) qui sont indépendants les uns des autres, et appelés *modes propres* du système, ou dans un milieu anisotrope les directions principales de propagation de la lumière. En mécanique quantique, comme vous l'apprendrez l'année prochaine, le rôle des vecteurs propres est absolument central.

4.1 Valeurs propres et vecteurs propres, spectre

Définition 1 (valeur propre). *Soit a un endomorphisme d'un K-espace vectoriel E. Un élément $\lambda \in K$ est appelé valeur propre de a si*

$$E_\lambda \stackrel{\text{déf.}}{=} \mathrm{Ker}\left(a - \lambda \mathcal{I}\right) \neq \left\{\vec{0}_E\right\} \tag{4.1}$$

Dans ce cas E_λ, qui est un sous-espace vectoriel de E, est appelé sous-espace propre associé à la valeur propre λ.

Définition 2 (spectre d'un endomorphisme) *Soit a un endomorphisme d'un K-espace vectoriel E. Le spectre de a, noté Spec(a), est l'ensemble de ses valeurs propres λ.*[1]

Définition 3 (vecteur propre) *Soit a un endomorphisme d'un K-espace vectoriel E. Un vecteur $\vec{v} \in E$, non nul, est appelé vecteur propre de a si, pour une certaine valeur propre λ, $v \in E_\lambda$. Autrement dit, il existe une valeur propre $\lambda \in K$ telle que*

$$a(\vec{v}) = \lambda \vec{v}. \tag{4.2}$$

Propriété 1 *Soit a un endomorphisme d'un K-espace vectoriel E et v un vecteur propre de a associé à la valeur propre λ. Nous avons les propriétés :*

1. *pour tout $N \in \mathbb{N}^\star$, $a^N(\vec{v}) = \lambda^n \vec{v}$;*
2. *pour tout polynôme $P \in \mathbb{K}_N[X]$,*

$$\begin{aligned} P(a)(\vec{v}) &\stackrel{\text{déf.}}{=} \left(c_0 + c_1 a + c_2 a^2 + \cdots c_N a^N\right)(\vec{v}) = \left(c_0 + c_1\lambda + c_2\lambda^2 + \cdots c_N\lambda^N\right)(\vec{v}) \\ &= P(\lambda)\vec{v} \end{aligned} \tag{4.3}$$

3. $\exp(a)\,(\vec{v}) = \exp(\lambda)\,\vec{v}$.

La première propriété est évidente par linéarité de a,

$$a^N(\vec{v}) = \underbrace{(a \circ a \circ \cdots \circ a)}_{N \text{ fois}}(\vec{v}) = a^{N-1}(\lambda\vec{v}) = \lambda a^{N-1}(\vec{v}) = \cdots = \lambda^N \vec{v}, \tag{4.4}$$

et les autres en découlent immédiatement. □

Propriété 2 *Soit a un endomorphisme d'un K-espace vectoriel E. Si λ et μ sont deux valeurs propres distinctes de a, alors les sous-espaces propres associés sont en somme directe, c.-à.-d. ont comme seul élément commun le vecteur nul (voir la définition 5 du cours 1).*

$$E_\lambda \cap E_\mu = \{\vec{0}_E\} \tag{4.5}$$

En conséquence, si $\vec{v} \in E_\lambda$ et $\vec{w} \in E_\mu$ avec λ et μ distincts, alors $\vec{v} \neq \vec{w}$.

1. Strictement parlant, ceci est vrai pour un espace vectoriel de dimension finie. Comme vous le verrez plus tard dans vos études, sur des espaces de Hilbert, généralisant les espaces vectoriels complexes en dimension infinie, tous les éléments du spectre ne correspondent pas nécessairement à des valeurs propres.

Supposons qu'il existe un vecteur non nul $\vec{u} \in E_\lambda \cap E_\mu$. Ce vecteur satisfait alors aux deux relations

$$a(\vec{u}) = \lambda\vec{u}\ , \quad a(\vec{u}) = \mu\vec{u}\,. \tag{4.6}$$

La différence de ces deux relations donne

$$(\lambda - \mu)\vec{u} = a(\vec{u}) - a(\vec{u}) = \vec{0}_E \tag{4.7}$$

donc, d'après la propriété 1.3 du cours 1, $\lambda - \mu = 0$, en contradiction avec l'hypothèse de départ. □

De cette propriété importante découle un non moins important théorème, dont nous ne donnerons pas la démonstration, un peu technique.

Théorème 1 *Soit E un K-espace vectoriel de dimension finie n et $a : E \to E$ un endomorphisme. Si on note $\{\lambda_1, \lambda_2, \dots, \lambda_p\} \in K^p$ l'ensemble des valeurs propres distinctes de a, alors les sous-espaces propres associés forment une somme directe de sous-espaces vectoriels de E :*

$$E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_p} \subset E\,. \tag{4.8}$$

En conséquence, le spectre de a, Spec$(a) = \{\lambda_1, \dots, \lambda_p\}$, contient au plus n éléments.

Une question importante sera de savoir si la somme directe des sous-espaces propres forme une partition de E, c.-à.-d. que $E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_p} = E$. Cela est évidemment vrai (mais pas uniquement) si le spectre de a contient n éléments distincts.

Remarque 1 Une fois choisie une base $\mathcal{B}$ de l'espace vectoriel E, on associe à a une matrice A et à un éventuel vecteur propre $\vec{v}$ pour une valeur propre λ on associe un vecteur colonne $\mathbf{V}$. Par extension on dit que V est vecteur propre de la matrice A car il satisfait l'équation

$$A \cdot \mathbf{V} = \lambda\mathbf{V}\,. \tag{4.9}$$

Exemple 1 Soit la matrice

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.10}$$

Il est aisé de voir que cette matrice admet une première valeur propre $\lambda_1 = 1$, en trouvant le vecteur propre

$$\mathbf{V}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \implies \sigma_1 \cdot \mathbf{V}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{V}_1\,. \tag{4.11}$$

Elle admet une deuxième valeur propre $\lambda_2 = -1$, en trouvant le vecteur propre

$$\mathbf{V}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \implies \sigma_1 \cdot \mathbf{V}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -\mathbf{V}_2\,. \tag{4.12}$$

* * *

4.2 Polynôme caractéristique

Étant donné un endomorphisme, il est crucial de développer une méthode permettant d'obtenir le spectre de ses valeurs propres. Nous partirons de la propriété évidente suivante.

Propriété 3 *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : E \to E$ *un endomorphisme. Si* $\lambda \in \text{Spec}(a)$ *alors*

$$\det\big(a - \lambda \mathcal{I}\big) = 0\,. \tag{4.13}$$

Si $\lambda \in \text{Spec}(a)$, alors il existe par définition un vecteur $\vec{v}$ distinct du vecteur nul de E tel que

$$\big(a - \lambda \mathcal{I}\big)(\vec{v}) = \vec{0}_E\,, \tag{4.14}$$

c.-à.-d. appartenant au noyau de $a - \lambda\mathcal{I}$. D'après le théorème du rang, $\text{Rang}(a - \lambda\mathcal{I}) \leqslant n - 1$, donc l'endomorphisme $(a - \lambda\mathcal{I})$ n'est pas inversible et son déterminant est nul. □

Définition 4 (polynôme caractéristique) *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : E \to E$ *un endomorphisme. Le polynôme caractéristique de* a *est le polynôme d'odre* n $P_a \in \mathbb{K}_n[X]$ *défini par*

$$P_a(X) \overset{\text{déf.}}{=} \det\Big(a - X\mathcal{I}\Big)\,. \tag{4.15}$$

Propriété 4 *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : E \to E$ *un endomorphisme de polynôme caractéristique* P_a*. Nous avons*

$$P_a(X) = (-1)^n X^n + (-1)^{n-1} \text{tr}(a)\, X^{n-1} + \cdots + \det a\,. \tag{4.16}$$

Le terme du plus haut degré provient du développement du terme provenant de la permutation identité dans la somme (2.61), soit

$$(a^1{}_1 - X)(a^2{}_2 - X)\cdots(a^n{}_n - X) = (-1)^n X^n + \cdots \tag{4.17}$$

Il s'agit donc bien d'un polynôme de degré n, de coefficient de plus haut degré $(-1)^n$. Le développement de (4.17) contient ensuite un terme de degré $n-1$, qui s'obtient, dans le développement, en choisissant l'un des facteurs $a^i{}_i$ au lieu de X. On obtient ainsi que $(a^1{}_1 - X)(a^2{}_2 - X)\cdots(a^n{}_n - X) = (-1)^n X^n + (-1)^{n-1}\text{tr}(a)\, X^{n-1} + \cdots$. Les termes du déterminant associés aux permutations autres que l'identité contribuent au plus au terme en X^{n-2}, comme une permutation non-triviale va, dans le meilleur des cas (transposition) faire perdre deux éléments diagonaux par rapport au produit (4.17). Finalement le terme d'ordre zéro en X s'obtient simplement en calculant $P_a(0) = \det a$. □

Propriété 5 *Soit* a *un endomorphisme d'un* K*-espace vectoriel* E*. Cet endomorphisme est singulier, c.-à.-d. non inversible, si et seulement si il admet* $\lambda = 0$ *comme valeur propre.*

Dans le sens direct, il suffit de réaliser que si a non inversible, alors $\det a = 0$, et $P_a(0) = \det a$ d'après la propriété 4. Pour montrer la réciproque, supposons que $\lambda = 0$ soit valeur propre. On a alors $\dim E_{\lambda=0} = \dim \operatorname{Ker} a \geqslant 1$ donc a n'est pas inversible. □

*
* *

Pour la suite quelques rappels sur les polynômes sont nécessaires.

Définition 5 (ordre d'une racine d'un polynôme). *Soit $P \in \mathbb{K}[X]$ un polynôme. On dit que λ est racine de P si $P(\lambda) = 0$, ou de manière équivalente si $(X - \lambda)$ divise P. L'ordre de la racine λ est l'entier $m \in \mathbb{N}^\star$ tel que $(X - \lambda)^m$ divise P mais $(X - \lambda)^{m+1}$ ne divise pas P. De manière équivalente, il existe $Q \in \mathbb{K}[X]$ tel que $Q(\lambda) \neq 0$ et $P(X) = (X - a)^m Q(X)$.*

Définition 6 (polynôme scindé) *Soit $P \in \mathbb{K}[X]$ un polynôme de degré n. Ce polynôme est dit scindé s'il peut se factoriser sous la forme :*

$$P = \alpha(X - \beta_1)(X - \beta_2) \cdots (X - \beta_n) \,, \tag{4.18}$$

où les coefficients $\beta_k \in K$ ne sont pas nécessairement distincts.

De manière évidente, tous les coefficients β_k correspondent à des racines de P, non nécessairement distinctes.

Théorème 2 *Tout polynôme $P \in \mathbb{C}[X]$, donc défini sur le corps des complexes, est scindé. En conséquence, il admet exactement n racines, non nécessairement distinctes.*

Si on regroupe tous les termes d'un polynôme scindé correspondant à des racines identiques, on obtient la décomposition suivante, unique à l'ordre des facteurs près :

$$P = \alpha(X - \lambda_1)^{d_1}(X - \lambda_2)^{d_2} \cdots (X - \lambda_p)^{d_p} \,, \tag{4.19}$$

où les entiers d_k correspondant à l'ordre (ou multiplicité) des racines correspondantes satisfont :

$$d_1 + d_2 + \cdots + d_p = n \tag{4.20}$$

et $p \leqslant n$.

Exemple 2 Le polynôme $P(X) = X^2 + 1$ est scindé si on le considére comme un polynôme sur le corps des complexes, c.-à.-d. comme un élément de $\mathbb{C}[X]$:

$$P(X) = X^2 + 1 = (X - i)(X + i) \,. \tag{4.21}$$

En revanche, vu comme un polynôme sur le corps des réels, c.-à.-d. comme un élélement de $\mathbb{R}[X]$, il est *irreductible.*

*
* *

Propriété 6 *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : E \to E$ *un endomorphisme de polynôme caractéristique* P_a*. Si* λ *est valeur propre de* a*, c.-à.-d.* $\lambda \in \mathrm{Spec}\,(a)$*, si et seulement si* λ *est racine du polynôme caractéristique, c.-à.-d. si* $P_a(\lambda) = 0$.

Le sens direct correspond à la propriété 3 déjà démontrée. Réciproquement, supposons que λ soit une racine de P_a. Alors l'endomorphisme $a - \lambda \mathcal{I}$ n'est pas inversible, donc son rang est inférieur à n et $\dim \mathrm{Ker}\,(a - \lambda\mathcal{I}) \geqslant 1$. En conséquence il existe $\vec{v} \neq \vec{0}_E$ tel que $(a - \lambda\mathcal{I})(\vec{v}) = \vec{0}_E$ donc λ est valeur propre de a. □

Propriété 7 *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : E \to E$ *un endomorphisme de polynôme caractéristique* P_a*. Si* d_λ *est la multiplicité de la racine* λ *du polynôme* P_a*, alors la dimension du sous-espace propre* E_λ *satisfait*

$$\dim E_\lambda \leqslant d_\lambda \,. \tag{4.22}$$

En particulier, si λ *est une racine simple, alors* $\dim E_\lambda = 1$.

Choisissons sur E une base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ telle que $\{\vec{e}_1, \vec{e}_2, \ldots \vec{e}_r\}$ forme une base du sous-espace vectoriel E_λ de dimension $r \leqslant n$; on pose également $q = n - r$. Dans une telle base la matrice A associée à l'endomorphisme a est de la forme

$$A = \begin{pmatrix} \lambda \mathbb{I}_r & B \\ \mathbb{O}_{q,r} & C \end{pmatrix} . \tag{4.23}$$

Nous avons donc, si on calcule le polynôme caractéristique dans cette base

$$P_a(X) = \det\,(A - X\mathbb{I}_n) = \begin{vmatrix} (\lambda - X)\mathbb{I}_r & B \\ \mathbb{O}_{q,r} & C - X\mathbb{I}_{q,q} \end{vmatrix} \tag{4.24}$$

Il s'agit du déterminant d'une matrice triangulaire par blocs et, d'après la propriété 8 du cours 5, nous avons

$$P_a(X) = (\lambda - X)^r \underbrace{\det\,(C - X\mathbb{I}_{q,q})}_{Q(X)} . \tag{4.25}$$

Étant donné que le polynôme $Q(X)$ pourrait éventuellement admettre λ comme racine, nous en déduisons que $r = \dim E_\lambda \leqslant m_r$. □

Exemple 3 Reprenons l'exemple de la matrice

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.26}$$

Son polynôme caractéristique est

$$P(x) = \begin{vmatrix} -X & 1 \\ 1 & -X \end{vmatrix} = X^2 - 1 = (X - 1)(X + 1) . \tag{4.27}$$

Ce polynôme est scindé sur $\mathbb{R}$, et admet deux racines simples $\lambda_\pm = \pm 1$, indiquant l'existence de deux sous-espaces propres $E_{\pm 1}$ de dimension 1. Une base de chacun de ses sous-espaces propres est donnée par les équations (4.10,4.12).

*
* *

4.3 Diagonalisation d'un endomorphisme

Définition 7 (endomorphisme diagonalisable). *Soit* E *un* K*-espace vectoriel de dimension* n *et* $a : \mathsf{E} \to \mathsf{E}$ *un endomorphisme. Cet endomorphisme est dit diagonalisable s'il existe une base de* E *formée de vecteurs propres de* a*, c.-à.-d. une base* $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ *telle que la matrice associée à* a *soit diagonale :*

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \mathrm{diag}\,(\lambda_1, \dots, \lambda_n)\,. \tag{4.28}$$

où les valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ *ne sont pas nécessairement distinctes.*

Théorème 3 *Soit* E *un* K*-espace vectoriel de dimension* n*,* $a : \mathsf{E} \to \mathsf{E}$ *un endomorphisme, et* *Spec*$\,(a) = \{\lambda_1, \dots, \lambda_r\}$ *l'ensemble de ses valeurs propres deux à deux distinctes. L'endomorphisme* a *est diagonalisable si et seulement si*

$$\mathsf{E} = \mathsf{E}_{\lambda_1} \oplus \mathsf{E}_{\lambda_2} \oplus \cdots \oplus \mathsf{E}_{\lambda_r}\,. \tag{4.29}$$

Supposons a diagonalisable, c.-à.-d. qu'il existe une base $\mathcal{B}$ telle que a soit représenté par une matrice de la forme (4.28). L'ordre des vecteurs étant arbitraire, on peut les réordonner par une permutation pour regrouper ceux correspondant à une valeur propre commune λ_i. Dans cette base, la matrice prend évidemment la forme

$$A' = \mathrm{diag}\,\big(\underbrace{\lambda_1, \dots, \lambda_1}_{s_1 \text{ fois}}, \underbrace{\lambda_2, \dots, \lambda_2}_{s_2 \text{ fois}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{s_r \text{ fois}}\big). \tag{4.30}$$

Pour toute valeur propre distincte λ_i, on note $\mathcal{B}_i = \{\vec{e}_{\lambda_i,k},\ k = 1, \dots, s_i\}$, le sous-ensemble des s_i vecteurs de cette base tels que $a(\vec{e}_{\lambda_i,k}) = \lambda_i \vec{e}_{\lambda_i,k}$. On note que, pour tout i, $s_i \leqslant \dim \mathsf{E}_{\lambda_i}$ car la famille $\mathcal{B}_i$ de vecteurs de E_{λ_i} est libre. Nous avons donc d'une part $s_1 + s_2 + \cdots + s_r = n$ (partage de la base en sous-familles) et d'autre part $s_1 + s_2 + \cdots + s_r \leqslant \dim \mathsf{E}_{\lambda_1} + \dim \mathsf{E}_{\lambda_2} + \cdots + \dim \mathsf{E}_{\lambda_r}$. Étant donné que les sous-espaces propres ont pour seule intersection deux à deux le vecteur nul, voir la propriété 2, on en déduit que, nécessairement, $\dim \mathsf{E}_{\lambda_1} + \dim \mathsf{E}_{\lambda_2} + \cdots + \dim \mathsf{E}_{\lambda_r} = n$ et ainsi les sous-espaces propres forment une partition de E et que, pour tout i, $s_i = \dim \mathsf{E}_{\lambda_i}$. La réciproque est évidente. Il suffit de remarquer que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$ forme une base de E composée de vecteurs propres de a qui est donc diagonalisable. □

Théorème 4 *(***critère de diagonalisabilité)** *Soit* E *un* K*-espace vectoriel de dimension* n*,* $a : \mathsf{E} \to \mathsf{E}$ *un endomorphisme de polynôme caractéristique* $\mathsf{P}_a \in \mathbb{K}_n[X]$*. Cet endomorphisme est diagonalisable si et seulement si les deux conditions suivantes sont réunies :*

1. *le polynôme* P_a *est scindé ;*
2. *pour toute valeur propre* λ_i *de* a*, la dimension du sous-espace propre* E_{λ_i} *associé est égale à la multiplicité de la racine* λ_i *du polynôme caractéristique.*

D'après le théorème 3, si a est diagonalisable alors il existe une base de vecteurs propres $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ avec $\mathcal{B}_i = \{\vec{e}_{\lambda_i,k},\ k = 1, \ldots, d_i\}$ vecteurs propres associés à la valeur propre λ_i et où $d_i = \dim E_{\lambda_i}$. On peut calculer le polynôme caractéristique de a dans cette base, et on obtient

$$\begin{aligned} P_a(X) &= \det\Big(\operatorname{diag}\big(\underbrace{\lambda_1 - X, \ldots, \lambda_1 - X}_{d_1 \text{ fois}}, \underbrace{\lambda_2 - X, \ldots, \lambda_2 - X}_{d_2 \text{ fois}}, \ldots, \underbrace{\lambda_r - X, \ldots, \lambda_r - X}_{d_r \text{ fois}}\big)\Big) \\ &= (\lambda_1 - X)^{d_1}(\lambda_2 - X)^{d_2} \cdots (\lambda_r - X)^{d_r} = (-1)^n (X - \lambda_1)^{d_1}(X - \lambda_2)^{d_2} \cdots (X - \lambda_r)^{d_r}, \end{aligned} \tag{4.31}$$

en utilisant à la dernière étape $d_1 + \cdots + d_r = n$. Les propriétés 1. et 2. sont donc satisfaites.

Pour démontrer la réciproque, il suffit de noter que, si la propriété 2. est satisfaite, pour tout i la dimension de E_{λ_i} est égale à multiplicité d_i de la racine λ_i correspondante du polynôme. Le polynôme étant scindé si la propriété 1. est satisfaite, nous avons $d_1 + d_2 + \cdots + d_r = n$ et, les sous-espaces propres E_{λ_i} étant en somme directe, on en déduit que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$ puis que, en utilisant le théorème 3, que a est diagonalisable. $\square$

Corollaire 1 *Soit* E *un* K*-espace vectoriel de dimension* n*,* $a : E \to E$ *un endomorphisme de polynôme caractéristique* $P_a \in \mathbb{K}_n[X]$*. Si* P_a *a exactement* n *racines distinctes alors* a *est diagonalisable.*

Si P_a possède exactement n racines distinctes alors P_a est scindé. En outre pour tout i nous avons $1 \leqslant \dim E_{\lambda_i} \leqslant d_i = 1$, d_i étant la multiplicité de la racine λ_i donc $\dim E_{\lambda_i} = 1$ et a diagonalisable. $\square$.

Remarque 2 *Le caractère scindé ou non du polynôme caractéristique, et donc la diagonalisabilité d'un endomorphisme, dépend crucialement du corps (réels ou complexe) sur lequel l'espace vectoriel est défini. Un endomorphisme diagonalisable sur* $\mathbb{C}$ *ne l'est pas nécessairement sur* $\mathbb{R}$*.* [2]

*
* *

Un intérêt évident d'avoir associé à l'endomorphisme une matrice diagonale provient des propriétés suivantes.

Propriété 8 *Soit* a *un endomorphisme d'un* K*-espace vectoriel* E *diagonalisable et* Λ *la matrice diagonale associée dans une base* $\mathcal{B}$ *de* E *formée de vecteurs propres de* a*. Nous avons les propriétés :*

1. *pour tout* $N \in \mathbb{N}^\star$,
$$\Lambda^N = \operatorname{diag}(\lambda_1^N, \ldots, \lambda_n^N) \tag{4.32}$$

2. À l'inverse un endomorphisme diagonalisable sur $\mathbb{R}$ l'est évidemment sur $\mathbb{C}$.

2. *pour tout polynôme* $P \in \mathbb{K}_N[X]$,

$$P(\Lambda) \overset{\text{déf.}}{=} c_0 + c_1\Lambda + c_2\Lambda^2 + \cdots c_N\Lambda^N = \text{diag}\left(P(\lambda_1), \ldots, P(\lambda_n)\right) \tag{4.33}$$

3. $\exp(\Lambda) = \text{diag}\left(\exp\lambda_1, \ldots, \exp\lambda_n\right)$.

Ces propriétés découlent immédiatement des propriétés 1 appliquées à chacun des vecteurs propres de a formant la base de E choisie. □

La forme simple prise, par exemple, par la puissance d'une matrice n'est vraie que dans une base où celle-ci est diagonale. Si notre problème mathématique ou physique fait intervenir plusieurs matrices de même dimension, une question importante est de savoir s'il est possible de trouver une telle forme pour toutes ces matrices simultanément.

Théorème 5 (endomorphismes commutants diagonalisables) *Soit* $\mathcal{A} = \{a_1, a_2, \ldots, a_p\}$ *une collection d'endomorphismes sur un* K*-espace vectoriel* E*. Si tout* $a_\ell \in \mathcal{A}$ *est diagonalisable, et si tous les endomorphismes de* $\mathcal{A}$ *commutent deux à deux, c.-à.-d.*

$$\forall \ell, m \ , \quad a_\ell \circ a_m = a_m \circ a_\ell \,, \tag{4.34}$$

alors il existe une base $\mathcal{B}$ *formée de vecteurs propres communs à tout les endomorphismes* $a_\ell \in \mathcal{A}$*.*

Ce théorème, que nous ne démontrerons pas, joue un rôle capital en mécanique quantique.

4.4 Diagonalisation de matrices

En pratique, un endomorphisme sur un espace vectoriel E de dimension n nous est donné, dans une certaine base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ par une matrice A de dimension $n \times n$.

Supposons que l'endomorphisme a soit diagonalisable. Il existe alors, d'après la définition 7, une base $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ de E telle que la matrice associée à l'endorphisme a soit de la forme

$$\Lambda = \text{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right) \tag{4.35}$$

Soit V la matrice de passage de la base $\mathcal{B}$ vers la base $\mathcal{B}'$. Nous avons alors d'après la propriété 18 du chapitre 1 :

$$\boxed{\Lambda = V^{-1} \cdot A \cdot V \quad \Leftrightarrow \quad A = V \cdot \Lambda \cdot V^{-1}} \tag{4.36}$$

Pour *diagonaliser* une matrice A, il faut donc trouver la matrice de passage V vers une base de vecteurs propres de l'endomorphisme a associé.

Algorithme de diagonalisation d'une matrice

Pour une matrice A donnée, l'algorithme permettant de déterminer si une matrice est diagonalisable et, le cas échéant, de la diagonaliser, se présente schématiquement de la façon suivante :

1. on calcule le polynôme caractéristique $P_A(X) = \det(A - X\mathbb{I})$;
2. si le polynôme P_A n'est pas scindé la matrice n'est pas diagonalisable (pour le corps considéré). Si le polynôme caractéristique est scindé, la matrice n'est pas nécessairement diagonalisable dans le cas où il possède des racines multiples ;
3. Pour chaque valeur propre distincte λ_i, on cherche à obtenir une famille libre de vecteurs propres de la plus grande dimension, correspondant ainsi à une base $\mathcal{B}_i = \{\vec{e}_{\lambda_i,k},\ k = 1, \ldots, r_i\}$ de E_{λ_i}. En pratique on résoud le système homogène :
$$(A - \lambda_i\mathbb{I}) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.37}$$
dont les solutions non triviales si elles existent, donnent des vecteurs colonne $\mathbf{X}_{\lambda_i,k}$, fournissant les composantes de vecteurs propres pour la valeur propre λ_i dans la base $\mathcal{B}$ de départ. L'ensemble de ces vecteurs engendre par définition le sous-espace-propre E_{λ_i}, et on cherche alors une base par un choix approprié de solutions du système.
4. si, pour toute valeur propre λ_i, la dimension r_i de E_{λ_i} est égale à la multiplicité d_i de la racine correspondante du polynôme, alors A est diagonalisable.
5. Si la matrice est diagonalisable, la matrice de passage est obtenue par juxtaposition des vecteurs colonnes déterminés précédemment.

Exemple 4 Soit la matrice
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{pmatrix} \tag{4.38}$$
associée à un endomorphisme a dans une base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Son polynôme caractéristique est donné par
$$P_A(X) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 4 & -1-\lambda & 3 \\ 4 & -2 & 4-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -1-\lambda & 3 \\ -2 & 4-\lambda \end{vmatrix} = -(\lambda-1)^2(\lambda-2) \tag{4.39}$$
Il est scindé sur $\mathbb{R}$ donc A est éventuellement diagonalisable sur $\mathbb{R}$. Pour trancher il faut déterminer une base associée à chacun des sous-espaces propres E_2 et E_1, plus particulièrement déterminer si E_1 est de dimension 1 (A non diagonalisable) ou 2 (A diagonalisable).

- **Base de** E_2
  On cherche à résoudre le système

$$\left(A - 2\mathbb{I}_3\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} -x & = & 0 \\ 4x - 3y + 3z & = & 0 \\ 4x - 2y + 2z & = & 0 \end{array} \right. \tag{4.40}$$

  Soit

$$\left\{ \begin{array}{lcl} x & = & 0 \\ -y + z & = & 0 \\ 0 & = & 0 \end{array} \right. \Longrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ y \end{pmatrix} \tag{4.41}$$

  Le système admet donc une infinité de solutions dépendant d'un unique paramètre, par exemple y. Ces vecteurs colonne étant tous colinéaires ils engendrent un sous-espace vectoriel E_2 de dimension 1, dont une base est donnée par

$$\vec{v}_2 = \vec{e}_2 + \vec{e}_3 \leftrightarrow \mathbf{V}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \tag{4.42}$$

- **Base de** E_1
  On cherche à résoudre le système

$$\left(A - \mathbb{I}_3\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} 0 & = & 0 \\ 4x - 2y + 3z & = & 0 \\ 4x - 2y + 3z & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{lcl} 4x - 2y + 3z & = & 0 \\ 0 & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{4.43}$$

  Le système admet donc une infinité de solutions dépendant de deux paramètres par exemple x et y. On peut facilement déterminer deux vecteurs colonne solution linéairement indépendants, par exemple

$$\vec{v}_1 = \vec{e}_1 - \vec{e}_2 - 2\vec{e}_3 \leftrightarrow \mathbf{W}_1 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \tag{4.44}$$

  et

$$\vec{w}_1 = \vec{e}_1 + 2\vec{e}_2 \leftrightarrow \mathbf{V}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \tag{4.45}$$

  Nous en déduisons que A est diagonalisable, car $\dim E_1 = 2$.

On peut ainsi écrire la matrice de passage de la base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ vers la base $\mathcal{B}' = \{\vec{v}_2, \vec{v}_1, \vec{w}_1\}$

$$V = (\mathbf{V}_2 \mathbf{V}_1 \mathbf{W}_1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & -2 & 0 \end{pmatrix} \tag{4.46}$$

puis calculer son inverse,

$$V^{-1} = \begin{pmatrix} 4 & -2 & 3 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \tag{4.47}$$

Puis vérifier que

$$V^{-1} \cdot A \cdot V = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{4.48}$$

Exemple 5 Soit la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \tag{4.49}$$

associée à un endomorphisme a dans une base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$. Son polynôme caractéristique est donné par

$$P_A(X) = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (\lambda - 1)^2 \tag{4.50}$$

Il est scindé sur $\mathbb{R}$ donc A est éventuellement diagonalisable sur $\mathbb{R}$. Cherchons les vecteurs propres associés à l'unique valeur propre $\lambda = 1$.

$$(A - \mathbb{I}_2) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} 2y & = & 0 \\ y & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{lcl} y & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{4.51}$$

Le sous-espace propre E_1 est donc de dimension 1, engendré par exemple par le vecteur

$$\vec{v} = \vec{e}_1 \;\leftrightarrow\; \mathbf{V} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \tag{4.52}$$

et la matrice A n'est pas diagonalisable.

Retrospectivement, un raisonnement simple et général aurait permis d'anticiper ce résultat. Supposons qu'une matrice A de taille $n \times n$, admettant une valeur propre λ de multiplicité n, soit diagonalisable. Alors, en passant à une base formée de vecteurs propres par la matrice de passage V associée, nous aurions

$$A = V \cdot \mathrm{diag}\,(\lambda, \ldots, \lambda) \cdot V^{-1} = \lambda V \cdot \mathbb{I}_n \cdot V^{-1} = \lambda V \cdot V^{-1} = \lambda \mathbb{I}_n \,, \tag{4.53}$$

et A aurait été dès le départ proportionnelle à la matrice identité, indépendamment de la base choisie.

Pour en revenir à notre exemple, la matrice (4.49) n'étant manifestement pas proportionnelle à l'identité et admettant une valeur propre de multiplicité 2, elle ne peut être diagonalisable.

4.5 Applications

Dans cette dernière section nous examinons quelques applications classiques de la diagonalisation de matrices, en mathématiques et en physique.

4.5.1 Puissances et exponentielles de matrices diagonalisables

Propriété 9 (puissance d'une matrice diagonalisable). *Soit* $A \in \mathcal{M}_{\mathbb{K}}(n,n)$ *une matrice diagonalisable, et* V *la matrice de passage telle que*

$$A = V \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot V^{-1}. \tag{4.54}$$

Pour tout $N \in \mathbb{N}^\star$, *nous avons*

$$A^N = V \cdot \operatorname{diag}(\lambda_1^N, \lambda_2^N, \ldots, \lambda_n^N) \cdot V^{-1}. \tag{4.55}$$

Pour démontrer cette propriété, il suffit d'écrire

$$\begin{aligned} A^N &= \big(V \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot \underbrace{V^{-1}\big) \cdot \big(V}_{=\mathbb{I}_n} \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot V^{-1}\big) \cdots \\ &= V \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot \mathbb{I}_n \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot \mathbb{I}_n \cdots (\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot V^{-1} \end{aligned} \tag{4.56}$$

donnant immédiatement le résultat. □

Corollaire 2 *Soit* $A \in \mathcal{M}_{\mathbb{K}}(n,n)$ *une matrice diagonalisable, et* V *la matrice de passage telle que*

$$A = V \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot V^{-1}. \tag{4.57}$$

Pour tout polynôme $P \in \mathbb{K}[X]$, *nous avons*

$$P(A) = V \cdot \operatorname{diag}\big(P(\lambda_1), P(\lambda_2), \ldots, P(\lambda_n)\big) \cdot V^{-1}. \tag{4.58}$$

Il suffit d'écrire

$$P(A) = V \cdot \Big(\sum_{k=1}^{N} c_k A^k\Big) \cdot V^{-1} = \sum_{k=1}^{N} c_k\, V \cdot A^k V^{-1} \tag{4.59}$$

puis d'appliquer la propriété 9 à chacun des monômes. □

Corollaire 3 (exponentielle d'une matrice diagonalisable). *Soit* $A \in \mathcal{M}_{\mathbb{K}}(n,n)$ *une matrice diagonalisable, et* V *la matrice de passage telle que*

$$A = V \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot V^{-1}. \tag{4.60}$$

Nous avons

$$\exp A = V \cdot \operatorname{diag}\big(\exp\lambda_1, \exp\lambda_2, \ldots, \exp\lambda_n\big) \cdot V^{-1}. \tag{4.61}$$

Il suffit d'utiliser la définition de l'exponentielle d'une matrice carrée,

$$\exp A \overset{\text{déf.}}{=} \sum_{N=0}^{\infty} \frac{A^N}{N!} \tag{4.62}$$

où, rappelons-le, on pose $A^0 = \mathbb{I}_n$, puis d'appliquer la propriété 9 à chacun des termes de cette série entière. □

Exemple 6 Reprenons la matrice étudiée à l'exemple 4. On obtient immédiatement que

$$\begin{aligned}
\begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{pmatrix}^{N} &= V \cdot \begin{pmatrix} 2^N & 0 & 0 \\ 0 & 1^N & 0 \\ 0 & 0 & 1^N \end{pmatrix} \cdot V^{-1} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2^N & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -2 & 3 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ -4+2^{N+2} & 3-2^{N+1} & 3(-1+2^N) \\ -4+2^{N+2} & 2-2^{N+1} & -2+3\times 2^N \end{pmatrix}
\end{aligned} \tag{4.63}$$

Comme vérification du calcul, on peut constater que pour $N = 1$, on retrouve bien la matrice A de départ.

On peut ensuite calculer de la même manière l'exponentielle de la matrice A en partant de l'exponentielle de la matrice diagonale. On trouve :

$$\begin{aligned}
\exp \begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{pmatrix} &= V \cdot \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e \end{pmatrix} \cdot V^{-1} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e \end{pmatrix} \cdot \begin{pmatrix} 4 & -2 & 3 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} e & 0 & 0 \\ 4e(e-1) & e(3-2e) & 3e(e-1) \\ 4e(e-1) & 2e(1-e) & e(3e-2) \end{pmatrix}
\end{aligned} \tag{4.64}$$

*
* *

4.5.2 Systèmes d'équations différentielles couplées du premier ordre

On considère un système de n équations différentielles couplées du premier ordre, pour les fonctions $\{x^i(t), i = 1, \ldots, n\}$, de la forme générale :

$$\left\{ \begin{array}{rcl}
\frac{dx^1(t)}{dt} &=& a^1_1 x^1(t) + a^1_2 x^2(t) + \cdots + a^1_n x^n(t) \\
\frac{dx^2(t)}{dt} &=& a^2_1 x^1(t) + a^2_2 x^2(t) + \cdots + a^2_n x^n(t) \\
\vdots & & \vdots \\
\frac{dx^n(t)}{dt} &=& a^n_1 x^1(t) + a^n_2 x^2(t) + \cdots + a^n_n x^n(t)
\end{array} \right. \tag{4.65}$$

avec des conditions initiales $\{x^i(0) = x^i_0 \ , \quad i = 1, \ldots, n\}$.

Afin d'utiliser les méthodes d'algèbre linéaire pour résoudre ce problème, introduisons les vecteurs colonne des inconnues et des conditions initiales,

$$\mathbf{X}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \\ \vdots \\ x^n(t) \end{pmatrix} \quad , \qquad \mathbf{X}_0 = \begin{pmatrix} x_0^1 \\ x_0^2 \\ \vdots \\ x_0^n \end{pmatrix} , \tag{4.66}$$

ainsi que la matrice A de taille $n \times n$ et de composantes $(A)^i{}_j = a^i{}_j$. Le système différentiel (4.65) peut alors se mettre sous la forme

$$\boxed{\frac{d}{dt}\mathbf{X}(t) = A \cdot \mathbf{X}(t) \ , \quad \mathbf{X}(0) = \mathbf{X}_0} \tag{4.67}$$

Propriété 10 *Soit A une matrice carrée de dimension $n \times n$. Définissons l'application*

$$\begin{cases} \mathbb{R} & \rightarrow & \mathcal{M}_K(n, n) \\ t & \mapsto & \exp(t\,A) \end{cases} \tag{4.68}$$

La dérivée de cette application est donnée par

$$\frac{d}{dt} \exp(t\,A) = A \cdot \exp(t\,A) \ . \tag{4.69}$$

Partons du développement de cette application en série entière :

$$\exp(t\,A) = \sum_{N=0}^{\infty} \frac{t^N A^N}{N!} \ . \tag{4.70}$$

En supposant que la dérivée de la série entière soit donnée par la séries des dérivées (ce qui se démontre en étendant les résultats de votre cours d'analyse aux séries entières de matrices), nous obtenons (par le changement d'indice $N' = N - 1$) :

$$\frac{d}{dt} \exp(t\,A) = \sum_{N=1}^{\infty} \frac{N t^{N-1} A^N}{N!} = A \cdot \sum_{N=1}^{\infty} \frac{t^{N-1} A^{N-1}}{(N-1)!} = A \cdot \sum_{N'=0}^{\infty} \frac{t^{N'} A^{N'}}{(N')!} = A \cdot \exp(t\,A) \ . \tag{4.71}$$

Notons que l'unicité de la solution se prouve de la même manière que pour une unique équation différentielle du premier ordre (théorème de Cauchy–Lipschitz).

Une fois ce résultat établi, vérifions que la solution du problème différentiel (4.67) est donnée par :

$$\boxed{\mathbf{X}(t) = \exp(t\,A) \cdot \mathbf{X}_0} \tag{4.72}$$

Nous avons en effet

$$\frac{d}{dt} \exp(t\,A) \cdot \mathbf{X}_0 \ = \ A \cdot \exp(t\,A) \cdot \mathbf{X}_0 = A \cdot \mathbf{X}(t) \tag{4.73a}$$

$$\mathbf{X}(0) \ = \ \exp(\mathbb{O}_n) \cdot \mathbf{X}_0 = \mathbf{X}_0 \tag{4.73b}$$

La solution (4.72) du problème différentiel (4.67) ne dépend pas de la diagonalisabilité de la matrice carrée A. Lorsque la matrice est diagonalisable, le calcul de $\exp(t\,A)$ en est évidemment grandement facilité ; cela permet également de donner une démonstration alternative de ce résultat.

Supposons que la matrice A soit diagonalisable, et soit V la matrice de passage associée, telle que

$$A = V \cdot \Lambda \cdot V^{-1}\,, \quad \Lambda = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{4.74}$$

Réécrivons l'équation (4.67) comme :

$$\frac{d}{dt}\mathbf{X}(t) = V \cdot \Lambda \cdot V^{-1} \cdot \mathbf{X}(t) \;\Leftrightarrow\; \frac{d}{dt}\left(V^{-1}\mathbf{X}(t)\right) = \Lambda \cdot \left(V^{-1} \cdot \mathbf{X}(t)\right) \tag{4.75}$$

et introduisons le vecteur colonne $\mathbf{Y}(t)$ des fonctions inconnues dans la base où la matrice A est diagonale :

$$\mathbf{Y}(t) = \begin{pmatrix} y^1(t) \\ y^2(t) \\ \vdots \\ y^n(t) \end{pmatrix} \stackrel{\text{déf.}}{=} V^{-1} \cdot \mathbf{X}(t)\,, \tag{4.76}$$

De telle sorte que

$$\frac{d}{dt}\begin{pmatrix} y^1(t) \\ y^2(t) \\ \vdots \\ y^n(t) \end{pmatrix} = \Lambda \cdot \mathbf{Y}(t) = \begin{pmatrix} \lambda_1 y^1(t) \\ \lambda_2 y^2(t) \\ \vdots \\ \lambda_n y^n(t) \end{pmatrix} \tag{4.77}$$

Nous avons donc obtenu un système de n équations différentielles linéaires du premier ordre *découplées*, dont la solution générale est bien connue :

$$\left\{\begin{array}{rcl} \frac{dy^1(t)}{dt} & = & \lambda_1 y^1(t) \\ \frac{dy^2(t)}{dt} & = & \lambda_2 y^2(t) \\ \vdots & & \vdots \\ \frac{dy^n(t)}{dt} & = & \lambda_n y^n(t) \end{array}\right. \Longrightarrow \left\{\begin{array}{rcl} y^1(t) & = & y_0^1\, e^{\lambda_1 t} \\ y^1(t) & = & y_0^2\, e^{\lambda_2 t} \\ \vdots & & \vdots \\ y^n(t) & = & y_0^n\, e^{\lambda_n t} \end{array}\right. \tag{4.78}$$

Il reste à déterminer les conditions initiales dans ces variables. Pour cela écrivons

$$\mathbf{Y}(t) = V^{-1}\mathbf{X}(t) \;\Longrightarrow\; \mathbf{Y}_0 \stackrel{\text{déf.}}{=} \begin{pmatrix} y_0^1 \\ y_0^2 \\ \vdots \\ y_0^n \end{pmatrix} = V^{-1}\mathbf{X}_0\,. \tag{4.79}$$

Une fois la solution du problème trouvée pour les fonctions $\{y^1(t), y^2(t), \ldots y^n(t)\}$, on peut revenir aux variables initiales en utilisant

$$\mathbf{X}(t) = V \cdot \mathbf{Y}(t) = V \cdot \begin{pmatrix} y_0^1\, e^{\lambda_1 t} \\ y_0^2\, e^{\lambda_2 t} \\ \vdots \\ y_0^n e^{\lambda_n t} \end{pmatrix}\,. \tag{4.80}$$

Exemple 7 Considérons le système différentiel

$$\left\{ \begin{array}{lcl} \frac{dx^1(t)}{dt} & = & -4x^1(t) + 3x^2(t) \\ \frac{dx^2(t)}{dt} & = & -6x^1(t) + 5x^2(t) \end{array} \right. \tag{4.81}$$

avec les conditions initiales $x^1(0) = 1$, $x^2(0) = -1$. On écrit premièrement le système sous forme matricielle :

$$\frac{d}{dt}\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \underbrace{\begin{pmatrix} -4 & 3 \\ -6 & 5 \end{pmatrix}}_{A} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \tag{4.82}$$

Puis on cherche à diagonaliser la matrice A. Son polynôme caractéristique est

$$P_A(X) = \begin{vmatrix} -4 - X & 3 \\ -6 & 5 - X \end{vmatrix} = X^2 - X - 2 = (X + 1)(X - 2)\,. \tag{4.83}$$

La matrice A est donc diagonalisable. Cherchons un vecteur de base du sous-espace propre E_{-1}. Nous avons

$$(A + \mathbb{I}_2)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} -3x + 3y & = & 0 \\ -6x + 6y & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{lcl} x - y & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{4.84}$$

donc un vecteur de base de E_{-1} correspond au vecteur colonne

$$\mathbf{V}_{-1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{4.85}$$

Cherchons un vecteur de base du sous-espace propre E_2. Nous avons

$$(A - 2\mathbb{I}_2)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{lcl} -6x + 3y & = & 0 \\ -6x + 3y & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{lcl} 2x - y & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{4.86}$$

donc un vecteur de base de E_2 correspond au vecteur colonne

$$\mathbf{V}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{4.87}$$

Nous avons donc la matrice de passage et son inverse

$$V = (\mathbf{V}_{-1}\,\mathbf{V}_2) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\,, \quad V^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\,. \tag{4.88}$$

Nous en déduisons premièrement que

$$\exp\,(At) = V \cdot \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{pmatrix} \cdot V^{-1} = \begin{pmatrix} 2e^{-t} - e^{2t} & -e^{-t} + e^{2t} \\ 2e^{-t} - 2e^{2t} & -e^{-t} + 2e^{2t} \end{pmatrix} \tag{4.89}$$

Puis la solution du problème,

$$\begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} = \exp\,(At) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3e^{-t} - 2e^{2t} \\ 3e^{-t} - 4e^{2t} \end{pmatrix}\,. \tag{4.90}$$

*
* *

4.5.3 Systèmes d'oscillateurs couplés

Un autre type d'application importante de la diagonalisation des matrices concerne l'étude des systèmes d'oscillateurs couplés. Considérons le système représenté sur la figure 4.1, consti-

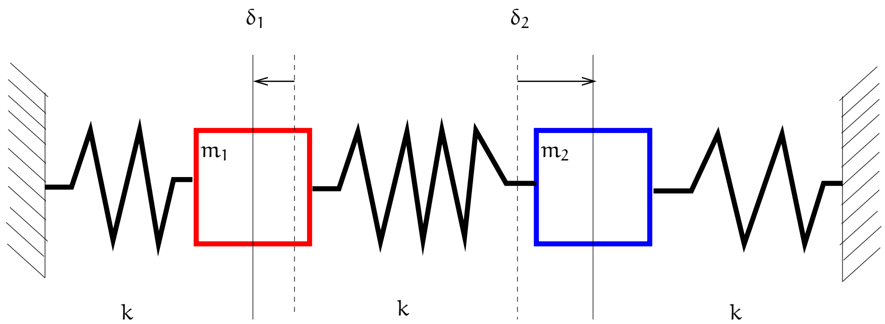

FIGURE 4.1 – Système de deux oscillateurs couplés

tué de deux masses m_1 et m_2, attachées par un ressort de raideur k et reliés de part et d'autre à un support rigide par un ressort de même raideur k. On note δ_1 et δ_2 les déplacements algébriques respectifs des deux masses par rapport à la position d'équilibre.

Il est évident que les déplacement des deux masses sont couplés entre eux. Si on écarte la masse m_1 de sa position d'équilibre, le ressort central va exercer une force sur la masse m_2 qui va se déplacer également. Notre objectif est de trouver des modes de vibration du système *indépendants*, c.-à.-d. qu'il est possible d'exciter l'un sans exciter l'autre.

Le principe fondamental de la dynamique appliqué au système constitué des deux masses, dans le reférentiel attaché au support supposé galiliéen, prend la forme

$$\left\{\begin{array}{rcl} m_1\ddot{\delta}^1 & = & -k\delta^1 + k(\delta^2 - \delta^1) \\ m_2\ddot{\delta}^2 & = & -k\delta^2 + k(\delta^1 - \delta^2) \end{array}\right. \tag{4.91}$$

Le système peut évidemment se mettre sous forme matricielle, comme :

$$\frac{d^2}{dt^2}\begin{pmatrix}\delta^1 \\ \delta^2\end{pmatrix} = \begin{pmatrix}-2k/m_1 & k/m_1 \\ k/m_2 & -2k/m_2\end{pmatrix}\begin{pmatrix}\delta^1 \\ \delta^2\end{pmatrix} \tag{4.92}$$

Pour simplifier les calculs, supposons que les deux masses sont égales, $m_1 = m_2 = m$, et posons $\omega^2 = k/m$. Nous avons alors

$$\frac{d^2}{dt^2}\begin{pmatrix}\delta^1 \\ \delta^2\end{pmatrix} = -\omega^2 \underbrace{\begin{pmatrix}2 & -1 \\ -1 & 2\end{pmatrix}}_{A}\begin{pmatrix}\delta^1 \\ \delta^2\end{pmatrix} \tag{4.93}$$

Diagonalisons la matrice A pour résoudre le problème. Nous avons immédiatement

$$P_A(X) = \begin{vmatrix}2-X & -1 \\ -1 & 2-X\end{vmatrix} = (X-3)(X-1)\,. \tag{4.94}$$

Pour le sous-espace propre $\mathsf{E}_{\sqrt{3}}$ nous avons

$$\begin{cases} -x-y &= 0 \\ -x-y &= 0 \end{cases} \Longrightarrow \mathbf{V}_{\sqrt{3}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{4.95}$$

Puis pour le sous-espace propre E_1 nous trouvons

$$\begin{cases} x-y &= 0 \\ -x+y &= 0 \end{cases} \Longrightarrow \mathbf{V}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{4.96}$$

d'où

$$V = \left(\mathbf{V}_{\sqrt{3}}\ \mathbf{V}_1\right) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad V^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{4.97}$$

Nous pouvons alors écrire le système (4.93) sous la forme :

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{pmatrix} \delta^1 \\ \delta^2 \end{pmatrix} = -\omega^2 V \cdot \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix} \cdot V^{-1} \begin{pmatrix} \delta^1 \\ \delta^2 \end{pmatrix} \tag{4.98}$$

Comme précédemment posons

$$\begin{pmatrix} \Delta^1 \\ \Delta^2 \end{pmatrix} = V^{-1} \cdot \begin{pmatrix} \delta^1 \\ \delta^2 \end{pmatrix} \tag{4.99}$$

de telle sorte que

$$\begin{cases} \ddot{\Delta}^1 &= 3\omega^2\Delta^1 \\ \ddot{\Delta}^2 &= \omega^2\Delta^2 \end{cases} \Longrightarrow \begin{cases} \Delta^1 &= \alpha_1 e^{i\sqrt{3}\omega t} + \beta_1 e^{-i\sqrt{3}\omega t} \\ \Delta^2 &= \alpha_2 e^{i\omega t} + \beta_2 e^{-i\omega t} \end{cases} \tag{4.100}$$

Choisissons comme conditions initiales du problème $\delta^1(0) = \delta$ et $\delta^2(0) = 0$, et pour les vitesses $\dot{\delta}^1(0) = \dot{\delta}^2(0) = \upsilon$. Nous avons alors

$$\begin{pmatrix} \Delta^1 \\ \Delta^2 \end{pmatrix}(0) = \begin{pmatrix} \alpha_1+\beta_1 \\ \alpha_2+\beta_2 \end{pmatrix} = V^{-1} \cdot \begin{pmatrix} \delta^1 \\ \delta^2 \end{pmatrix}(0) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \delta \\ 0 \end{pmatrix} = \delta \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \tag{4.101}$$

Concernant les vitesses, nous avons

$$\begin{pmatrix} \dot{\Delta}^1 \\ \dot{\Delta}^2 \end{pmatrix}(0) = \begin{pmatrix} i\sqrt{3}\omega(\alpha_1-\beta_1) \\ i\omega(\alpha_2-\beta_2) \end{pmatrix} = V^{-1} \cdot \begin{pmatrix} \dot{\delta}^1 \\ \dot{\delta}^2 \end{pmatrix}(0) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \upsilon \\ \upsilon \end{pmatrix} = \upsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.102}$$

Nous obtenons finalement les quatre équations

$$\begin{cases} \alpha_1+\beta_1 &= \delta/2 \\ \alpha_2+\beta_2 &= \delta/2 \\ \alpha_1-\beta_1 &= 0 \\ \alpha_2-\beta_2 &= -i\upsilon/\omega \end{cases} \Leftrightarrow \begin{cases} \alpha_1 &= \delta/4 \\ \beta_1 &= \delta/4 \\ \alpha_2 &= (\delta/2 - i\upsilon/\omega)/2 \\ \beta_2 &= (\delta/2 + i\upsilon/\omega)/2 \end{cases} \tag{4.103}$$

soit

$$\begin{pmatrix} \Delta^1(t) \\ \Delta^2(t) \end{pmatrix} = \begin{pmatrix} \frac{\delta}{2}\cos\sqrt{3}\omega t \\ \frac{\delta}{2}\cos\omega t + \frac{\upsilon}{\omega}\sin\omega t \end{pmatrix} \tag{4.104}$$

et au final

$$\begin{aligned}\begin{pmatrix}\delta^1(t)\\ \delta^2(t)\end{pmatrix} = V\cdot\begin{pmatrix}\Delta^1(t)\\ \Delta^2(t)\end{pmatrix} &= \begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\cdot\begin{pmatrix}\frac{\delta}{2}\cos\sqrt{3}\omega t\\ \frac{\delta}{2}\cos\omega t+\frac{v}{\omega}\sin\omega t\end{pmatrix}\\ &= \begin{pmatrix}\frac{\delta}{2}(\cos\sqrt{3}\omega t+\cos\omega t)+\frac{v}{\omega}\sin\omega t\\ \frac{\delta}{2}(\cos\omega t-\cos\sqrt{3}\omega t)+\frac{v}{\omega}\sin\omega t\end{pmatrix}.\end{aligned} \tag{4.105}$$

On peut vérifier finalement, que les conditions initiales du problème sont bien satisfaites par cette solution.

Au-delà du calcul proprement dit, il est intéressant d'analyser physiquement la solution obtenue. Nous voyons que le mouvement des masses correspond à la superposition de deux modes de vibration :

— un mode de pulsation $\sqrt{3}\omega$. Le vecteur propre colonne associé étant $\mathbf{V}_1 = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$, on peut l'interpréter comme un mouvement des deux masses dans des directions opposées et d'amplitudes égales, voir la figure 4.2.a ;

— un mode de pulsation ω. Le vecteur propre colonne associé étant $\mathbf{V}_2 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$, on peut l'interpréter comme un mouvement d'ensemble des deux masses, voir la figure 4.2.b. Dans ce mode le ressort central garde toujours sa longueur au repos, expliquant qualitativement que ce mode soit de pulsation plus faible que le premier.

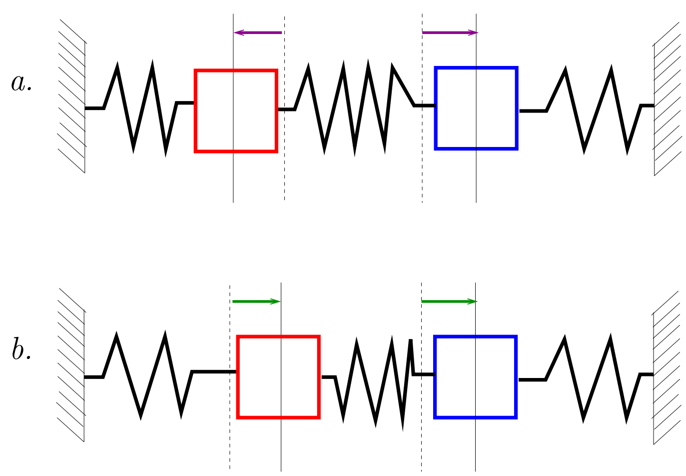

Figure 4.2 – Modes propres de vibration.

Avec des conditions initiales génériques, les deux modes sont excités simultanément. Cependant il est possible de choisir des conditions initiales n'excitant que l'un de ces modes. Par exemple, pour exciter le mouvement d'ensemble des deux masses, on peut choisir $\delta^1 = \delta^2 = \delta$ et $\dot{\delta}^1 = \dot{\delta}^2 = 0$. Cela donne

$$\begin{pmatrix}\Delta^1\\ \Delta^2\end{pmatrix}(0) = \begin{pmatrix}\alpha_1+\beta_1\\ \alpha_2+\beta_2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\cdot\begin{pmatrix}\delta\\ \delta\end{pmatrix} = \delta\begin{pmatrix}0\\ 1\end{pmatrix}. \tag{4.106}$$

Concernant les vitesses, nous avons

$$\begin{pmatrix} \dot{\Delta}^1 \\ \dot{\Delta}^2 \end{pmatrix}(0) = \begin{pmatrix} \sqrt{3}i\omega(\alpha_1 - \beta_1) \\ i\omega(\alpha_2 - \beta_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.107}$$

d'où

$$\begin{pmatrix} \Delta^1(t) \\ \Delta^2(t) \end{pmatrix} = \begin{pmatrix} 0 \\ \delta \cos \omega t \end{pmatrix}. \tag{4.108}$$

Les deux modes de vibration sont ainsi indépendants, et peuvent être considérés comme la base naturelle des degrés de liberté du système ; ils sont appelés *modes propres*.

Ce principe se généralise à des systèmes plus compliqués constitués d'un nombre arbitraire d'oscillateurs. Les modes propres jouent par exemple un rôle important dans la thermodynamique des solides cristallins, dans lesquels les modes de vibration du réseau sont excités par l'agitation thermique.

Cours n°5

Réduction des endomorphismes II : trigonalisation

Dans le chapitre précédent, nous avons étudié les endomorphismes diagonalisables et donné quelques unes de leurs propriétés, en particulier la simplification de nombreux calculs associés comme l'évaluation de puissances ou d'exponentielles de ces endomorphismes. Lorsqu'un endomorphisme n'est pas diagonalisable, nous montrerons ici qu'il existe néanmoins toujours, sur le corps des complexes, un choix de base permettant de simplifier grandement de tels calculs, ainsi que, par exemple, la résolution de systèmes différentiels.

Avant de considérer la théorie générale prenons l'exemple d'une matrice réelle 2×2 de la forme suivante :

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} , \quad \lambda \in \mathbb{R} . \tag{5.1}$$

Une telle matrice est triangulaire, et n'est pas diagonalisable. Nous avons en effet immédiatement (voir également l'exemple 5 du cours 4) :

$$P_A = \det\left(A - X\mathbb{I}\right) = (\lambda - X)^2 \tag{5.2}$$

La valeur propre λ correspond à une racine double, mais la dimension du sous-espace propre E_λ est égale à 1 :

$$\left(A - \lambda\mathbb{I}\right) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{ccc} y & = & 0 \\ 0 & = & 0 \end{array} \right. , \tag{5.3}$$

donc ce sous-espace vectoriel de $\mathbb{R}^2$ est engendré par le vecteur colonne $\mathbf{V} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$.

Même s'il n'est pas possible de mettre cette matrice triangulaire sous une forme plus « simple » par changement de base, une telle forme se prête néanmoins relativement aisément aux calculs. Écrivons

$$A = \lambda\mathbb{I} + D \ , \quad D = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \tag{5.4}$$

où la matrice D est nilpotente, car elle satisfait $D^2 = \mathbb{O}_{2,2}$. Pour cette raison, et comme D commute avec l'identité (comme toute matrice carrée), nous avons pour tout $n \in \mathbb{N}^\star$:

$$A^n = (\lambda\mathbb{I} + D)^n = \lambda^n \mathbb{I} + n\lambda^{n-1}\mathbb{I}\cdot D + \sum_{k=2}^{n} \binom{n}{k} \lambda^{n-k} \underbrace{D^k}_{=0}$$
$$= \lambda^n \mathbb{I} + n\lambda^{n-1} D = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \tag{5.5}$$

De la même manière le calcul de l'exponentielle d'une telle matrice n'est guère difficile. En vue des applications aux systèmes différentiels linéaires, calculons, pour $t \in \mathbb{R}$:

$$\exp(t\,A) = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{t^n\lambda^n}{n!} & \sum_{n=1}^{\infty} t^n \frac{n\lambda^{n-1}}{n!} \\ 0 & \sum_{n=0}^{\infty} \frac{t^n\lambda^n}{n!} \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & t\,e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}. \tag{5.6}$$

Dans chapitre nous verrons à quelles conditions un endomorphisme quelconque peut être représenté, en faisant un choix de base approprié, par une matrice triangulaire par blocs permettant de généraliser ce type de calculs ; un tel programme est appelé *trigonalisation* d'un endomorphisme.

5.1 Théorème de Cayley–Hamilton

Définition 1 (polynôme annulateur). *Soit* $a : E \to E$ *un endomorphisme sur un* K*-espace vectoriel* E *et* $P \in K[X]$ *un polynôme.* P *est appelé polynôme annulateur de* a *si* $P(a) = 0$*, où* 0 *désigne ici l'application nulle.*

Exemple 1 Soit $\mathcal{P} : E \to E$ un projecteur, c.-à.-d. un endomorphisme satisfaisant $\mathcal{P}^2 = \mathcal{P}$. Le polynôme de degré 2 défini par $Q(X) = X^2 - X$ est un polynôme annulateur de $\mathcal{P}$ car $Q(\mathcal{P}) = \mathcal{P}^2 - \mathcal{P} = 0$.

*
* *

Propriété 1 *Soit* E *un* K*-espace vectoriel de dimension* n*. Tout endomorphisme* $a : E \to E$ *admet un polynôme annulateur de degré au plus égal à* n^2*.*

L'ensemble des applications linéaires de E dans E est lui-même un K-espace vectoriel de dimension n^2 ; une fois une base de E choisie il est en effet isomorphe à l'espace vectoriel $\mathcal{M}_K(n,n)$ des matrices $n \times n$ à coefficients dans K, spécifées par la donnée de leur n^2 composantes. En conséquence, la famille de (n^2+1) endomorphismes $\{\mathcal{I}, a, a^2, \ldots, a^{n^2}\}$ est une famille liée. Il existe donc des coefficients $(\alpha_0, \alpha_1, \ldots, \alpha_{n^2}) \in K^{n^2}$, non tous nuls, tels que

$$\alpha_0 \mathcal{I} + \alpha_1 a + \alpha_2 a^2 + \cdots + \alpha_{n^2} a^{n^2} = 0\,. \tag{5.7}$$

Si on définit le polynôme $Q(x) = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \cdots \alpha_{n^2} X^{n^2} \in K[X]$ nous avons évidemment $Q(a) = 0$. □

Définition 2 (polynôme minimal). *Soit* $a : E \to E$ *un endomorphisme sur un* K*-espace vectoriel* E *de dimension* n*. Le polynôme minimal de* a *est l'unique polynôme de plus bas degré, et de coefficient dominant égal à 1, annulateur de* a*.*

Nous savons d'après la propriété 1 que l'ensemble des polynômes annulateurs de l'endomorphisme a est non vide.[1] Supposons qu'il existe dans cet ensemble deux polynômes P_1 et P_2 de plus bas degré d et de coefficient dominant égal à 1, c.-à.-d. que $P_1(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_{d-1}X^{d-1} + X^d$ et $P_2(X) = \beta_0 + \beta_1 X + \cdots + \beta_{d-1}X^{d-1} + X^d$. Les deux polynômes étant annulateurs de a, nous avons

$$0 = P_1(a) - P_2(a) = \big(P_1 - P_2\big)(a) = (\alpha_0 - \beta_0)\mathcal{I} + (\alpha_1 - \beta_1)a + \cdots + (\alpha_{d-1} - \beta_{d-1})a^{d-1} \quad (5.8)$$

donc soit $P_1 - P_2 = 0$, soit $P_1 - P_2$ est un polynôme annulateur de a de degré au plus $d-1$, contredisant l'hypothèse. □

Exemple 2 (homothétie). Une homothétie sur un K-espace vectoriel est un endomorphisme de la forme $a = \lambda\mathcal{I}$, avec $\lambda \in K^\star$. Le polynôme $X - \lambda$ est un polynôme annulateur de a. Étant donné qu'un polynôme annulateur est de degré au moins égal à 1, il s'agit du polynôme minimal de a. Réciproquement, si un endomorphisme admet un polynôme minimal de degré 1, celui-ci est de la forme $P(X) = X + \alpha_0$ et il s'agit donc nécessairement d'une homothétie.

*
* *

Théorème 1 (théorème de Cayley–Hamilton). *Soit* $a : E \to E$ *un endomorphisme sur un* K*-espace vectoriel* E*. Le polynôme minimal de* a*, noté* π_a*, divise son polynôme caractéristique* P_a*, c.-à.-d. qu'il existe* $Q \in K[X]$ *tel que* $P_a(X) = \pi_a(X)Q(X)$*. En particulier, cela implique que* $P_a(a) = 0$*, c.-à.-d. que le polynôme caractéristique est un polynôme annulateur de* a*.*

Soit A la matrice associée à a dans une base $\mathcal{B}$ de E. Les coefficients de la matrice $\mathrm{Com}\big(A - \mathbb{I}_n X\big)$ sont des déterminants $(n-1) \times (n-1)$ donc des polynômes en X de degré au plus $(n-1)$; nous pouvons donc écrire

$$\Big(\mathrm{Com}\big(A - \mathbb{I}_n X\big)\Big)^{\mathrm{T}} = \sum_{\ell=0}^{n-1} A_\ell X^\ell \,, \quad (5.9)$$

où les A_ℓ sont des matrices $(n-1) \times (n-1)$. D'après le théorème 4 du chapitre 5 nous avons :

$$\big(A - \mathbb{I}_n X\big) \cdot \Big(\mathrm{Com}\big(A - \mathbb{I}_n X\big)\Big)^{\mathrm{T}} = \underbrace{\det\big(A - \mathbb{I}_n X\big)}_{= P_a(X)} \mathbb{I}_n \quad (5.10)$$

soit

$$\big(A - \mathbb{I}_n X\big) \cdot \sum_{\ell=0}^{n-1} A_\ell X^\ell = P_a(X)\mathbb{I}_n \,. \quad (5.11)$$

1. Il est aisé de voir, pour ceux d'entre vous familiers avec cette notion, que l'ensemble des polynômes annulateurs est un idéal de l'anneau des polynômes $K[X]$.

Si on note $P_a(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_n X^n$, avec $\alpha_n = (-1)^n$, et en identifiant les coefficients de même degré en X nous trouvons :

$$A \cdot A_0 = \alpha_0 \mathbb{I}_n \,, \quad A \cdot A_1 - A_0 = \alpha_1 \mathbb{I}_n \,, \quad \ldots \,, \quad A \cdot A_{n-1} - A_{n-2} = \alpha_{n-1} \mathbb{I}_n \,, \quad -A_{n-1} = \alpha_n \mathbb{I}_n \,. \tag{5.12}$$

Nous avons donc finalement

$$\begin{aligned} P_a(A) &= \alpha_0 \mathbb{I}_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_n A^n \\ &= \alpha_0 \mathbb{I}_n + \alpha_1 \mathbb{I}_n \cdot A + \alpha_2 \mathbb{I}_n \cdot A^2 + \cdots + \alpha_n \mathbb{I}_n \cdot A^n \\ &= A \cdot A_0 + A \cdot \big(A \cdot A_1 - A_0\big) + \cdots + A^{n-1} \cdot \big(A \cdot A_{n-1} - A_{n-2}\big) - A^n \cdot A_{n-1} \\ &= 0 \end{aligned} \tag{5.13}$$

car tous les termes s'annulent deux à deux. Ceci étant vrai dans n'importe quelle base, nous en déduisons que $P_a(a) = 0$. Montrons en suite que π_a divise P_a. De manière générale la division euclidienne de P_a par π_a s'écrit

$$P_a(X) = \pi_a(X) Q(X) + R(X) \,, \quad Q, R \in K[X] \,. \tag{5.14}$$

Étant donné que $P_a(a) = 0$ et $\pi_a(a) = 0$ par hypothèse, nous avons également $R(a) = 0$ donc R est un polynôme annulateur de a. Dans l''equation (5.14) le degré de R est nécessairement inférieur ou égal au degré de π_a (par définition de la division euclidienne de polynômes). Étant donné que π_a est le polynôme annulateur de plus bas degré, nous avons alors nécessairement $R = 0$, donc il existe un polynôme $Q \in K[X]$ tel que $P_a(X) = \pi_a(X) Q(X)$. □

Corollaire 1 *Le polynôme minimal π_a d'un endomorphisme a sur un K-espace vectoriel E de dimension n est de degré au plus égal à n.*

Cela découle directement du fait que le polynôme caractéristique d'un endomorphisme est de degré $n = \dim E$. □

Proposition 1 *Soit $a : E \to E$ un endomorphisme sur un K-espace vectoriel E. λ est valeur propre de a si et seulement si λ est racine du polynôme minimal π_a. En conséquence si P_a est scindé et n'admet que des racines simples alors $P_a = (-1)^n \pi_a$.*

Soit λ une valeur propre de a. λ est alors une racine du polynôme caractéristique P_a, avec $P_a(X) = \pi_a(X) Q(X)$, c.-à.-d. que $\pi_a(\lambda) Q(\lambda) = 0$ donc $\pi_a(\lambda) = 0$ ou $Q(\lambda) = 0$. Supposons que $\pi_a(\lambda) \neq 0$, c.-à.-d. qu'il existe une constante $c \neq 0$ et un polynôme $S \in K[X]$ tels que

$$\pi_a(X) = (X - \lambda) S(X) + c \quad \Longrightarrow \quad 0 = \pi_a(a) = S(a) \circ \big(a - \lambda \mathcal{I}\big) + c \mathcal{I} \,. \tag{5.15}$$

Si on applique cette relation à un vecteur propre v de a associé à la valeur propre λ on trouve

$$0 = S(a) \Big[\underbrace{\big(a - \lambda \mathcal{I}\big)(v)}_{=0} \Big] + c v \implies c = 0 \,, \tag{5.16}$$

contredisant l'hypothèse de départ. Ainsi, $\pi_a(a) = 0$.

Réciproquement si λ est racine du polynôme minimal π_a c'est également d'après le théorème de Cayley–Hamilton une racine du polynôme caractéristique P_a donc une valeur propre de a. □

Exemple 3 Soit la matrice triangulaire

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \tag{5.17}$$

dont le polynôme caractéristique est $P_A(X) = (1-X)^2(2-X)$. Concernant le polynôme minimal nous avons soit $\pi_a = (X-1)(X-2)$ soit $\pi_A = (X-1)(X-2)^2$. Pour trancher il faut utiliser le fait que π_a doit être un polynôme annulateur de A, c.-à.-d. que $\pi_A(A) = \mathbb{O}_3$. Nous constatons que

$$(A - \mathbb{I}_3) \cdot \big(A - 2\mathbb{I}_3\big) = \begin{pmatrix} 0 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbb{O}_3 \tag{5.18}$$

donc $\pi_A(X) = (X-a)(X-2)^2$. On vérifie évidemment que $\pi_A(A) = \mathbb{O}_3$.

*
* *

Afin d'exploiter les notions de polynôme annulateur et de polynôme minimal pour la diagonalisation/trigonalisation d'endomorphismes, nous aurons besoin du lemme suivant que nous admettrons : [2]

Lemme 1 (lemme de décomposition des noyaux). *Soit E un K-espace vectoriel de dimension n, $a : E \to E$ un endomorphisme et P un polynôme quelconque scindé, c.-à.-d. que P peut se factoriser comme :*

$$P(X) = \alpha_n (X-\lambda_1)^{d_1}(X-\lambda_2)^{d_2}\cdots(X-\lambda_r)^{d_r} \quad , \qquad \forall i,j\ \lambda_i \neq \lambda_j\,. \tag{5.19}$$

Nous avons alors

$$\mathrm{Ker}\, P(a) = \mathrm{Ker}\,(X-\lambda_1)^{d_1} \oplus \mathrm{Ker}\,(X-\lambda_2)^{d_2} \oplus \cdots \mathrm{Ker}\,(X-\lambda_r)^{d_r}\,. \tag{5.20}$$

Nous en déduisons immédiatement un nouveau critère de diagonalisabilité utilisant le polynôme minimal.

Théorème 2 (critère de diagonalisabilité). *Soit E un K-espace vectoriel et $a : E \to E$ un endomorphisme. Cet endomorphisme est diagonalisable si et seulement si son polynôme minimal est scindé et n'admet que des racines simples, c.-à.-d.*

$$\pi_a(X) = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_r) \quad , \qquad \forall i,j\ \lambda_i \neq \lambda_j\,. \tag{5.21}$$

2. Il s'agit en réalité d'un cas particulier d'un lemme plus général stipulant que, pour tout polynôme $P \in K[X]$ se factorisant comme $P = P_1 P_2 \cdots P_r$ ou les P_i sont des polynômes premiers entre eux deux à deux, et pour tout endomorphisme a d'un K-espace vectoriel E, $\mathrm{Ker}\, P(a) = \mathrm{Ker}\, P_1(a) \oplus \cdots \oplus \mathrm{Ker}\, P_r(a)$.

Supposons que l'endormophisme a soit diagonalisable. D'après la proposition 1 toutes les valeurs propres de a, $\{\lambda_1, \lambda_2, \dots, \lambda_r\}$ sont racines du polynôme minimal, et on peut ainsi l'écrire comme

$$\pi_a(X) = \underbrace{(X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_r)}_{=M(X)} Q(X) \ , \quad Q \in K[X] \, . \tag{5.22}$$

Soit v_i un vecteur propre de a associé à une valeur propre λ_i. Nous avons alors

$$M(a)(v_i) = \Big(\prod_{\ell \neq i} (a - \lambda_\ell \mathcal{I})\Big) \Big[\underbrace{(a - \lambda_i \mathcal{I})(v_i)}_{=0_E}\Big] = 0_E \, . \tag{5.23}$$

Si a est diagonalisable, nous savons d'après le théorème 3 du chapitre 4 que E se décompose en somme directe des sous-espaces propres de a :

$$E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_r} \, , \tag{5.24}$$

c.-à.-d. que tout vecteur $v \in E$ se décompose de manière unique comme

$$v = v_1 + v_2 + \cdots v_r \ , \quad v_i \in E_{\lambda_i} \, , \tag{5.25}$$

voir le théorème 1 du chapitre 1. Nous avons donc

$$M(a)(v) = \sum_{i=1}^{r} M(a)(v_i) = 0 \, . \tag{5.26}$$

Le vecteur v étant arbitraire, nous en déduisons que $M(a) = 0$, c.-à.-d. que M est un polynôme annulateur de a. Par définition du polynôme minimal de a, et étant donné que le coefficient de plus haut degré de $M(X)$ est égal à 1, on en déduit que $\mu_a(X) = M(X)$ et la propriété est démontrée.

Pour montrer la réciproque, supposons que le polynôme minimal de a soit de la forme $\pi_a(X) = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_r)$. Le lemme 1 implique alors que

$$E = \operatorname{Ker} \pi_a(a) = \operatorname{Ker}(a - \lambda_1) \oplus \operatorname{Ker}(a - \lambda_2) \oplus \cdots \operatorname{Ker}(X - \lambda_r) \, . \tag{5.27}$$

Pour chaque λ_i tel que $\operatorname{Ker}(X - \lambda_i)$ n'est pas réduit au vecteur nul, ce noyau s'identifie avec le sous-espace propre E_{λ_i} correspondant. Le théorème 3 du chapitre 4 indique alors que, l'espace vectoriel se décomposant en somme directe de sous-espaces propres, l'endomorphisme a est diagonalisable. □

Exemple 4 Reprenons l'exemple de la matrice triangulaire

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \tag{5.28}$$

dont le polynôme caractéristique est $P_A(X) = (1-X)^2(2-X)$. Nous avions montré que son polynôme minimal est donné par $\pi_A(X) = (X-1)(X-2)^2$. La racine $X = 2$ étant double, la matrice A n'est pas diagonalisable.

*
* *

Considérons un endomorphisme $a : E \to E$ non nécessairement diagonalisable, mais dont le polynôme caractéristique est scindé (ce qui est toujours le cas pour un $\mathbb{C}$-espace vectoriel), c.-à.-d. qu'il s'exprime comme

$$P_a(X) = (-1)^n (X-\lambda_1)^{d_1} (X-\lambda_2)^{d_2} \cdots (X-\lambda_r)^{d_r}\,, \tag{5.29}$$

où les valeurs propres λ_i sont toutes distinctes. Le théorème de Cayley–Hamilton indique que le polynôme minimal associé est de la forme

$$\pi_a(X) = (X-\lambda_1)^{c_1} (X-\lambda_2)^{c_2} \cdots (X-\lambda_r)^{c_r}\,, \tag{5.30}$$

où, étant donné que π_a divise P_a, la multiplicité des racines satisfait $c_i \leqslant d_i$ pour tout $i \in \{1, \dots, r\}$.

Définition 3 (sous-espace caractéristique). *Soit* E *un* K*-espace vectoriel,* $a : E \to E$ *un endomorphisme dont le polynôme caractéristique est scindé, c.-à.-d. de la forme donnée par l'éqn. (5.29). Le sous-espace caractéristique associé à la valeur propre* λ_i*, noté* χ_{λ_i}*, est défini par*

$$\chi_{\lambda_i} \overset{\text{déf.}}{=} \operatorname{Ker}\left(\left(a - \lambda_i \mathcal{I}\right)^{d_i}\right). \tag{5.31}$$

Remarquons que lorsque $d_i = 1$ (racine simple), nous avons $\chi_{\lambda_i} = E_{\lambda_i}$, c.-à.-d. que nous retrouvons le sous-espace propre associé à λ_i. De manière générale nous pouvons montrer la propriété suivante.

Propriété 2 *Soit* E *un* K*-espace vectoriel,* $a : E \to E$ *un endomorphisme dont le polynôme caractéristique est scindé. Pour toute racine* λ_i *du polynôme caractéristique, le sous-espace propre* E_{λ_i} *est inclus dans le sous-espace caractéristique* χ_{λ_i}*. En conséquence,* $\dim \chi_{\lambda_i} \geqslant 1$.

Soit un vecteur propre $v_i \in E_{\lambda_i}$. Soit $d_i = 1$, et alors $E_{\lambda_i} = \chi_{\lambda_i}$, soit nous avons

$$\left(a - \lambda_i \mathcal{I}\right)^{d_1}(v_i) = \left(a - \lambda_i \mathcal{I}\right)^{d_1 - 1}\Big[\underbrace{\left(a - \lambda_i \mathcal{I}\right)(v_i)}_{=0_E}\Big] = 0_E \tag{5.32}$$

donc $v \in \chi_{\lambda_i}$ et $E_{\lambda_i} \subseteq \chi_{\lambda_i}$. □

Proposition 2 *Soit* a *un endomorphisme et* χ_{λ_i} *l'un de ses sous-espaces caractéristiques. Si* $v \in \chi_{\lambda_i}$*, alors* $a(v) \in \chi_{\lambda_i}$*, c.-à.-d. que* χ_{λ_i} *est stable sous l'action de* a*.*

En effet, si $\nu \in \chi_{\lambda_i}$ alors $(a-\lambda_i\mathcal{I})^{d_i}(\nu) = 0_E$. On en déduit que

$$\left(a-\lambda_i\mathcal{I}\right)^{d_i}\left[a(\nu)\right] = a\left[\left(a-\lambda_i\mathcal{I}\right)^{d_i}(\nu)\right] = 0_E \tag{5.33}$$

donc $a(\nu) \in \chi_{\lambda_i}$. □

Définition 4 (restriction d'un endomorphisme). *Soit* E *un espace vectoriel,* $F \subset E$ *un sous-espace vectoriel de* E *et* $a : E \to E$ *un endomorphisme. La restriction de* a *à* F*, notée* $a|_F$*, est l'application linéaire :*

$$a|_F \ : \ \begin{cases} F & \to & E \\ \nu & \mapsto & a|_F(\nu) = a(\nu) \end{cases} \tag{5.34}$$

Attention, de manière générale la restriction d'un endomorphisme à un sous-espace vectoriel F n'est pas un endomorphisme de F car l'image de F par $a|_F$ n'est pas nécessairement contenue dans F. En revanche, la proposition précédente nous indique que, pour tout sous-espace caractéristique χ_λ, $a|_{\chi_\lambda}$ est un endomorphisme.

Théorème 3 (dimension des sous-espaces caractéristiques). *Soit* E *un* K*-espace vectoriel,* $a : E \to E$ *un endomorphisme, de polynôme caractéristique scindé, qui s'exprime comme* $P_a(X) = (-1)^n(X-\lambda_1)^{d_1}\cdots(X-\lambda_r)^{d_r}$*. La dimension du sous-espace caractéristique* χ_{λ_i} *associé à la valeur propre* λ_i *est égale à la multiplicité* d_i *de la racine* λ_i *dans le polynôme caractéristique.*

La demonstration, un peu longue, ne sera pas donnée dans le cadre de ce cours.

Propriété 3 (indice de nilpotence). *Chaque sous-espace vectoriel* χ_{λ_i} *étant stable sous l'action de* a*, il l'est également sous l'action de l'endomorphisme* $a-\lambda_i\mathcal{I}$*. La multiplicité* c_i *de la racine* λ_i *dans le polynôme minimal* π_a *peut alors être caractérisée comme,*

$$c_i = \mathit{min}\left\{k \in \mathbb{N}^\star \ \mathit{t.q.} \quad (a-\lambda_i\mathcal{I})^k|_{\chi_{\lambda_i}} = 0\right\}, \tag{5.35}$$

autrement dit la plus petite puissance telle que $(a-\lambda_i\mathcal{I})^k|_{\chi_{\lambda_i}}$ *est un opérateur nilpotent, appelée indice de nilpotence de l'opérateur* $(a-\lambda_i\mathcal{I})|_{\chi_{\lambda_i}}$*.*

En effet s'il existait une puissance $k < c_i$ telle que $(a-\lambda_i\mathcal{I})^k|_{\chi_{\lambda_i}} = 0$ alors en remplaçant dans le polynôme minimal $(X-\lambda_i)^{c_i}$ par $(X-\lambda_i)^k$ on obtiendrait un polynôme annulateur de a de plus bas degré que π_a, ce qui est contraire à la définition du polynôme minimal. □

Théorème 4 *Soit* E *un* K*-espace vectoriel,* $a : E \to E$ *un endomorphisme de polynôme caractéristique est scindé, dont les sous-espaces caractéristiques sont donnés par* $\{\chi_{\lambda_1}, \ldots, \chi_{\lambda_r}\}$*. L'espace vectoriel* E *se décompose alors comme somme directe des sous-espaces caractéristiques,*

$$E = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r}\,. \tag{5.36}$$

Partons du polynôme caractéristique de a,

$$P_a(X) = (-1)^n (X-\lambda_1)^{d_1}(X-\lambda_2)^{d_2}\cdots(X-\lambda_r)^{d_r}\,. \tag{5.37}$$

D'après le théorème de Cayley–Hamilton, $P_a(a) = 0$ donc $E = \operatorname{Ker}\big(P_a(a)\big)$. Le lemme 1 indique alors que

$$E = \operatorname{Ker}\big(P_a(a)\big) = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r}\,. \tag{5.38}$$

$\square$

Nous déduisons de ce dernier théorème un autre critère important de diagonalisabilité d'un endomorphisme.

Propriété 4 *Soit E un K-espace vectoriel et $a : E \to E$ un endomorphisme. Cet endomorphisme est diagonalisable si et seulement si les deux conditions suivantes sont réunies :*

1. *son polynôme caractéristique P_a est scindé ;*
2. *pour chaque valeur propre λ_i, $\chi_{\lambda_i} = E_{\lambda_i}$, c.-à.-d. que chaque sous-espace propre se confond avec le sous-espace caractéristique associé à la même valeur propre.*

Si a est diagonalisable, alors nous savons que P_a est scindé (théorème 4 du chapitre 4) et que

$$E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_r}\,. \tag{5.39}$$

En outre le théorème 4 indique que, P_a étant scindé,

$$E = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r}\,. \tag{5.40}$$

Comme la propriété 2 indique que $E_{\lambda_i} \subseteq \chi_{\lambda_i}$ nous avons alors nécessairement $E_{\lambda_i} = \chi_{\lambda_i}$. Réciproquement, si P_a est scindé alors

$$E = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r}\,, \tag{5.41}$$

et si nous avons également $\chi_{\lambda_i} = E_{\lambda_i}$ pour tout i, alors

$$E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_r}\,. \tag{5.42}$$

et a est diagonalisable. $\square$

Théorème 5 *Soit E un K-espace vectoriel et $a : E \to E$ un endomorphisme de polynôme caractéristique scindé, et de polynôme minimal*

$$\pi_a(X) = (X-\lambda_1)^{c_1}(X-\lambda_2)^{c_2}\cdots(X-\lambda_r)^{c_r}\,. \tag{5.43}$$

Pour tout $i \in \{1,\dots,r\}$, le sous-espace caractéristique χ_{λ_i} est donné par

$$\chi_{\lambda_i} = \operatorname{Ker}\Big[\big(a - \lambda_i \mathcal{I}\big)^{c_i}\Big]\,, \tag{5.44}$$

où c_i est donc la multiplicité de la racine λ_i dans le polynôme minimal π_a.

Étant donné que, pour tout i, $c_i \leqslant d_i$, nous avons d'après la propriété 2 :

$$\operatorname{Ker}\left[\left(a-\lambda_i\mathcal{I}\right)^{c_i}\right] \subseteq \operatorname{Ker}\left[\left(a-\lambda_i\mathcal{I}\right)^{d_i}\right] = \chi_{\lambda_i} . \tag{5.45}$$

Étant donné que $\mu_a(a) = 0$, $\operatorname{Ker}\left(\mu_a(a)\right) = E$ et le lemme 1 indique que

$$E = \operatorname{Ker}\left(\mu_a(a)\right) = \operatorname{Ker}\left[\left(a-\lambda_1\mathcal{I}\right)^{c_1}\right] \oplus \operatorname{Ker}\left[\left(a-\lambda_2\mathcal{I}\right)^{c_2}\right] \oplus \cdots \oplus \operatorname{Ker}\left[\left(a-\lambda_r\mathcal{I}\right)^{c_r}\right] . \tag{5.46}$$

Comme nous avons également

$$E = \operatorname{Ker}\left(P_a(a)\right) = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r} , \tag{5.47}$$

on en déduit que, pour tout i, $\chi_{\lambda_i} = \operatorname{Ker}\left[\left(a-\lambda_i\mathcal{I}\right)^{c_i}\right]$. □

Ce théorème important va être mis à profit dans la section suivante pour comprendre comment trigonaliser de manière systématique tout endomorphisme dont le polynôme caractéristique est scindé.

5.2 Trigonalisation d'un endomorphisme

Nous considérons dans cette section des endomorphismes dont le polynôme caractéristique est scindé, ce qui est toujours le cas sur le corps des complexes. Nous nous intéresserons plus spécialement à ceux dont le polynôme caractéristique admet des racines multiples.

Définition 5 (endomorphisme trigonalisable). *Soit E un espace vectoriel de dimension n et $a : E \to E$ un endomorphisme. Cet endomorphisme est dit trigonalisable s'il existe une base $\mathcal{B}$ de E dans laquelle la matrice A associée à a est triangulaire, supérieure ou inférieure, c.-à.-d. de la forme (ou sa transposée) :*

$$A = \begin{pmatrix} {a^1}_1 & \cdots & \cdots & {a^1}_n \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & {a^n}_n \end{pmatrix} \tag{5.48}$$

Théorème 6 (critère de trigonalisabilité). *Soit E un K-espace vectoriel de dimension n et $a : E \to E$ un endomorphisme. Cet endomorphisme est trigonalisable si et seulement si son polynôme caractéristique P_a est scindé ; en particulier, tout endomorphisme sur un $\mathbb{C}$-espace vectoriel de dimension finie est trigonalisable.*

La preuve dans un sens est évidente. Si a est trigonalisable, calculons son polynôme caractéristique dans une base où la matrice associée est triangulaire :

$$P_A(X) = \begin{vmatrix} {a^1}_1 - X & \cdots & \cdots & {a^1}_n \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & {a^n}_n - X \end{vmatrix} = \left({a^1}_1 - X\right)\left({a^2}_2 - X\right)\cdots\left({a^n}_n - X\right) . \tag{5.49}$$

La réciproque se montre par récurrence. Une matrice 1×1 est évidemment trigonalisable. Supposons que la propriété soit vraie pour des endomorphismes dans des espaces vectoriels de dimension inférieure ou égale à $n-1$. Considérons dans un espace vectoriel E de dimension n un endomorphisme a dont le polynôme caractéristique est scindé. Ce polynôme admet au moins une racine λ, avec un vecteur propre $v \in E$ associé. Soit $\mathcal{B} = \{v, e_2, \dots, e_n\}$ une base de E formée en complétant ce vecteur v en une famille libre et génératrice. Dans cette base, la matrice $n \times n$ associée à a est de la forme

$$A = \begin{pmatrix} \lambda & {a^1}_2 & \cdots & {a^1}_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \tag{5.50}$$

Soit $F \subset E$ le sous-espace vectoriel engendré par la famille libre $\{e_2, \dots, e_n\}$ et l'endomorphisme :

$$b : \begin{cases} F & \longrightarrow & F \\ v & \longmapsto & b(v) = \mathcal{P}_F \circ a|_F \end{cases} , \tag{5.51}$$

où, rappelons-le, $\mathcal{P}_F$ désigne le projecteur sur le sous-espace vectoriel F et $a|_F$ la restriction de a à F. Nous pouvons alors écrire le polynôme caractéristique de a comme

$$P_a(X) = \det\left(A - X\mathbb{I}_n\right) = (\lambda - X)\det(B - X\mathbb{I}_{n-1}) = (\lambda - X)P_b(X)\,. \tag{5.52}$$

et, étant donné que P_a est scindé, P_b l'est également. En utilisant l'hypothèse de récurrence, nous en déduisons qu'il existe une base $\{f_2, f_3, \dots, f_n\}$ du sous-espace vectoriel F telle que la matrice B' associée à l'endomorphisme b soit triangulaire supérieure. Dans la base $\mathcal{B}' = \{v, f_2, \dots, f_n\}$ la matrice associée à l'endomorphisme a est aussi triangulaire supérieure. $\square$

Exemple 5 Soit un endomorphisme a associé dans une certaine base $\mathcal{B}$ d'un espace vectoriel de dimension 3 à la matrice

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 4 & -2 \\ 1 & 2 & 0 \end{pmatrix} \tag{5.53}$$

Son polynôme caractéristique est

$$P_A(X) = -(X-3)(X-2)^2 \tag{5.54}$$

Déterminons également son polynôme minimal. Nous avons

$$\left(A - 3\mathbb{I}\right) \cdot \left(A - 2\mathbb{I}\right) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \neq \mathbb{O} \tag{5.55}$$

donc $\pi_A(X) = P_A(X) = -(X-3)(X-2)^2$. Le polynôme minimal n'admettant pas que des racines simples, nous savons que A n'est pas diagonalisable (cf. théorème 2). Cherchons tout d'abord une base du sous-espace propre E_3, en resolvant le système

$$\left(A - 3\mathbb{I}\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{rcl} y - z &=& 0 \\ x + y - 2z &=& 0 \\ x + 2y - 3z &=& 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} x + y - 2z &=& 0 \\ y - z &=& 0 \\ 0 &=& 0 \end{array} \right. \tag{5.56}$$

dont une solution est $\mathbf{V}_3 = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right)$. Examinons le sous-espace propre E_2. Étant donné que A n'est pas diagonalisable, nous savons d'avance que $\dim E_2 = 1$. Nous avons

$$\left(A - 2\mathbb{I}\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z &=& 0 \\ x + 2y - 2z &=& 0 \\ x + 2y - 2z &=& 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z &=& 0 \\ y - z &=& 0 \\ 0 &=& 0 \end{array} \right. \tag{5.57}$$

dont une solution est $\mathbf{V}_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$; toutes les autres solutions sont proportionnelles à celle-ci. Pour trigonaliser la matrice, il suffit de compléter $\{\mathbf{V}_3, \mathbf{V}_2\}$ par un troisième vecteur colonne quelconque linéairement indépendant de ceux-ci, par exemple $\mathbf{W} = \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right)$. Nous avons alors la matrice de passage et son inverse :

$$V = \left(\mathbf{V}_3\, \mathbf{V}_2\, \mathbf{W}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad , \quad V^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{5.58}$$

et finalement

$$T = V^{-1} \cdot A \cdot V = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} , \tag{5.59}$$

qui est bien une matrice triangulaire.

Il existe cependant un moyen d'obtenir une trigonalisation plus « optimisée » de cette matrice, que nous exposerons sur cet exemple avant de généraliser à la section suivante. Nous partons de l'observation que la troisième colonne de la matrice T s'interprète comme

$$T(w) = v_3 + v_2 + 2w\,, \tag{5.60}$$

où v_3, v_2 et w sont les vecteurs associés respectivement aux vecteurs colonne $\mathbf{V}_3$, $\mathbf{V}_2$ et $\mathbf{W}$, et indique ainsi un « mélange » entre le vecteur v_3 engendrant le sous-espace propre E_3 et les deux autres vecteurs de base sous l'action de l'endomorphisme a.

Pour éviter ce mélange, et donc obtenir une forme plus simple de la matrice triangulaire, nous partons du théorème 4 qui indique (utilisant le fait que le sous-espace caractéristique χ_3 s'identifie au sous-espace propre E_3, la racine correspondante étant une racine simple) :

$$E = \chi_3 \oplus \chi_2 = E_3 \oplus \chi_2\,. \tag{5.61}$$

Au lieu de choisir le vecteur w de manière arbitraire, cherchons à compléter v_2 pour obtenir une base du sous-espace caractéristique

$$\chi_2 = \operatorname{Ker}\left(A - 2\mathbb{I}\right)^2. \tag{5.62}$$

Nous avons :

$$\left(A-2\mathbb{I}\right)^2\cdot\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&1&-1\\1&1&-1\\1&1&-1\end{pmatrix}\cdot\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\left\{\begin{array}{rcl}x+y-z&=&0\\0&=&0\\0&=&0\end{array}\right. \tag{5.63}$$

dont une solution est donnée par le vecteur propre $\mathbf{V}_2=\left(\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right)$, car $E_2\subset\chi_2$. Nous cherchons ensuite une autre solution, linéairement indépendante de $\mathbf{V}_2$, du système (5.63) ; elle est donnée par exemple par $\mathbf{U}=\left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right)$. Nous trouvons alors la matrice de passage

$$\widehat{V}=\left(\mathbf{V}_3\,\mathbf{V}_2\,\mathbf{U}\right)=\begin{pmatrix}1&0&1\\1&1&0\\1&1&1\end{pmatrix}\quad,\quad\widehat{V}^{-1}=\begin{pmatrix}1&1&-1\\-1&0&1\\0&-1&1\end{pmatrix}$$

$$\Longrightarrow\quad\widehat{V}^{-1}\cdot A\cdot\widehat{V}=\begin{pmatrix}3&0&0\\0&2&-1\\0&0&2\end{pmatrix}\quad, \tag{5.64}$$

qui est effectivement plus simple que la matrice (5.59) obtenue précédemment. L'objectif de la section suivante sera d'obtenir de manière systématique la trigonalisation d'un endomorphisme sous une forme canonique, optimisée comme celle-ci.

5.3 Réduction de Jordan

Débutons cette section par des rappels sur les notions acquises jusque là. Soit $a:E\to E$ d'un $\mathbb{C}$-espace vectoriel E de dimension n. À cet endormorphisme peuvent être associés :

— son polynôme caractéristique, de la forme

$$P_a=\det\left(a-X\mathbb{I}\right)=(-1)^n(X-\lambda_1)^{d_1}(X-\lambda_2)^{d_2}\cdots(X-\lambda_r)^{d_r}$$

dont les racines distinctes sont associées à des sous-espace propres $E_{\lambda_i}=\operatorname{Ker}(a-\lambda_i)$, de dimension $1\leqslant\dim E_{\lambda_i}\leqslant d_i$, et formant une somme directe de sous-espaces vectoriels,

$$E_{\lambda_1}\oplus E_{\lambda_2}\oplus\cdots\oplus E_{\lambda_r}\subset E\,,$$

donnant une partition de E lorsque a est diagonalisable ;

— le polynôme minimal π_a, polynôme de plus bas degré et de coefficient dominant 1 tel que $\pi_a(a)=0$; il est de la forme

$$\pi_a(X)=(X-\lambda_1)^{c_1}(X-\lambda_2)^{c_2}\cdots(X-\lambda_r)^{c_r}$$

où, pour tout $i\in\{1,\ldots,r\}$, $1\leqslant c_i\leqslant d_i$ car π_a divise P_a (Cayley–Hamilton).

Le polynôme minimal possède les mêmes racines que le polynôme caractéristique mais leurs multiplicités peuvent être différentes. Le polynôme minimal intervient en particulier dans la décomposition de E en sous-espaces caractéristiques. Le théorème 5 nous apprend en effet que

$$E = \operatorname{Ker}\left(P_a(a)\right) = \chi_{\lambda_1} \oplus \chi_{\lambda_2} \oplus \cdots \oplus \chi_{\lambda_r}\,, \tag{5.65}$$

où, pour tout $i \in \{1, \ldots, r\}$, $\chi_{\lambda_i} = \operatorname{Ker}\left[\left(a - \lambda_i \mathcal{I}\right)^{c_i}\right]$ et ce sous-espace est de dimension d_i.

Définition 6 (matrice diagonale par blocs). *Soit $a : E \to E$ un endomorphisme sur un K-espace vectoriel E de dimension n et $D \in \mathcal{M}_k(n, n)$ la matrice associée à a dans une base $\mathcal{B}$ de E. Cette matrice est dite diagonale par blocs si elle est de la forme*

$$D = \begin{pmatrix} A_1 & 0 & \cdots & \cdots & 0 \\ 0 & A_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_r \end{pmatrix} \tag{5.66}$$

où $A_i \in \mathcal{M}_k(k_i, k_i)$ avec $k_1 + k_2 + \cdots + k_r = n$.

Naturellement, une matrice carrée arbitraire est un exemple trivial de matrice diagonale par blocs, composée d'un seul bloc. À l'autre extrême une matrice diagonale est également diagonale par blocs, tout les blocs étant de dimension 1×1.

Propriété 5 *Soit $a : E \to E$ un endomorphisme sur un K-espace vectoriel E de dimension n dont le polynôme caractéristique est scindé. À chaque sous-espace caractéristique χ_{λ_i}, on associe une base $\mathcal{B}_i$. Soit la famille de vecteurs de E définie par l'union $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$. Nous avons :*

1. *$\mathcal{B}$ est une base de E ;*
2. *la matrice A associée à l'endomorphisme a dans la base $\mathcal{B}$ est diagonale par blocs ;*
3. *chaque bloc A_i correspond à la restriction de a au sous-espace caractéristique χ_{λ_i} correspondant à la valeur propre λ_i ;*
4. *le bloc A_i est de taille $d_i \times d_i$, où d_i est la multiplicité de la racine λ_i dans le polynôme caractéristique $P_a(X)$.*

Le premier point est évident d'après le théorème 4, car $E = \chi_{\lambda_1} \oplus \cdots \oplus \chi_{\lambda_r}$. La notion de décomposition en somme directe (voir la définition 5 du cours 1) implique alors que tout vecteur $v \in E$ se décompose de manière unique commme $v = v_1 + v_2 + \cdots + v_r$ avec $v_i \in \chi_{\lambda_i}$, et chaque v_i se décompose de manière unique sur la base $\mathcal{B}_i$ choisie sur χ_{λ_i}.

Le deuxième point découle directement de la proposition 2 ; si $v_i \in \chi_{\lambda_i}$ alors $a(v_i) \in \chi_{\lambda_i}$. Cela est en particulier vrai si $v_i \in \mathcal{B}_i$, et l'image d'un vecteur de base de χ_{λ_i} se décompose ainsi comme une combinaison linéaire de vecteurs de $\mathcal{B}_i$.

La restriction $a|_{\chi_{\lambda_i}}$ est un endomorphisme sur le sous-espace vectoriel χ_{λ_i}. Il est donc entièrement caractérisé par l'image des vecteurs de base de $\mathcal{B}_i$:

$$a|_{\chi_{\lambda_i}}(e_{\ell,i}) = \sum_{k=1}^{d_i} e_{k,i}\big(A_i\big)^k{}_i\,, \quad \mathcal{B}_i = \left\{e_{1,i}, e_{2,i}, \ldots, e_{d_i,i}\right\}, \tag{5.67}$$

associée à la matrice A_i apparaissant dans la décomposition diagonale par blocs de la matrice A associée à l'endomorphisme a dans la base $\mathcal{B}$.

Enfin, le quatrième point provient du théorème 3 indiquant que la dimension de χ_{λ_i} est égale à la multiplicité d_i de la racine λ_i dans le polynôme caractéristique de a. ☐

Remarque 1 *La matrice A étant diagonale par blocs, la matrice $A - X\mathcal{I}$ l'est aussi et nous avons (voir la propriété 8 du cours 5) :*

$$\det A = \det(A_1 - X\mathbb{I}_{d_1})\det(A_2 - X\mathbb{I}_{d_2})\cdots\det(A_r - X\mathbb{I}_{d_r})\,. \tag{5.68}$$

avec

$$\det\left(A_i - X\mathbb{I}_{d_i}\right) = (-1)^{d_i}(X-\lambda_i)^{d_i}\,. \tag{5.69}$$

Propriété 6 *Soit a un endomorphisme trigonalisable sur un espace vectoriel $E = \chi_{\lambda_1} \oplus \chi_{\lambda_2}\cdots\oplus\chi_{\lambda_r}$, où les χ_{λ_i} sont les sous-espaces caractéristiques de a. La restriction de a à chaque sous-espace caractéristique :*

$$a|_{\chi_{\lambda_i}} \;:\; \left\{\begin{array}{rcl} \chi_{\lambda_i} & \to & \chi_{\lambda_i} \\ v & \mapsto & a(v) \end{array}\right. \tag{5.70}$$

est un un endomorphisme trigonalisable. Dans toute base $\mathcal{B}_i$ de χ_{λ_i} telle que la matrice A_i associée à l'endomorphisme $a|_{\chi_{\lambda_i}}$ est triangulaire supérieure, elle est de la forme

$$A_i = \begin{pmatrix} \lambda_i & \star & \cdots & \star \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix} \tag{5.71}$$

c.-à.-d. triangulaire supérieure avec tous les éléments diagonaux égaux à λ_i.

On note premièrement, cf. remarque 1, que le polynôme caractéristique de $a|_{\chi_{\lambda_i}}$ et donné par

$$P_i \overset{\text{déf.}}{=} \det\Big((a - X\mathcal{I})|_{\chi_{\lambda_i}}\Big) = (-1)^{d_i}(X-\lambda_i)^{d_i} \tag{5.72}$$

qui est scindé, et la matrice A_i est donc trigonalisable d'après le théorème 6. En utilisant l'expression générale du déterminant d'une matrice triangulaire, nous avons

$$P_i = \det\left(A_i - X\mathbb{I}_{d_i}\right) = \Big((A_i)^1{}_1 - X\Big)\Big((A_i)^2{}_2 - X\Big)\cdots\Big((A_i)^{d_i}{}_{d_i} - X\Big)\,. \tag{5.73}$$

En comparant avec l'expression (5.72) nous obtenons par identification des racines du polynôme que ${(A_i)^1}_1 = {(A_i)^2}_2 = \cdots = {(A_i)^{d_i}}_{d_i} = \lambda_i$. Ainsi, tous les éléments diagonaux du bloc de la matrice associé à la restriction au sous-espace caractéristique χ_{λ_i} sont égaux à λ_i. □

*
* *

Récapitulons ce que nous avons obtenu jusqu'ici avant de mener la démarche d'optimisation de la trigonalisation à son terme. Nous avons montré que, pour tout endomorphisme a dont le polynôme caractéristique est scindé :

1. il est possible de lui associer une matrice D diagonale par blocs, voir l'éqn. (5.66), chaque bloc correspondant à un des sous-espaces caractéristiques χ_{λ_i} ;
2. pour chaque bloc A_i, il est possible de choisir une base telle que ce bloc soit une matrice triangulaire supérieure, et ses éléments diagonaux sont toujours tous égaux à λ_i, voir l'éqn. (5.71).

Il n'est pas possible de faire mieux d'une manière générique, et il faut alors rentrer dans la discussion des différents cas de figure, suivant la dimension du sous-espace caractéristique concerné. Avant de décrire le résultat général, nous allons examiner ce qu'il est possible d'obtenir pour les plus petites dimensions de ces blocs.

Bloc 1x1.

Il s'agit de la situation pour laquelle $\dim E_\lambda = \dim \chi_\lambda = 1$. Le sous-espace propre, de dimension 1, est engendré par tout vecteur distinct du vecteur nul solution de l'équation $(a - \lambda \mathcal{I})(v) = 0_E$.

Bloc 2x2.

Il s'agit de la situation pour laquelle le polynôme caractéristique de a peut se factoriser sous la forme

$$P_a(X) = (X - \lambda)^2 Q(X) \ , \quad Q(X) \neq 0 \, . \tag{5.74}$$

Le sous-espace propre E_λ et le sous-espace caractéristique χ_λ associés à la valeur propre λ sont tels que :

$$1 \leqslant \dim E_\lambda \leqslant \dim \chi_\lambda = 2 \, . \tag{5.75}$$

Donc soit $\dim E_\lambda = \dim \chi_\lambda = 2$, et le bloc est alors diagonalisable, soit $\dim E_\lambda = 1$ et le bloc est trigonalisable. Si l'endomorphisme est caractérisé initialement par une matrice A, on étudie l'espace vectoriel des solutions du système homogène

$$\left(A - \lambda \mathbb{I}\right) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \, . \tag{5.76}$$

Si cet espace de solutions est de dimension 2, on choisit 2 vecteurs colonne solutions linéairement indépendants, $\mathbf{V}_\lambda$ et $\mathbf{W}_\lambda$, qui sont par construction générateurs du sous-espace propre E_λ de dimension 2. Après changement de base nous aurons alors un bloc de la forme

$$A_i = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \tag{5.77}$$

Comme nous avons dans ce cas $E_\lambda = \chi_\lambda$, donc $\chi_\lambda = \operatorname{Ker}(a - \lambda\mathcal{I})$, le polynôme minimal associé à a est nécessairement de la forme

$$\pi_a(X) = (X - \lambda)\widehat{Q}(X) \;, \quad \widehat{Q}(\lambda) \neq 0 \,. \tag{5.78}$$

*
* *

Si l'espace vectoriel des solutions du système (5.76) est de dimension 1 , nous sommes dans la situation où le polynôme minimal est de la forme

$$\pi_a(X) = (X - \lambda)^2\widehat{Q}(X) \;, \quad \widehat{Q}(\lambda) \neq 0 \,. \tag{5.79}$$

et donc

$$E_\lambda = \operatorname{Ker}\left(a - \lambda\mathcal{I}\right) \subsetneq \chi_\lambda = \operatorname{Ker}\left(a - \lambda\mathcal{I}\right)^2 . \tag{5.80}$$

Nous considérons un vecteur colonne propre $\mathbf{V}_{\lambda_i}$ correspondant à une solution du système (5.76), et cherchons à compléter ce vecteur par un deuxième vecteur qui lui est linéairement indépendant pour obtenir une base du sous-espace caractéristique χ_λ, de dimension 2. On considère alors le système

$$\left(A - \lambda\mathbb{I}\right)^2 \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} . \tag{5.81}$$

Évidemment toute solution du système (5.76) est aussi solution de celui-ci, ce qui s'applique en particulier au vecteur colonne déjà choisi $\mathbf{V}_\lambda$. On choisit alors une solution $\mathbf{W}_\lambda$ de (5.81), linéairement indépendante de $\mathbf{V}_\lambda$, pour compléter la base de χ_λ et ainsi pouvoir mettre le bloc A_i sous la forme

$$A_\lambda = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix} \tag{5.82}$$

où α est un nombre quelconque, distinct de zéro par hypothèse, qui se comprend en examinant l'action de l'endomorphisme sur le vecteur w_λ associé à $\mathbf{W}_\lambda$, de la forme :

$$a(w_\lambda) = \underline{\alpha} v_\lambda + \lambda w_\lambda \tag{5.83}$$

étant donné que la première colonne de (5.82) correspond à la décomposition de $a(v_\lambda)$ dans la base (v_λ, w_λ) du sous-espace χ_λ, et la deuxième colonne à la décomposition de $a(w_\lambda)$.

Pour obtenir une forme plus canonique de ce bloc, il est possible d'obtenir toute valeur non-nulle de α par un choix approprié de $\mathbf{W}_\lambda$. Le choix standard correspond à :

$$A_\lambda = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} , \tag{5.84}$$

provenant de l'équation

$$a(w_\lambda) = v_\lambda + \lambda w_\lambda \tag{5.85}$$

Sous forme matricielle, une fois déterminé le vecteur colonne propre $\mathbf{V}_\lambda$, on cherche alors un vecteur colonne $\mathbf{W}_\lambda$ solution du système :

$$\left(A - \lambda \mathbb{I}\right) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \mathbf{V}_\lambda \,. \tag{5.86}$$

Il est utile de faire deux remarques à ce stade.

1. Contrairement aux systèmes rencontrés jusqu'ici dans le cadre de la diagonalisation ou trigionalisation, ce système n'est pas homogène car il possède un membre de droite correspondant au vecteur propre colonne $\mathbf{V}_\lambda$ choisi.
2. Toute solution du système (5.86) est également solution du système (5.81). En effet si on multiplie à gauche l'éqn. (5.86) par $(A - \lambda_i \mathbb{I})$, on obtient

$$\left(A - \lambda \mathbb{I}\right)^2 \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \left(A - \lambda \mathbb{I}\right) \mathbf{V}_\lambda = 0 \,,$$

car $\mathbf{V}_\lambda$ est vecteur propre. Cela confirme que la contrainte (5.85) imposée au vecteur w_λ induit que $w_\lambda \in \chi_\lambda$ comme demandé.

Exemple 6 Reprenons la matrice déjà étudiée à l'exemple 5 :

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 4 & -2 \\ 1 & 2 & 0 \end{pmatrix} \tag{5.87}$$

plus particulièrement l'étude du sous-espace caractéristique χ_2 de dimension 2. Nous avions déterminé un premier vecteur colonne propre, en étudiant le système

$$\left(A - 2\mathbb{I}\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z & = & 0 \\ x + 2y - 2z & = & 0 \\ x + 2y - 2z & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z & = & 0 \\ y - z & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{5.88}$$

dont une solution est $\mathbf{V}_2 = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$. Pour compléter la base de χ_2 tout en mettant le bloc correspondant sous forme canonique, on étudie le système :

$$\left(A - 2\mathbb{I}\right) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z & = & 0 \\ x + 2y - 2z & = & 1 \\ x + 2y - 2z & = & 1 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} x + y - z & = & 0 \\ y - z & = & 1 \\ 0 & = & 0 \end{array} \right. \tag{5.89}$$

Cet espace des solutions, qui n'est plus un espace vectoriel, contient en particulier le vecteur colonne $\mathbf{U}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Nous avons alors pour la matrice de passage et son inverse

$$V = \big(\mathbf{V}_3\,\mathbf{V}_2\,\mathbf{U}_2\big) = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \;, \quad V^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & -1 & 2 \\ 0 & 1 & -1 \end{pmatrix} \tag{5.90}$$

d'où on tire, comme prévu, que

$$V^{-1} \cdot A \cdot V = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \tag{5.91}$$

*
* *

Bloc 3x3.

Il s'agit de la situation pour laquelle le polynôme caractéristique de a peut se factoriser sous la forme

$$P_a(X) = (X-\lambda)^3 Q(X) \;, \quad Q(\lambda) \neq 0\,, \tag{5.92}$$

et donc que

$$1 \leqslant \dim E_\lambda \leqslant \dim \chi_\lambda = 3\,. \tag{5.93}$$

Donc soit $\dim E_\lambda = \dim \chi_\lambda = 3$, et le bloc est alors diagonalisable, soit $\dim E_\lambda < 3$ et le bloc est trigonalisable. Si l'endomorphisme est caractérisé initialement par une matrice A, on étudie l'espace des solutions du système homogène :

$$\big(A - \lambda \mathbb{I}\big) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} . \tag{5.94}$$

Dans le **premier cas**, cet espace de solutions est de dimension 3, et on peut choisir 3 vecteurs colonne solutions linéairement indépendants, $\mathbf{U}_\lambda$, $\mathbf{V}_\lambda$ et $\mathbf{W}_\lambda$, qui sont par construction générateurs du sous-espace propre $E_\lambda = \chi_\lambda$ de dimension 3. Le polynôme minimal est donc dans cette situation de la forme

$$\pi_a(X) = (X-\lambda)\widehat{Q}(X) \;, \quad \widehat{Q}(\lambda) \neq 0\,, \tag{5.95}$$

car $\chi_\lambda = E_\lambda = \operatorname{Ker}(a - \lambda \mathcal{I})$. Après changement de base nous aurons un bloc diagonal de la forme

$$A_\lambda = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{5.96}$$

Le **second cas** est celui où l'espace vectoriel des solutions du système (5.94) est de dimension 2, c.-à.-d. qu'il n'est possible de trouver que deux vecteurs colonne solutions linéairement indépendants, $\mathbf{U}_\lambda$ et $\mathbf{W}_\lambda$; nous avons alors $\dim \mathsf{E}_\lambda = 2$. Nous chercherons à obtenir le bloc 3×3 correspondant au sous-espace caractéristique χ_λ sous une forme canonique donnée par :

$$A_\lambda = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} . \tag{5.97}$$

Commençons par choisir un certain vecteur colonne propre $\mathbf{U}_\lambda \in \mathsf{E}_\lambda$, c.-à.-d. solution du système (5.94). On cherche alors un vecteur colonne $\mathbf{V}_\lambda$ tel que

$$\big(A - \lambda \mathbb{I}\big)\mathbf{V}_\lambda = \mathbf{U}_\lambda . \tag{5.98}$$

Remarquons que

$$\big(A - \lambda \mathbb{I}\big)^2\mathbf{V}_\lambda = \big(A - \lambda \mathbb{I}\big)\mathbf{U}_\lambda = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.99}$$

car $\mathbf{U}_\lambda$ et vecteur propre, et donc $\mathbf{V}_\lambda \in \operatorname{Ker}(\mathfrak{a} - \lambda \mathfrak{I})^2 \subset \chi_\lambda = \operatorname{Ker}(\mathfrak{a} - \lambda \mathfrak{I})^3$. Il suffit alors de choisir un autre vecteur propre $W_\lambda \in \mathsf{E}_\lambda$, afin que $\{\mathbf{U}_\lambda, \mathbf{V}_\lambda, \mathbf{W}_\lambda\}$ forment une famille libre et donc une base du sous-espace caractéristique χ_λ, dans laquelle le bloc A_λ sera de la forme souhaitée, donnée par l'éqn. (5.97)

Étant donné que $\mathbf{U}_\lambda$ et $\mathbf{V}_\lambda$ appartiennent à $\mathsf{E}_\lambda = \operatorname{Ker}(\mathfrak{a} - \lambda \mathfrak{I})$, et que $\mathbf{V}_\lambda \in \operatorname{Ker}(\mathfrak{a} - \lambda \mathfrak{I})^2$, le cas de figure examiné ici correspond à

$$\chi_\lambda = \operatorname{Ker}\big(\mathfrak{a} - \lambda \mathfrak{I}\big)^2 \quad \Leftrightarrow \quad \pi_\mathfrak{a}(X) = (X - \lambda)^2 \widehat{Q}(X) \ , \quad \widehat{Q}(\lambda) \neq 0 . \tag{5.100}$$

Remarque 2 *Crucialement, qu'il n'est pas garanti que notre choix de vecteur propre* $\mathbf{U}_\lambda$ *conduise à l'existence d'une solution* $\mathbf{V}_\lambda$ *pour l'équation (5.99) ! Il se peut en effet que, dans une base contenant* $\mathbf{U}_\lambda$*, la matrice associée à l'endomorphisme ne puisse prendre la forme (5.97) qui est diagonale par blocs.*

À la lumière de cette remarque il est préférable de présenter l'algorithme sous une forme différente, qui fonctionne à coup sûr :

1. On commence par choisir un vecteur colonne $\mathbf{V}_\lambda \in \operatorname{Ker}(\mathfrak{a} - \lambda \mathfrak{I})^2$, c.-à.-d. solution de l'équation

$$\big(A - \lambda \mathbb{I}\big)^2 \mathbf{V}_\lambda = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} , \tag{5.101}$$

avec la contrainte que

$$\mathbf{U}_\lambda \overset{\text{déf.}}{=} \big(A - \lambda \mathbb{I}\big)\mathbf{V}_\lambda \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} ; \tag{5.102}$$

on constate que $\mathbf{U}_\lambda \in \mathsf{E}_\lambda$, c.-à.-d. vecteur propre pour la valeur propre λ.

2. On cherche un autre vecteur propre $\mathbf{W}_\lambda \in \mathsf{E}_\lambda$ tel que la famille $\{\mathbf{U}_\lambda, \mathbf{V}_\lambda, \mathbf{W}_\lambda\}$ soit une famille libre.

Le **troisième cas** est celui où l'espace vectoriel des solutions du système (5.94) est de dimension 1, c.-à.-d. qu'il n'est possible de trouver qu'un vecteur colonne $\mathbf{U}_\lambda$ solution linéairement indépendant. Il s'agit de la situation pour laquelle

$$\chi_\lambda = \operatorname{Ker}\left(a - \lambda\mathcal{I}\right)^3 \quad \Leftrightarrow \quad \pi_a(X) = (X-\lambda)^3\widehat{Q}(X)\ , \quad \widehat{Q}(\lambda) \neq 0\,. \tag{5.103}$$

Pour trigonaliser le bloc 3×3 correspondant nous chercherons à former une base de χ_λ avec $\mathbf{U}_\lambda$ et deux autres vecteurs colonnes $\mathbf{V}_\lambda$ et $\mathbf{W}_\lambda$ afin d'obtenir le bloc 3×3 sous une forme canonique donnée par :

$$A_\lambda = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}\ , \tag{5.104}$$

correspondant aux trois équations :

$$\left(a - \lambda\mathcal{I}\right)(u_\lambda) = 0_E \tag{5.105a}$$
$$\left(a - \lambda\mathcal{I}\right)(v_\lambda) = u_\lambda \tag{5.105b}$$
$$\left(a - \lambda\mathcal{I}\right)(w_\lambda) = v_\lambda \tag{5.105c}$$

Des vecteurs obéissant à ces équations forment nécessairement une famille libre. Supposons qu'il existe α, β, γ tels que

$$\alpha u_\lambda + \beta v_\lambda + \gamma w_\lambda = 0_E \tag{5.106}$$

En appliquant $(a-\lambda\mathcal{I})^2$ à cette équation on trouve $\gamma = 0$, puis en lui appliquant $(a-\lambda\mathcal{I})$ on trouve $\beta = 0$ et enfin $\alpha = 0$.

L'algorithme pour obtenir la base de χ_λ souhaitée se présente alors sous une forme proche du cas précédent. On commence par choisir un vecteur colonne $\mathbf{W}_\lambda \in \operatorname{Ker}(a-\lambda\mathcal{I})^3$, c.-à.-d. solution de l'équation

$$\left(A - \lambda\mathbb{I}\right)^3 \mathbf{W}_\lambda = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\ , \tag{5.107}$$

avec la contrainte que

$$\mathbf{U}_\lambda \overset{\text{déf.}}{=} \left(A - \lambda\mathbb{I}\right)^2 \mathbf{W}_\lambda \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\,. \tag{5.108}$$

Nous constatons alors que

$$\left(A - \lambda\mathbb{I}\right)\mathbf{U}_\lambda = \left(A - \lambda\mathbb{I}\right)^3 \mathbf{W}_\lambda = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\ , \tag{5.109}$$

donc $\mathbf{U}_\lambda$ est vecteur colonne propre associé à la valeur propre λ. Il suffit alors de définir le vecteur colonne

$$\mathbf{V}_\lambda \stackrel{\text{déf.}}{=} \big(A - \lambda\mathbb{I}\big)\mathbf{W}_\lambda\,, \tag{5.110}$$

par construction non-nul, pour que la famille libre $\{\mathbf{U}_\lambda, \mathbf{V}_\lambda, \mathbf{W}_\lambda\}$ soit une base du sous-espace caractéristique χ_λ tel que le bloc 3×3 associé soit de la forme (5.104) souhaitée.

Exemple 7 Soit b un endomorphisme sur un $\mathbb{R}$-espace vectoriel de dimension 3 représenté dans une certaine base par la matrice

$$B = \begin{pmatrix} 3 & -2 & -1 \\ -1 & 3 & 1 \\ 2 & -3 & 0 \end{pmatrix} \tag{5.111}$$

On calcule son polynôme caractéristique,

$$P_B(X) = -(X-2)^3 \tag{5.112}$$

donc B, qui n'est manifestement par proportionnelle à la matrice identité, n'est pas diagonalisable. Les vecteurs propres sont solutions du système

$$\big(B - 2\mathbb{I}\big)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{\begin{array}{rcl} x - 2y - z & = & 0 \\ -x + y + z & = & 0 \\ 2x - 3y - 2z & = & 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{rcl} x - 2y - z & = & 0 \\ y & = & 0 \\ 0 & = & 0 \end{array}\right. \tag{5.113}$$

L'espace des solutions est donc de dimension 1 ; nous somme dans le troisième cas de figure ci-dessus, tel que $\chi_2 = \operatorname{Ker}(b - \lambda\mathcal{I})^3$. Remarquons en passant que, dans ce cas précis, comme nous sommes en dimension 3 et que $\dim \chi_2 = 3$, tout vecteur appartient à χ_2, soit $E = \chi_2$. Choisissons-donc un vecteur colonne $\mathbf{W}_2$ arbitraire tel que

$$\big(B - 2\mathbb{I}\big)^2\mathbf{W}_2 = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow x - y - z \neq 0\,. \tag{5.114}$$

On choisit par exemple $\mathbf{W}_2 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$. Nous avons alors

$$\mathbf{U}_2 = \big(B - 2\mathbb{I}\big)^2\mathbf{W}_2 = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{5.115}$$

ainsi que

$$\mathbf{V}_2 = \big(B - 2\mathbb{I}\big)\mathbf{W}_2 = \begin{pmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 2 & -3 & -2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \tag{5.116}$$

On forme alors la matrice de passage

$$V = \left(\mathbf{U}_2\,\mathbf{V}_2\,\mathbf{W}_2\right) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & -2 & 0 \end{pmatrix} \tag{5.117}$$

et on vérifie que

$$\begin{aligned} V^{-1} \cdot B \cdot V &= \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -2 & -1 \\ -1 & 3 & 1 \\ 2 & -3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & -2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned} \tag{5.118}$$

comme attendu.

Exemple 8 Soit un endomorphisme sur un $\mathbb{R}$-espace vectoriel de dimension 3 représenté dans une certaine base par la matrice

$$C = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 0 \\ 1 & -1 & 4 \end{pmatrix} \tag{5.119}$$

On calcule son polynôme caractéristique,

$$P_C(X) = -(X-3)^3 \tag{5.120}$$

donc C, qui n'est manifestement par proportionnelle à la matrice identité, n'est pas diagonalisable. Les vecteurs propres sont solutions du système

$$\left(C - 3\mathbb{I}\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{ \begin{array}{rcl} -x+y-z & = & 0 \\ 0 & = & 0 \\ x-y+z & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} x-y+z & = & 0 \\ 0 & = & 0 \\ 0 & = & 0 \end{array} \right. \tag{5.121}$$

L'espace des solutions est donc de dimension 2 ; nous somme dans le second cas de figure ci-dessus, tel que $\chi_2 = \operatorname{Ker}(c - \lambda\mathcal{I})^2$. Remarquons que, ici aussi, comme nous sommes en dimension 3 et que $\dim \chi_2 = 3$, tout vecteur appartient à χ_2, soit $E = \chi_2$.On cherche donc premièrement un vecteur colonne $\mathbf{V}_3$ arbitraire tel que $(C - 3\mathbb{I})\mathbf{V}_3$ soit non-nul, c.-à.-d.

$$\left(C - 3\mathbb{I}\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow -x+y-z \neq 0 \tag{5.122}$$

Choisissons par exemple $\mathbf{V}_3 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$. Nous avons ensuite

$$\mathbf{U}_3 = \left(C - 3\mathbb{I}\right)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \tag{5.123}$$

qui est par construction un vecteur propre de C. On complète alors la base par un vecteur propre de C linéairement indépendant de $\mathbf{V}_3$, par exemple

$$\mathbf{W}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \tag{5.124}$$

et on forme la matrice de passage

$$\mathsf{V} = \big(\mathbf{U}_3\,\mathbf{V}_3\,\mathbf{W}_3\big) = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{5.125}$$

et on vérifie que

$$\begin{aligned} \mathsf{V}^{-1} \cdot \mathsf{C} \cdot \mathsf{V} &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 0 \\ 1 & -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \end{aligned} \tag{5.126}$$

comme attendu.

*
* *

La généralisation de cette méthode de trigonalisation en dimension quelconque porte le nom de *reduction de Jordan*; nous nous contenterons de l'esquisser ici.

Définition 7 (bloc de Jordan) *Un bloc de Jordan de dimension* k *pour la valeur propre* λ*, noté* $\mathsf{J}_\mathsf{k}(\lambda)$*, est une matrice triangulaire* $\mathsf{k} \times \mathsf{k}$ *de la forme*

$$\mathsf{J}_\mathsf{k}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix} \tag{5.127}$$

Propriété 7 *Tout bloc de Jordan se décompose comme*

$$\mathsf{J}_\mathsf{k}(\lambda) = \lambda \mathbb{I}_\mathsf{k} + \mathsf{N}_\mathsf{k}\,, \tag{5.128}$$

où N_k *est une matrice nilpotente de degré* k*, c.-à.-d. telle que* $\big(\mathsf{N}_\mathsf{k}\big)^\mathsf{k} = \mathbb{O}_\mathsf{k}$.

Il suffit d'écrire

$$J_k(\lambda) = \underbrace{\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix}}_{\lambda \mathbb{I}_k} + \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}}_{N_k} \tag{5.129}$$

puis de calculer le polynôme caractéristique de la matrice N_k :

$$P_{N_k}(X) = \begin{vmatrix} -X & 1 & 0 & \cdots & 0 \\ 0 & -X & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & -X \end{vmatrix} = (-1)^k X^k . \tag{5.130}$$

Le théorème de Cayley–Hamilton nous indique alors que

$$P_{N_K}(N_K) = 0 \implies \left(N_k\right)^k = \mathbb{O}_k . \tag{5.131}$$

Reste à se convaincre que $\left(N_k\right)^{k-1} \neq \mathbb{O}_k$. Considérons le vecteur colonne

$$\mathbf{U}_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{5.132}$$

et on constate que

$$U_{k-1} \overset{\text{déf.}}{=} N_k \mathbf{U}_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{5.133}$$

puis

$$U_{k-2} \overset{\text{déf.}}{=} \left(N_k\right)^2 \mathbf{U}_k = N_k \mathbf{U}_{k-1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{5.134}$$

et ainsi de suite jusqu'à atteindre

$$U_1 \overset{\text{déf.}}{=} \left(N_k\right)^{k-1} \mathbf{U}_k = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.135}$$

Pour comprendre cette propriété différemment cherchons les vecteurs propres de N_k, pour l'unique valeur propre qui est égale à zéro car $P_{N_k}(X) = (-1)^k X^k$. Nous avons

$$N_k \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow x^2 = x^3 = \cdots = x^n = 0\,. \tag{5.136}$$

Le sous-espace propre E_0 est de dimension 1, engendré par le vecteur colonne $\mathbf{U}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Pour construire une base du sous-espace caractéristique χ_0, qui est de dimension k, on part d'un vecteur colonne $\mathbf{U}_k$ tel que

$$\left(N_k\right)^{k-1} \mathbf{U}_k \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.137}$$

et on construit les autres vecteurs de base par applications successives de N_k, comme ci-dessus. □

Théorème 7 (décomposition de Jordan) *Soit E un K-espace vectoriel de dimension n et $a : E \to E$ un endormophisme de polynôme caractéristique scindé. Il existe alors une base $\mathcal{B}$ de E telle que la matrice associée à l'endomorphisme A soit de la forme dite de Jordan :*

$$J = \begin{pmatrix} J_1(\lambda_1) & 0 & \cdots & \cdots & 0 \\ 0 & J_2(\lambda_2) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & J_p(\lambda_p) \end{pmatrix}, \tag{5.138}$$

dans laquelle les valeurs propres λ_i apparaissant dans les des différents blocs de Jordan ne sont pas nécessairement distinctes.

Nous ne donnerons pas la preuve de ce théorème, généralisant les cas étudiés précédemment concernant les blocs de Jordan jusqu'à la taille 3×3. On notera que, pour une valeur propre λ donnée, il existe autant de blocs de Jordan dans la décomposition de la matrice que de vecteurs propres linéairement indépendants associés à cette valeur propre.

Exemple 9 Reprenons les matrices étudiées dans les exemples 6 à 8. Nous avons trouvé premièrement que

$$V^{-1} \cdot A \cdot V = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} J_1(3) & 0 \\ 0 & J_2(2) \end{pmatrix}, \tag{5.139}$$

puis

$$V^{-1} \cdot B \cdot V = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = J_3(2) \tag{5.140}$$

et enfin

$$V^{-1} \cdot C \cdot V = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} J_2(3) & \begin{matrix} 0 \\ 0 \end{matrix} \\ 0 \quad 0 & J_1(3) \end{pmatrix} \tag{5.141}$$

Remarquons, sur ce dernier exemple, la présence de deux blocs de Jordan associés à la même valeur propre $\lambda = 3$.

À la lumière de nos études précédentes il serait possible de proposer une méthode générale pour obtenir la réduction de Jordan d'un endomorphisme a quelconque de polynôme caractéristique scindé :

1. On calcule le polynôme caractéristique de a, $P_a = (-1)^n (X - \lambda_1)^{c_1} \cdots (X - \lambda_n)^{c_n}$.
2. Pour chaque valeur propre λ, on détermine la dimension du sous-espace propre $E_\lambda = \mathrm{Ker}\,(a - \lambda \mathfrak{I})$, qui donne le nombre de blocs de Jordan associés à λ, en étudiant l'espace des solutions du système
$$\big(A - \lambda \mathbb{I}_n\big) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.142}$$
On choisit alors une base de solutions associés aux vecteurs colonne $\{\mathbf{U}_1, \mathbf{V}_1, \ldots\}$.
3. Partant du premier vecteur propre de la base $\mathbf{U}_1$, on cherche s'il existe un vecteur colonne $\mathbf{U}_2$ tel que
$$\big(A - \lambda \mathbb{I}\big) \mathbf{U}_2 = \mathbf{U}_1 \tag{5.143}$$
4. si un tel vecteur existe, on cherche alors un vecteur $\mathbf{U}_3$ tel que
$$\big(A - \lambda \mathbb{I}\big) \mathbf{U}_3 = \mathbf{U}_2\,, \tag{5.144}$$
et ainsi de suite.
5. le processus s'arrête[3] avec un certain vecteur $\mathbf{U}_\ell$ lorsqu'il n'existe pas de solution à

3. Dans tous les cas la taille du bloc de Jordan est inférieure à la dimension du sous-espace caractéristique χ_λ, donnée par la multiplicité de la racine λ dans le polynôme caractéristique $P_a(X)$.

l'itération suivante $\big(A - \lambda\mathbb{I}\big)\mathbf{U}_{\ell+1} = \mathbf{U}_\ell$. On a alors

$$\begin{aligned} A\mathbf{U}_1 &= \lambda\mathbf{U}_1 \\ A\mathbf{U}_2 &= \mathbf{U}_1 + \lambda\mathbf{U}_2 \\ \vdots\quad & \quad\vdots \\ A\mathbf{U}_\ell &= \mathbf{U}_{\ell-1} + \lambda\mathbf{U}_\ell \end{aligned}$$

engendrant le bloc de Jordan $J_\ell(\lambda)$ de taille $\ell \times \ell$.

6. On repète le même processus aveec le deuxième vecteur propre $\mathbf{V}_1$ de la base choisie pour le sous-espace propre E_λ, et ainsi de suite pour les autres vecteurs de base.
7. On recommence pour les autres valeurs propres.

Cependant, comme nous l'avons noté plus haut, voir la remarque 2, cet algorithme ne conduit pas nécessairement à une solution du problème, si le vecteur propre initial est mal choisi. Pour être certain d'aboutir, il faut généraliser aux valeurs propres d'ordre arbitraire la méthode exposée après la remarque 2 pour les valeur propres d'ordre 3. *Dans le cadre de ce cours, nous nous contenterons de savoir réduire les endomorphismes ayant au plus des valeurs propres d'ordre 3.*

Exemple 10 Soit a un endomorphisme sur un $\mathbb{R}$-espace vectoriel de dimension 5 représenté dans une certaine base par la matrice

$$A = \begin{pmatrix} 5 & 0 & 4 & -2 & -3 \\ -2 & 3 & -3 & 2 & 4 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 1 & 0 & 2 & -1 & 1 \end{pmatrix} \tag{5.145}$$

Le polynôme caractéristique associé est $P_a(X) = -(X-3)^5$. Cherchons une base de l'unique sous-espace propre E_3, en résolvant le système

$$\big(A - 3\mathbb{I}_3\big)\begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{5.146}$$

dont on trouve deux solutions linéairement indépendantes, par exemple

$$\mathbf{U}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \mathbf{V}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{5.147}$$

La réduction de Jordan de la matrice A fera donc apparaître deux blocs.

Pour obtenir le bloc associé au vecteur colonne propre $\mathbf{U}_1$ on cherche un vecteur $\mathbf{U}_2$ tel que

$$\big(A - 3\mathbb{I}_3\big)\mathbf{U}_2 = \mathbf{U}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.148}$$

dont une solution est

$$\mathbf{U}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{5.149}$$

On vérifie ensuite que le système $\big(A-3\mathbb{I}_3\big)\mathbf{U}_3 = \mathbf{U}_2$ n'admet pas de vecteur $\mathbf{U}_3$ solution, donc le processus s'arrête et on a obtenu un premier bloc de Jordan $J_2(3)$.

Pour obtenir le bloc associé au vecteur colonne propre $\mathbf{V}_1$ on cherche un vecteur $\mathbf{V}_2$ tel que

$$\big(A - 3\mathbb{I}_3\big)\mathbf{V}_2 = \mathbf{V}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{5.150}$$

dont une solution est

$$\mathbf{V}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{5.151}$$

On cherche ensuite un vecteur $\mathbf{V}_3$ tel que

$$\big(A - 3\mathbb{I}_3\big)\mathbf{V}_3 = \mathbf{V}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{5.152}$$

dont une solution est

$$\mathbf{V}_3 = \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \tag{5.153}$$

Le processus s'arrête nécessairement là, car nous avons obtenu assez de vecteurs pour générer l'espace vectoriel de dimension 5. En utilisant la matrice de passage

$$V = \left(\mathbf{U}_1\,\mathbf{U}_2\,\mathbf{V}_1\,\mathbf{V}_2\,\mathbf{V}_3\right), \tag{5.154}$$

on vérifie que

$$V^{-1}\cdot A\cdot V = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} J_2(3) & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & J_3(3) \end{pmatrix} \tag{5.155}$$

donnant la réduction de Jordan demandée.

5.4 Applications

Les applications du processus de trigonalisation sont identiques à celles de la diagonalisation, à savoir les calculs de puissances et d'exponentielles d'endomorphismes, et l'étude de systèmes d'équations différentielles linéaires couplées.

Évoquons tout d'abord brièvement le calcul des puissances d'un endomorphisme qui n'est pas diagonalisable. La première méthode disponible est d'utiliser le théorème de Cayley–Hamilton, qui indique que, si P_a est le polynôme caractéristique de a, alors

$$P_a(a) = (-1)^n a^n + \cdots = 0\,. \tag{5.156}$$

On peut ainsi exprimer a^n en termes de puissances inférieures de cet endomorphisme.

Exemple 11 Reprenons l'exemple 10, où nous avons étudié l'endomorphisme associé à la matrice :

$$A = \begin{pmatrix} 5 & 0 & 4 & -2 & -3 \\ -2 & 3 & -3 & 2 & 4 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 1 & 0 & 2 & -1 & 1 \end{pmatrix} \tag{5.157}$$

Son polynôme caractéristique est

$$P_A(X) = -(X-3)^5 = -X^5 + 15X^4 - 90X^3 + 270X^2 - 405X + 243\,, \tag{5.158}$$

D'où nous tirons que

$$A^5 = 15A^4 - 90A^3 + 270A^2 - 405A + 243\mathbb{I}_5\,. \tag{5.159}$$

Nous pouvons aussi en déduire l'inverse de A, en réécrivant cette équation comme

$$A(A^4-15A^3+90A^2-270A+405\mathbb{I}) = (A^4-15A^3+90A^2-270A+405\mathbb{I})A = 243\mathbb{I}_5\,. \tag{5.160}$$

d'où

$$A^{-1} = \frac{1}{243}\left(A^4 - 15A^3 + 90A^2 - 270A + 405\mathbb{I}\right). \tag{5.161}$$

*
* *

Plus généralement, considérons une la réduction de Jordan d'une matrice A trigonalisable. Il existe une matrice de passage V telle que

$$A = V \cdot \underbrace{\begin{pmatrix} J_1(\lambda_1) & 0 & \cdots & \cdots & 0 \\ 0 & J_2(\lambda_2) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & J_p(\lambda_p) \end{pmatrix}}_{J} \cdot V^{-1} \tag{5.162}$$

Le calcul des puissances de A se ramène au calcul des puissances de J, étant donné que

$$A^k = (V \cdot J \cdot V^{-1}) \cdot (V \cdot J \cdot V^{-1}) \cdots (V \cdot J \cdot V^{-1}) = V \cdot J^k \cdot V^{-1}. \tag{5.163}$$

Nous remarquons ensuite que, la matrice J étant diagonale par blocs,

$$J^k = \begin{pmatrix} J_1(\lambda_1)^k & 0 & \cdots & \cdots & 0 \\ 0 & J_2(\lambda_2)^k & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & J_p(\lambda_p)^k \end{pmatrix}. \tag{5.164}$$

Le problème se ramène donc au calcul de la puissance d'un unique bloc de Jordan. Rappelons, d'après la propriété 7, qu'un bloc de Jordan $J_m(\lambda)$ peut se décomposer comme

$$J_m(\lambda) = \underbrace{\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix}}_{\lambda\mathbb{I}_m} + \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}}_{N_m} \tag{5.165}$$

où N_m est nilpotente de degré m, c.-à.-d. que $(N_m)^{m-1} \neq \mathbb{O}$ et $(N_m)^m = \mathbb{O}_m$.

Étant donné que la matrice identité commute avec toutes les autres matrices carrée de même taille, nous pouvons alors utiliser la formule du binôme de Newton :

$$\left(J_m(\lambda)\right)^p = \left(\lambda\mathbb{I}_m + N_m\right)^p = \sum_{\ell=0}^{p} \binom{p}{l} \lambda^{p-\ell}\left(N_m\right)^{\ell} \tag{5.166}$$

où, rappelons-le, $(N_m)^0 = \mathbb{I}_k$ par convention. Si $p \geqslant m$ la somme est tronquée étant donné que $(N_m)^\ell = 0$ si $\ell \geqslant m$.

*
* *

Une fois connue l'expression générale de la puissance d'un bloc de Jordan, on peut également calculer l'exponentielle d'une matrice trigonalisable via sa réduction de Jordan :

$$\exp A = V \cdot \exp J \cdot V^{-1} = V \cdot \begin{pmatrix} \exp J_1(\lambda_1) & 0 & \cdots & \cdots & 0 \\ 0 & \exp J_2(\lambda_2) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \exp J_p(\lambda_p) \end{pmatrix} \cdot V^{-1} \tag{5.167}$$

où

$$\exp J_m(\lambda) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \big(J_m(\lambda)\big)^\ell . \tag{5.168}$$

Exemple 12 Reprenons la matrice étudiée à l'exemple 7,

$$B = \begin{pmatrix} 3 & -2 & -1 \\ -1 & 3 & 1 \\ 2 & -3 & 0 \end{pmatrix} \tag{5.169}$$

Nous avions trouvé une matrice de passage

$$V = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & -2 & 0 \end{pmatrix} \tag{5.170}$$

telle que

$$V^{-1} \cdot B \cdot V = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = J_3(2) \tag{5.171}$$

Pour calculer les puissances de B, nous utilisons donc

$$B^p = V \cdot J_3(2)^p \cdot V^{-1} \tag{5.172}$$

et nous avons, pour tout $p \geqslant 2$,

$$J_3(2)^p = \big(2\mathbb{I}_3 + N_3\big)^p = 2^p + p2^{p-1}N_2 + \frac{p(p-1)}{2}2^{p-2}N_2^2 = \begin{pmatrix} 2^p & 2^{p-1}p & 2^{p-3}(p-1)p \\ 0 & 2^p & 2^{p-1}p \\ 0 & 0 & 2^p \end{pmatrix} \tag{5.173}$$

car $N_2^3 = \mathbb{O}$. Nous avons donc

$$B^p = V \cdot J_3(2)^p \cdot V^{-1} = \begin{pmatrix} 2^{p-3}\left(p^2+3p+8\right) & -2^{p-3}p(p+7) & -2^{p-3}p(p+3) \\ -2^{p-1}p & 2^{p-1}(p+2) & 2^{p-1}p \\ 2^{p-3}p(p+7) & -2^{p-3}p(p+11) & -2^{p-3}\left(p^2+7p-8\right) \end{pmatrix}. \tag{5.174}$$

On peut aussi vérifier que

$$\exp(tB) = V\cdot\exp\big(tJ_3(2)\big)\cdot V^{-1} = \begin{pmatrix} e^{2t}\left(\frac{t^2}{2}+t+1\right) & e^{2t}\left(-\frac{t^2}{2}-2t\right) & e^{2t}\left(-\frac{t^2}{2}-t\right) \\ -e^{2t}t & e^{2t}(t+1) & e^{2t}t \\ e^{2t}\left(\frac{t^2}{2}+2t\right) & e^{2t}\left(-\frac{t^2}{2}-3t\right) & e^{2t}\left(-\frac{t^2}{2}-2t+1\right) \end{pmatrix} \tag{5.175}$$

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.